DOCTORAL THESIS

A Generic Framework for Collective Intelligence Systems

Shweta Suran

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2022

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
15/2022

A Generic Framework for Collective Intelligence Systems

SHWETA SURAN

TALTECH PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Department of Software Science

The dissertation was accepted for the defence of the degree of Doctor of Philosophy (Computer Science) on 14 April 2022

Supervisor: Prof. Dr. Dirk Draheim,
Information Systems Group
Department of Software Science, School of Information Technologies,
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Assoc.-Prof. Dr. Alexander Horst Norta,
Department of Software Science, School of Information Technologies,
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Assoc.-Prof. Dr. Ingrid Pappel,
Department of Software Science, School of Information Technologies,
Tallinn University of Technology
Tallinn, Estonia

Opponents: Prof. Dr. Gerhard Schwabe,
Department of Informatics,
University of Zurich
Zurich, Switzerland

Prof. Dr. Anna De Liddo,
Knowledge Media Institute,
The Open University
Milton Keynes, United Kingdom

Defence of the thesis: 26 May 2022, Tallinn

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.

Shweta Suran ______________________
signature

ISSN 2585-6898 (publication)
ISBN 978-9949-83-817-2 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-818-9 (PDF)
Printed by Koopia Niini & Rauam

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
15/2022

Kollektiivse intelligentsuse süsteemide üldine raamistik

SHWETA SURAN

TALTECH KIRJASTUS

Contents

List of Publications

The present Ph.D. thesis is based on the following publications that are referred to in the text by Roman numbers.

I S. Suran, V. Pattanaik, S. B. Yahia, and D. Draheim. Exploratory analysis of collective intelligence projects developed within the EU Horizon 2020 framework. In N. T. Nguyen, R. Chbeir, E. Exposito, P. Aniorté, and B. Trawiński, editors, *Proceedings of ICCCI 2019 – the 11th International Conference on Computational Collective Intelligence*, Lecture Notes in Artificial Intelligence 11684, pages 285–296. Springer, 2019

II V. Pattanaik, S. Suran, and D. Draheim. Enabling social information exchange via dynamically robust annotations. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2019 – the 21st International Conference on Information Integration and Web-based Applications and Services*, pages 176–184. ACM, 2019

III S. Suran, V. Pattanaik, and D. Draheim. Frameworks for collective intelligence: A systematic literature review. *ACM Computing Surveys*, 53(1):1–36, 2020

IV S. Suran, V. Pattanaik, and D. Draheim. CommunityCare: Tackling mental health issues with the help of community. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2020 – the 22nd International Conference on Information Integration and Web-based Applications and Services*, pages 377–382. ACM, 2020

V S. A. Peious, S. Suran, V. Pattanaik, and D. Draheim. Enabling sensemaking and trust in communities. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2021 – the 23rd International Conference on Information Integration and Web-based Applications and Services*, pages 1–9. ACM, 2021

Author's Contributions to the Publications

I I was the main author of this publication. I conducted an analysis of 10 collective intelligence (CI) projects and analyzed the results, prepared the figures, and wrote the manuscript.

II In this publication, I was a co-author. I wrote about the role of the Web and CI in today's digital world, prepared the figures, and revised the manuscript.

III[1–4] I was the main author of this publication. I conducted the systematic literature review (SLR) of 9,418 research articles, contributed the generalized framework and conducted the comparative case studies of 6 CI platforms, analyzed the results, prepared the figures and tables, and wrote the manuscript.

IV I was the main author of this publication. I proposed a web-based (CI) platform (i.e., CommunityCare), as a use case, prepared the figures and tables, and wrote the manuscript.

V In this publication, I was the second author. I formulated the research questions, wrote about the role CI in business intelligence (BI), prepared the figures, wrote Section 2 and Section 3 and revised the manuscript.

[1] As of 29th Oct. 2021, the article has received 41 citations (Google scholar) resp. 17 citations (Scopus).

[2] The article has been listed in The Living Library [2] of The GovLab [1].

[3] The developed CI framework is currently used by the Information Systems Group of Tallinn University of Technology to design a collaborative platform for the silver economy based on endeavors of Butt et al. [22, 23, 21] in the EU Interreg project "Supporting Smart Specialization Approach in Silver Economy for Increasing Regional Innovation Capacity and Sustainable Growth"[7].

[4] The developed CI framework will also be used in an urban planning project of Chaves et al. [27], where citizens will be engaged in developing and mapping new ideas and discussing their points of view to build future societies.

Abbreviations

BI	business intelligence
CI	collective intelligence
DS	design science
HCI	human-computer interaction
ICT	information and communications technology
IS	information systems
RQ	research question
SLR	systematic literature review

Terms

collective (adjective)	done or shared by all members of a group of people; involving a whole group or society. (Oxford Learner's Dictionaries [6])
collective (noun)	a singular noun, such as committee or team, that refers to a group of people, animals or things and, in British English, can be used with either a singular or a plural verb. In American English it must be used with a singular verb. (Oxford Learner's Dictionaries [6])
collective intelligence	1) "a form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skill" [61] 2) "groups of individuals acting collectively in ways that seem intelligent" [68]
intelligence	1) the ability to learn, understand and think in a logical way about things; the ability to do this well (Oxford Learner's Dictionaries [6]) 2) a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings – "catching on," "making sense" of things, or "figuring out" what to do. [39]

Summary

This research summary is structured as follows. In Sect. 1, we establish the motivation, provide the problem statement behind this work, and state the addressed research questions (RQs). In Sect. 2, we provide an overview of the methodologies used in this work. In Sect. 3, we state the contributions made in this work. In Sects. 4, 5 and 6, we explain how the publications [I], [II], [III], [IV] and [V] contribute to the work, i.e., the development and evaluation of a generic framework for CI systems in information and communications technology (ICT). In Sect. 7, we discuss the related work studied in the publications [I], [III] and [IV]. In Sect. 8, we present directions for future research. We conclude the research summary in Sect. 9.

1 Motivation and Problem Statement

1.1 Research Relevance

Collective intelligence (CI) is not a new concept and has been in the focus of research such as computer science, management, organizations, social science, biology, and psychology for many years [III]. CI can be found in several forms on the web, and looking at the CI literature, web-based CI applications can be mainly grouped into three distinct paradigms.

First, it has become easy to connect with other people around the world; we are now capable to learn about other cultures, history, medicine, art and more, thereby, improving our beliefs and knowledge about each other [65, 74]. Through CI platforms like InnoCentive and OpenIDEO, people from different countries can support, promote and contribute to each other's ideas and initiatives, and can offer not only economic support but also mental support in generating innovative ideas [65, 74].

Second, organizations are now able to harness the knowledge of web users, e.g., to gain insights about market trends, user requirements and perceptions, or to predict upcoming changes in the consumer markets [V]. Within organizations, the use of CI techniques such as crowdsourcing has the potential to enable knowledge creation and exchange [14]. Stafford Beer's viable systems theory [12, 13] explains the critical functions of any viable organization. At the next level, the theory has been used in visions to shape organizational ecosystems [11], a notion that is currently discussed as *real-time economics* [112, 111]. The integration of CI with the critical functions of viable organizations shows an immense potential [32, 31, 30, 70]. This potential has been understood by major players such as IBM with its vision of a *Cogniculture* [87] and BlackRock (world largest shadow bank with more than 7.3 trillion USD assets under management) with its CI platform Alladin [8]. Both Cogniculture and Alladin are examples of how today's established CI [65, 74] can be transcended (taken to a next level) by *organizational "human-machine co-evolution" [87]*, compare with Bruno Latour's *actor-network theory* (ANT) [59].

Finally, the overall rise of the social web (i.e., social media) and CI has encouraged government bodies and public organizations to empower citizens with more decision-making capabilities and power in policy making and the implementation of new laws; by understanding citizens' opinions about *what* changes they want and *why*. Thanks to platforms such as Decidim.Barcelona [9] and CitizenLab [86], crowds are now able to contribute more actively in their own governance and can even guide their governments in the development of new guidelines and policies. Taking into consideration the aforementioned aspects, it has become visible that CI plays an essential role in our society; unfortunately, the design and development of CI platforms is still a time-consuming task, e.g., need months for requirement analysis and planning alone [III]. Therefore, it is hard for small organisations and governing bodies to design such platforms. Our research aims at providing a ‘generic’ CI framework that allows stakeholders to easily merge distinct components needed to enable CI, this way allowing them to develop their own CI platforms more effectively and efficiently. Utilizing the results of our research, it becomes simpler for researchers, developers and other stakeholders to build new CI platforms while cutting the amount of money and time needed [III]. Furthermore, the findings of our research enable researchers who are new to the field to understand faster, how CI is enabled in the many different kind of systems (based on ICT) that exist today.

1.2 Main Challenges

There are three mutually dependent main challenges currently being studied in the field of CI system development:

- Existing CI frameworks suffer a specific reproducibility crisis.
- Existing CI models are domain-specific.
- It lacks a generic CI framework.

These three challenges are described in Sects. 1.2.1, 1.2.2 and 1.2.3.

1.2.1 Reproducibility Crisis The concept of CI [61, 66], popularly known as “wisdom of crowds” [110] has been a part of scientific discussions for centuries, actually, ever since the Aristotelian era [74]. What started as a group of individuals arguing over collective issues and policies, over time has co-evolved through various phases in concert with the novel communication means of the different eras (e.g., printing press, sound telegraph, ARPANET). And this indeed, has allowed the concept to transcend into a plethora of scientific strands such as human-computer interaction (HCI) [III], citizen science [65], swarm intelligence [65, 74], collaborative AI [87] and open innovation [65, 74]. Examples of CI systems that started in the early 1990s include Goldcorp (which aimed to identify new mining locations, using the wisdom of web users) [65] and WikiWikiWeb (the first Wiki, that enabled exchange of ideas) [65]. Similarly, the Linux software community [90] is an early CI example of artifact and knowledge creation [65]. Since then, scientific advancement in ICT technologies such as the social web [91, 45], have led to the development of numerous novel innovative CI systems; to name a few, Wikipedia (for knowledge exchange) [65], Climate CoLab (to tackle climate issues) [49], Reddit (to share passions, ideas, and opinions) [115], Kaggle (to propose new machine learning solutions) [I], and StackOverflow (to propose new coding questions and solutions) [99]. The social web (i.e., social media) as a predominant means for communication today, has enabled web users to interact and collaborate with each other in unprecedented ways, empowering them by allowing effective and efficient ways for mobilizing crowds and harnessing the crowd’s intelligence [98, 95]. Such systems have allowed problem solving, learning, information sharing, prediction and decision making among the collectives, through real-time collaborations [III], [65, 74]. Sadly, many of these systems are commercial and their underlying architectures (models and frameworks) are often not available in scientific literature [III]. CI systems that have been examined and described in scientific literature have only been discussed with focus on their reliability and popularity. And most studies have only investigated parts of these CI systems, and have thus focused on questions such as, how these systems and their users are contributing to the society; rather than focusing on the underlying architecture on which these systems were built [III]. This absence of explicit and comprehensive knowledge of the underlying architectural principles of CI systems has led to a *reproducibility crisis* with respect to existing CI frameworks.

1.2.2 CI Models are Domain-Specific In order to understand the underlying architectures of CI platforms, researchers have proposed several CI models. Unfortunately, many of these models have been designed for specific types of CI systems, each being domain- or use case specific. Furthermore, CI models that have been described in literature, are usually described using different terminologies, vocabularies, semantics and metaphors such as "system-specific elements, principles, attributes, requirements, or their combinations" [III]. And therefore, these models can be interpreted as being distinct entities; for example, in the CI genome model, Malone et al. describe CI systems by using the analogy of biological genes [67]. They explain that CI systems can be viewed as being made of distinct building blocks (namely *What*, *Who*, *Why*, and *How*), and that, similar to how biological systems can be distinctly identified based on their genes, all CI systems can be examined by answering the questions "What is being done? Who is doing it? Why are they doing it? How is it being done?" [68, 67]. Other CI models available in literature describe CI systems in several different ways, for instance, some scholars view CI systems from an organizational perspective [69, 31], whereas others focus on knowledge creation and exchange, and some only focus on CI systems from an engineering and development perspective [63].

1.2.3 Lack of a Generic CI Framework Due to rising interest in the applications of CI, especially in prediction and problem solving [65, 74], private organizations and government bodies world-wide are developing many CI platforms; e.g., privately owned platforms like InnoCentive and OpenIDEO (where web users contribute by proposing new innovative ideas and solutions [65, 74]), or government-owned platforms like MyGov.in [52] and vTawain [3] that aim to improve decision making and policy creation by engaging citizens in these processes. Many such examples of CI systems have been enumerated and organized based on specific domains by NESTA [4]. Furthermore, developing CI systems requires lots of planning and requirement analysis. Unfortunately, most CI systems resp. platforms or initiatives are developed independent of previous knowledge, due to the lack of fundamental understanding of the concepts of CI. Therefore, many CI systems become obsolete after a few years. We found that one of the reasons for CI systems becoming obsolete is that many systems lack the required features that enable collectives to self-organize, adapt, and evolve [III]; thus making the development and maintenance of such CI systems very expensive. This lack of well defined and systematic knowledge about the concepts and models of CI in ICT [III] has not yet lead to the development of a complete generic framework for CI systems [I].

1.3 Research Questions

In this work, we focus on two major research gaps: first, the lack of systematic and well-described knowledge regarding the theoretical and technological aspects of CI systems, and second, the lack of a generic CI framework that would empower such systems' stakeholders and developers by making it easier for them to design, develop and maintain new CI systems, irrespective of their application domains.

We aim to answer the three primary research questions (RQs) stated below. Table 1 presents a mapping of each of these (RQs) with respect to the publications presented in this work.

- *RQ1* What are the underlying models of existing CI systems?
  - *RQ1.1* What are the common terminologies used to describe CI models?

  - *RQ1.2* What are the components of CI models? And, how are these components associated to each other?

- *RQ2* Do any of the available CI models appropriately define all CI systems, irrespective of their applications?
  - *RQ2.1* Can these models be used to create CI systems for novel challenges?

- *RQ3* How can the available knowledge of CI models and systems be combined to create a generic model that defines all CI systems?
  - *RQ3.1* How would such a generic model be evaluated?

Table 1 – Mapping of associated RQs and publications.

Research Question	*Publications*
RQ1	[I], [III]
RQ1.1	[I], [II], [III]
RQ1.2	[III]
RQ2	[I], [III]
RQ2.1	[I], [II], [III]
RQ3	[III]
RQ3.1	[III], [IV], [V]

2 Research Methodology

The main structure of this work is composed of three original peer-reviewed research articles. Publication [III] is a journal article, while [I] and [IV] are articles published in conference proceedings. Furthermore, two additional research articles, [II] and [V], were published (in conference proceedings) as part of this work. Each research article supports the claims presented in this work and is based on distinct studies as illustrated in Fig. 1. Altogether, these research articles have allowed for the creation of a novel conceptual framework for CI systems that is complete and comprehensive while at the same time being generic. The generic framework is based on an the amalgamation of existing CI research in the field of ICT, including existing domain-specific CI models (covered by [I] and [III]), plus several currently used CI systems (both proprietary and open) that have not been discussed in scientific literature so far (discussed in [I] and [III]) and supported by own experience in implementing CI concepts (covered by [II]). The development process of the novel framework is covered in [III], while the evaluation of the framework through case studies [III] and use case scenarios is described in [IV] and [V].

The work presents a novel artifact designed as a conceptual framework. The proposed framework is developed based on a rigorous study of the available scientific literature, and has been evaluated (and cross-validated) using multiple different approaches, namely case studies and use case scenarios. The designed artifact contributes a crucial tool to the domain of group-based decision making and problem solving, which gets ever more relevant for the positive development of today's societies. Given this rigor and relevance, the research efforts fulfill usual best practices and principles of high-quality design science (DS) research [43].

The particulars of the research methodologies adapted in this work are described in Sects. 2.1, 2.2, 2.3 and 2.4.

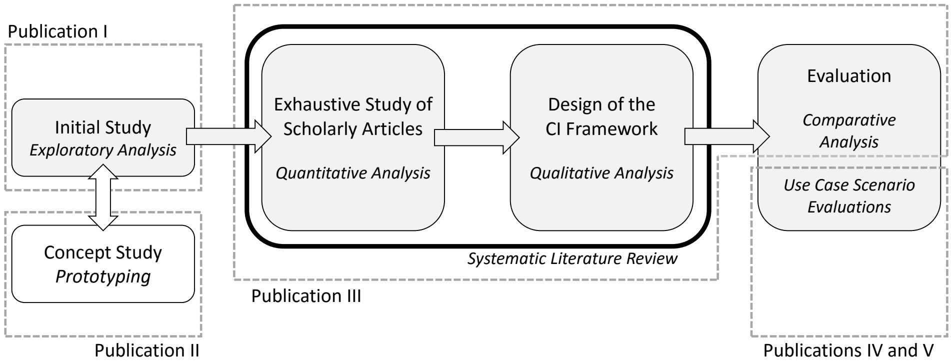


Figure 1 – Design science process. The figure shows the design science process of the work (grey) plus an auxiliary concept study (white), together with corresponding methodologies (in italics) and corresponding publications (dashed lines).

2.1 Exploratory Analysis

In our initial study [I], we found that the genome model for CI proposed by Malone et al. in 2010 [67] is the most prominent model for development of CI systems. Through our examination of the model we found that among the several CI models that have been presented in scientific literature so far, the genome model is the only one that is considered general enough to be able to describe all CI systems. However, as pointed out by scholars (including Malone et al. themselves), the model provides only an abstract view for understanding CI systems [67]. Furthermore, as pointed out in [I], the platforms that have been studied by Malone et al. during the development of the genome model were all developed between 2000 and 2010, i.e., more then a decade ago. We argue that given the several technological advancements that have occurred in the past ten years (and the several social features that have emerged due to increasing interest in social media), the model may not be able to adequately describe today's CI systems and hence there is a need to develop a more comprehensive generic CI framework. To evaluate the genome model in context of current CI systems, we asked the question: *"Can the genome model comprehensively describe recent CI platforms? If not, what new genes could be proposed to improve the model?"*[I]. In order to answer this question, we conducted an exploratory analysis [116, 38] of 10 CI platforms deployed after 2015 (for specifics please refer to [I]). To identify the platforms for the study, we followed a predefined set of criteria; first, we chose only those platforms that were currently in use, and whose documentations (such as technical reports, research articles, and deliverables) were available on the open web, and second, the platforms had to have more than 3,000 users. Following these criteria, we were able to identify 10 CI systems that we examined during the study. As part of the study, we participated on the systems as a passive, observing user over a duration of four weeks.

The efforts of the exploratory study have been supported by (and at the same time supported) a concept study on enabling social information exchange on the web through the use of web annotations [II].

At the end of the study, we concluded that the CI genome model could not describe new CI platforms and their new components and features [I]. Moreover, we hypothesized a more complete and generic framework for CI systems can be achieved by aggregation of the genome model with other existing CI models [I].

2.2 Quantitative and Qualitative Analysis

To answer the research questions, we started by conducting an initial literature review of scholarly articles discussing CI platforms and frameworks. Through this review process, we found that CI is studied in a wide variety of domains, and thus, the subject is often confused with other similar concepts such as citizen science, swarm intelligence and more. Since the focus of the work is CI in ICT, it was critical to extract and analyse only those scholarly articles that discussed exactly this. Therefore, to answer the RQs, we decided to conduct a comprehensive and exhaustive literature review using the well-established Kitchenham's guidelines for conducting systematic literature reviews (SLRs) (in software engineering) [54]. As explained by Kitchenham et al. [102, 53], SLRs are a suitable research method for accumulating completed and documented research regarding a particular research topic. They can be utilized "to identify any gaps in current research in order to suggest areas for further investigation" [54] and "to provide a framework/background in order to appropriately position new research activities" [54]. Following these guidelines and insights, we conducted an SLR of CI-related literature published in the past two decades. We identified about 10,000 scholarly articles, out of which some 200 were selected for a deeper review. Out of these 200 articles, we identified 12 research studies that adhered to the predefined inclusion and exclusion criteria. After further examining these 12 research articles and their references in detail, we identified 24 unique properties that were discussed in the studied articles. By combining and reorganizing these unique properties, we were able to propose a novel, complete and generic framework for development of CI systems [III]. The details of the SLR methodology as described in [III] are presented in Table 2.

Table 2 – Results of the four phases of the SLR.

Phase	Outcome
Phase 1	Based on the search string, 9,418 research articles were identified from selected academic databases (ACM, Elsevier, IEEE and Springer).
Phase 2	Pre-defined inclusion criteria were applied on identified articles of Phase-1, and 219 research articles have been selected for Phase-3.
Phase 3	After the second phase, study quality assessment criteria were applied on the 219 articles of Phase-2 and, finally, 12 CI models have been identified.
Phase 4	Finally, relevant data were extracted from the selected CI models and synthesized in order to answer all RQs.

2.3 Comparative Case Study

In order to evaluate the developed generic CI framework, we conducted a comparative case study [116, 38] in [III]. This study involved six CI platforms, that were examined over a duration of six months. These six CI platforms were selected based on a pre-defined selection criteria. These criteria were: the selected CI systems had to be ongoing during the period of the study, each of the platforms had to be focused on different disciplines and areas of application, and finally, their details had to be available in the scientific literature. To investigate these platforms, we created new user profiles for every system and then observed different user activities such as how users collaborated with each other, how they made decisions (both as individuals and groups), how the users participated within the system (i.e., actively or passively) and more. Overall, each selected platforms was studied over six months. We also studied the available research articles related to these platforms. At the end of the study, we mapped the activities we observed on these platforms with our proposed CI model and found that all components of these platforms directly corresponded to specific components of our generic CI framework in [III].

2.4 Use Case Scenarios

To evaluate the generic CI framework further, we modeled two scenarios, i.e., novel artifacts designed as web-based crowd-oriented systems. The first platform is called CommunityCare [IV]. The platform concept was proposed to tackle mental health issues by empowering citizens and bring them together to help each other. Through the design of the web-based platform, we explained how several diverse humans and machine components of a crowd-oriented system as well as their actions corresponded to the different components of the proposed CI framework. The second platform [V] has been designed and developed with focus on bridging the gap between CI and business intelligence (BI).

3 Contributions

The main contributions of this work can be summarized as follows:

- The work *addresses the issue of reproducibility crisis in regards of the absence of explicit and comprehensive knowledge of the underlying architectural principles of CI systems (i.e., CI frameworks)*. By investigating ten ongoing CI initiatives developed within the EU Horizon 2020 framework and by analysing their corresponding scientific literature, we were able to deduce that not even the most cited CI model (i.e., the CI genome model [67]) is able to describe all components of today's CI systems [I]; there by, partially addressing the Research Question (RQ2): "Do any of the available CI models appropriately define all CI systems, irrespective of their applications?". It should be noted that during our exploratory analysis we found four frameworks, three of which (i.e., all except the CI genome model [67]) focused on domain-specific CI applications. Considering the lack of systematic knowledge about CI models available in published literature, we found it imperative to conduct a reproducible and exhaustive search of CI frameworks; and to this end, we conducted an SLR [III].

- The work *presents a first and foremost exhaustive study of CI research articles*. By methodologically and transparently conducting an exhaustive meta-study of 9,418 CI related research articles, we were able to identify a total of 12 CI models (and frameworks) that were proposed since the year 2000 [III]. By strictly adhering to Kitchenham's guidelines [54] for conducting systematic literature reviews, we ensured that the process of identifying the said models/frameworks was reproducible. Similar to our finding from the exploratory analysis, (through the SLR) we found that none of identified CI models were able to "appropriately define all CI systems" (thereby addressing the Research Question, RQ2). By studying the individual components of each of the identified CI models (and frameworks), we were able to identify the "common terminologies" and "components" used to describe these state-of-the-art CI models [III], thereby addressing the Research Question (RQ1): "What are the underlying models of existing CI systems?". The breaking down and mapping the components of the 12 CI models, we were able to identify 24 unique attributes that could be used to describe all CI systems, irrespective of their applications (thus addressing the Research Question, RQ2).

- The work *proposes and develops a novel, generic CI framework*. In regard to the Research Question (RQ3): "How can the available knowledge of CI models and systems be combined to create a generic model that defines all CI systems?"; by combining the identified 24 unique attributes, into a unified model (and framework) based on Malone et al.'s [67] building block analogy, we were able to develop a novel 'generic' CI framework [III]. To validate the completeness and genericness of the proposed CI framework, we opted for a multi-tiered approach. We first examined six CI platforms belonging to different domains and attempted to describe the platforms using the newly developed 'generic' framework [III]. The proposed framework was found to be successful at describing all six of the studied CI platforms [III]. As a second evaluation, we proposed two novel CI systems [IV], [V] (both with focus on distinct domains) and demonstrated how the 'generic' framework could appropriately define CI systems regardless of their domains and applications.

- Finally, the work *highlights the new challenges for future CI research, based on theoretical studies and real-world artifacts (i.e., CI projects and initiatives).* As a consequence of the investigations conducted as part of this work [I], [II], [III], [IV], [V], we were able to identify new challenges in CI research that have yet to be deeply explored; for instance, the role of critical mass, the effect user reputation in decision making, and more.

4 Initial Study – An Exploratory Analysis of CI Platforms

In [I], we present an exploratory analysis of ten CI projects developed within the EU Horizon 2020 framework. In this context, we discuss the motivation for the development of new genes (resp. components) for the "CI genome" [67], and introduce three novel genes, i.e., (i) *beneficiaries*, (ii) *knowledge and social cause*, and (iii) *collaboration-based contest*. Moreover, we propose that examining and accumulating additional CI models along with the "CI genome" [67], could help us accomplish a novel, generic CI framework.

As part of work related to [I], we review four CI models/frameworks, i.e., "Toward CI of online communities: A primitive conceptual model" [62], "Harnessing crowds: Mapping the genome of CI" [68], "Collective intelligence systems: Classification and modeling" [64] and "Intelligent collectives: Theory, applications, and research challenges" [76]. Also we discuss the limitation of all these models, i.e., all models are domain-specific and use different types of terminologies to describe the components of their models; for specifics of related work of [I], see Sect. 7.

Next, in [I], we describe the Kaggle CI platform [5] in terms of the "CI genome" suggested by Malone et al. [67], i.e., by using all of the genes suggest by Malone et al. [68]: (i) the *Who* gene: "*Who is performing the task*" [68] (it represents the two types of users, namely *crowd* and *hierarchy*), the *Why* gene: "*Why are they doing it?*" [68] (it represents the motivation which can be *intrinsic* or *extrinsic*), the *What* gene: "*What is being done?*" [68] (it represents the *goal* of the crowd, resp. its *hierarchy*) and the *How* gene: "*How is it being done?*" [68] (*collection*, *contest* or *collaboration*) [68, 65]. For instance, let us consider the question *why* people participate in the challenges posted by Kaggle. To answer this question, we map Kaggle's user motivations to the three categories of the *Why* gene [68], i.e., money, love and glory (money motivates the crowd to participate as collectives in innovation challenges, and love and glory are the intrinsic motivations that inspire members of a collective to help others in the collective).

Moreover, we provide an analysis of ten CI platforms in [I], which all have been developed after 2015, namely:

- CIPTEC (Collective Innovation for Public Transport in European Cities) [75]

- POWER: political and social awareness on water environmental challenges [58]
- Crowd4Roads (Crowd sensing and ride-sharing for road sustainability) [24]
- Open4Citizens: empowering citizens to make meaningful use of data [73]
- Saving Food 2.0: a solution to tackle food waste through collaborative power [25]
- CAPTOR (Collective Awareness Platform for Tropospheric Ozone Pollution) [57]
- COMRADES: platform for community resilience and social innovation during crises [26]
- SOCRATIC (Social Creative Intelligence) [28]
- ChildRescue: a platform for missing children investigation and rescue [71]
- Share4Rare: promoting citizen science in rare disease research [89]

Additionally, to acquire "a deeper understanding of the inner functioning of these platforms"[I], we participated on the systems as a passive user (lurker), i.e., we did not participate in any activities but only observed other users of the platforms over a duration of four weeks. Based on our comprehensive exploratory analysis, we are able to categorize several components of the examined platforms based on the "CI genome" [67]; and the results show that there are some components of these projects that cannot be described using the "CI genome" [67]. Finally, in [I], we propose three new genes, i.e., *beneficiaries* (for the *Who* gene), *knowledge and social cause* (for the *Why* gene), and *collaboration-based contest* (for the *How* gene).

The work in the exploratory study [I] has been further supported by practical efforts (described in [II]) in prototyping a novel web annotation platform for social information exchange. This platform is the outcome of a larger design science project [33, 80, 81, 82] in its own right, with [II] as a concrete work package. The efforts in [II] helped streamlining the efforts (in support of DS *relevance* and *rigor* [43]) in [I], and *vice versa*. Publication [II] presents, how a CI generic framework (guidelines) can be used for a crowdsourcing information system in order to enable real-time collaborations via. annotations on web pages. To do so, first, we explain new challenges related to social media such as misinformation and echo chambers, and how web-based CI platforms can harness the intelligence of collectives to solve some of these challenges.

As results of this initial study, we are able to (i) conclude that the genome model is not generic enough to describe all CI systems, and (ii), state the hypothesis, that a generic framework for CI systems can be achieved by combining the components of the genome model with components of other CI models [I].

5 The Contributed Generic CI Framework

Publication [III] proposes a *generic* framework for CI systems in ICT, as such a framework potentially would make it easier for stakeholders to design new CI platforms and understand existing ones. We discuss the motivation for the development of a generic CI framework, conduct an exhaustive study of research articles related to CI based on Kitchenham's guidelines [54] and propose a novel, generic CI framework. Subsequently, we conduct a comparative case study of six CI platforms to evaluate the proposed framework. Figure 2 shows the details of the CI generic model.

Based on our initial findings of CI models (see Sect. 4), we decided to explore more CI models and conducted an exhaustive SLR (utilizing an SLR, one can accumulate, critically

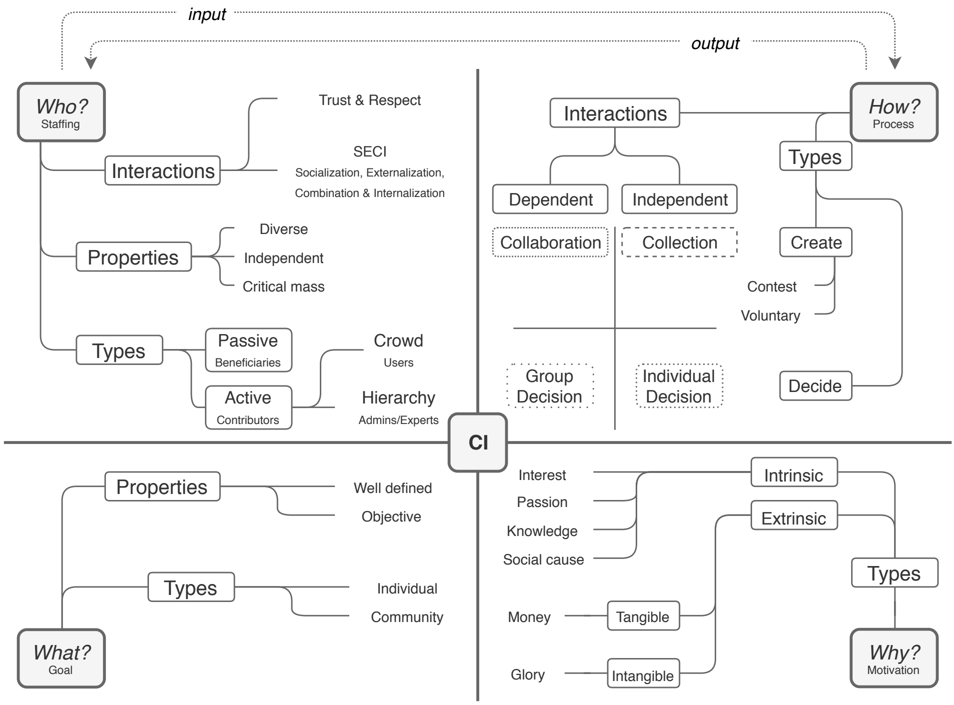


Figure 2 – The generic model for collective intelligence systems from [III].

assess and integrate scientific research studies for a particular RQ) of 9,418 research articles that were selected from four academic databases (i.e., ACM, Springer, Elsevier and IEEE). Based on pre-defined selection criteria, we selected 219 relevant studies, studied them in depth (including their references) and found 12 CI models from different domains. Then, we performed quality assessment and data synthesis to accumulate and sum up the contributions of the relevant studies.

Next, in [III], we elaborate our novel CI generic framework. Our CI framework organizes knowledge about CI systems with respect to the following three aspects:

- A *generic* model that defines all CI systems. This component is the core of our generic framework.
- Additional requisites for CI systems.
- Consideration of CI as a complex adaptive system.

We outline these three aspects of the framework in Sects. 5.1, 5.2 and 5.3.

5.1 A Generic Model that Defines All CI Systems

Taking inspiration from the "CI genome" [67, 68], combined with the results of our SLR (expressed through twenty-four unique attributes), we propose a generic model that explains CI systems by the means of four questions (as components), i.e., "*Who is Performing the Task?*" (Staff), "*How is It Being Done?*" (Process), "What is Being Accomplished?" (Goal), and "Why They are Doing It?" (Motivation). Moreover, we provide new classifications, i.e., properties and interactions for *Staff*, *Process* and *Goal*; and describe all four genes in a fine-grained manner.

To explain every component (its types, properties, and interactions) of the proposed CI generic model, we use different examples of CI platforms:

- "*Who is Performing the Task?" (Staff)*. The proposed model defines the *staff* of a CI system (i.e., the collectives) as being of two types: passive actors and active actors. The passive actors are CI system users who use the information and knowledge produced by other members of the collective, but do not contribute to the system. The active actors on the other hand, contribute artifacts and information to the system, and also give feedback to other members of the collective if required. According to the model, the active actors are further classified as crowd and hierarchy.

  The model states that the *staff* must have the three properties, namely, diversity, independence and critical mass; and ought to interact with trust and respect, while using SECI (i.e., socialization, externalization, combination, and internalization).

- *"Why They are Doing It?" (Motivation)*. In CI systems, users' *motivation* can be categorized into two types: intrinsic and extrinsic. Intrinsic *motivation* can be of four types, namely, interest, passion, knowledge, and social cause; while extrinsic *motivation* can either be a tangible object (such as money), or an intangible object (such as glory). The model explains the users' *motivation* using examples of CI platforms like: DDtrac, Threadless, InnoCentive, Goldcorp, WikiCrimes, and more.

- *"What is Being Accomplished?" (Goal)*. The *goal* of CI systems can be understood as being of two types: individual goals, i.e., the goals of an individual system user, or community goals, i.e., the goals of the system's collective. As explained by model, these *goals* should be well defined and objective.

- *"How is It Being Done?" (Process)*. The *process* of achieving CI in CI systems can be broken down into two types of actions; these are create and decide. The action of creation can be further classified as contest oriented and voluntary. Interaction of the process types can be classified as dependent and independent. Together these process types and interactions can be understood as collaboration (i.e., dependent-create actions), collection (i.e., independent-create actions), group decision (i.e., dependent-decide actions) and individual decision (i.e., independent-decide actions).

  It is important to note here that the process by which individuals and collectives make decisions (i.e., collective behaviour) is still a rather nascent research domain that has gained interest only in recent years [10].

5.2 Additional Requisites for CI Systems

There are other further requisites that should be consider when designing a CI system.

- *System state*: It represents the least number of variables that describe a CI system. It can include different motivations, actors and processes, or the exclusive grouping of the same.

- *Data is the key*: Data (i.e., information and knowledge) provided by the users of a CI system plays an important role in developing new ideas and solutions. Therefore, the system should permit its users to gather, share and make changes to the data.

- *Aggregate knowledge*: As the efficacy of the CI system depends primarily on user-generated content, thus such a system must provide methods (like social tagging or aggregating functions) to aggregate this data (i.e., information and knowledge).

- *Access to decentralized knowledge*: Taking into consideration the user's interests, it is essential that a CI system should be accessible on multiple devices and must provide access to the aggregated knowledge.
- *Task and workload allocation*: During the designing and developing of a CI system, it is important to define (in advance) the task and workload allocation, as this enables effective coordination and collaboration activities.
- *Task-specific representation*: To boost knowledge creation and information exchange among the users, a CI system should offer task-specific representations and also allow users to visualize the same in different forms.
- *Robustness*: As CI systems have multiple components, actors, processes and resources, it is necessary that a system should be able to deal with redundant and erroneous inputs and should also have suitable methods for data backup and recovery.

5.3 CI as a Complex Adaptive System

CI systems are complex dynamic networks of interactions and by complex means, they show the essential properties of complex adaptive systems [19, 46, 72, 12, 13, 30, 32, 31, 70], i.e., (i) *self-organization* (the systems coordinates and manage their inner structure without any external control), (ii) *adaptivity* (components of the system continuously change according to the requirements of its collective) and (iii) *emergence* (over time, the system shows new patterns and properties). These three properties are critical for any CI system [III], without them, any CI system would become obsolete after a few years. Furthermore, studies related to self-organisation [100], adaptivity [96] and emergence [100] are found in different disciplines of research such as collective animal behaviour [104, 103], or swarm behaviour [78, 16]. Researchers also suggest that insights from these disciplines can also help us to understand the self-organization behaviour in humans [78, 51].

6 Evaluation of the Contributed Generic CI Framework

6.1 Evaluation Through a Comparative Case Study

Publication [III] contains an evaluation of the proposed CI generic framework. To evaluate the genericness of our framework, we conducted a comparative case study [116, 38] of six CI platforms, namely:

- CAPSELLA (Collective Awareness Platform for Environmentally-sound Land Management based on Data Technologies and Agrobiodiversity) [60]
- hackAIR: a platform for outdoor air pollution awareness [56]
- openIDEO: a platform to harness collaboration for social good [35]
- ClimateCoLab: open problem-solving platform for climate issues [49]
- WikiCrimes: collective intelligence in law enforcement [36]
- Threadless: global community-driven design platform featuring designs created by various designers, artists and general consumers [17]

The duration of our case study was six months and the platforms were selected on the basis of the following pre-defined criteria: a CI platform should (i) be available online,

(ii) its scientific knowledge should be available in the literature and (iii) it should allow new users to participate in an event or an activity. During the case study, we built ideas (as a passive user) to examine the procedures (as a part of the creation activity of our model), but we never presented the ideas for assessment. We studied the proposals of other users and examined how the groups work together in proposing new innovative solutions. Based on our six months observations, we mapped several characteristics of these platforms with our generic framework and found that these platforms could be explained using our proposed framework.

Interestingly, we also found that when collaborating, CI system users utilized a wide range of communication methodologies, even the ones beyond the scope of the CI platforms. In most cases, CI systems often provide their users with discussion forums, comments, and chats services to interact and communicate with each other; however, as we learned during the case study, CI users often tend to use communication tools that they are more comfortable with (for example, services such as Hangout, Discord, Telegram) even when these communication channels are not included within the system. This observation is in line with Levy's definition of CI in ICT, that states that CI activities must be "coordinated in real-time" [61], however the definition (similar to Malone et al. [68]) does not specify what types of communication channels and methodologies should be used. We would argue that this is so as, communication methodologies change with time and can depend upon the goal and the environment of a CI system, also on the structure of a CI system's crowd; therefore, defining communication methods for CI systems can be detrimental to a CI framework's generalizability. And this why when defining CI, we explain it as having three main components, the third being only a "means/platform for real-time communication (viz., hardware/software)", and not as its specific types, properties, or interactions.

6.2 Evaluation Through Use Case Scenarios

Publications [IV] and[V] present further evaluations of our generic CI framework.

Publication [IV] presents a crowd-oriented web-based CI platform for mental well-being as a use case for our generic CI model. Mental health issues refer to a broad range of problems including depression, anxiety, and hypertension. With the rising interest in applications of CI (such as learning and knowledge exchange), many private organizations started building CI platforms to tackle mental health-related problems. In our initial investigation, we found that these platforms are popular and offer several features to help the individuals and their doctors and family members, however, each of them has some limitations; for example, all the platforms are owned by private organizations, and are closed behind paywalls, furthermore, most of these systems only offer their services only in specific areas (i.e., specific cities and countries). Furthermore, most of the systems only offer personalized services instead of providing overall support and counselling in an open environment. By keeping all these aforementioned points, we designed the CommunityCare platform for "tackling mental health issues with the help of community" [IV]. The platform is intended to tackle mental health issues by empowering citizens and bring them together to help each other. To do so, we describe the various components of the CommunityCare platform through the components (i.e., twenty-four unique attributes) of our 'generic CI framework and thus, evaluate the framework. Figure 3 shows the details of the mapping of four components of the CI generic model as described in [IV].

Publication [V] demonstrates the potential of CI for business intelligence (BI) in today's organizations [29, 34, 31, 32]. To do so, first, we explain why CI applications (such as problem-solving, decision making and prediction) are so popular and how can we harness

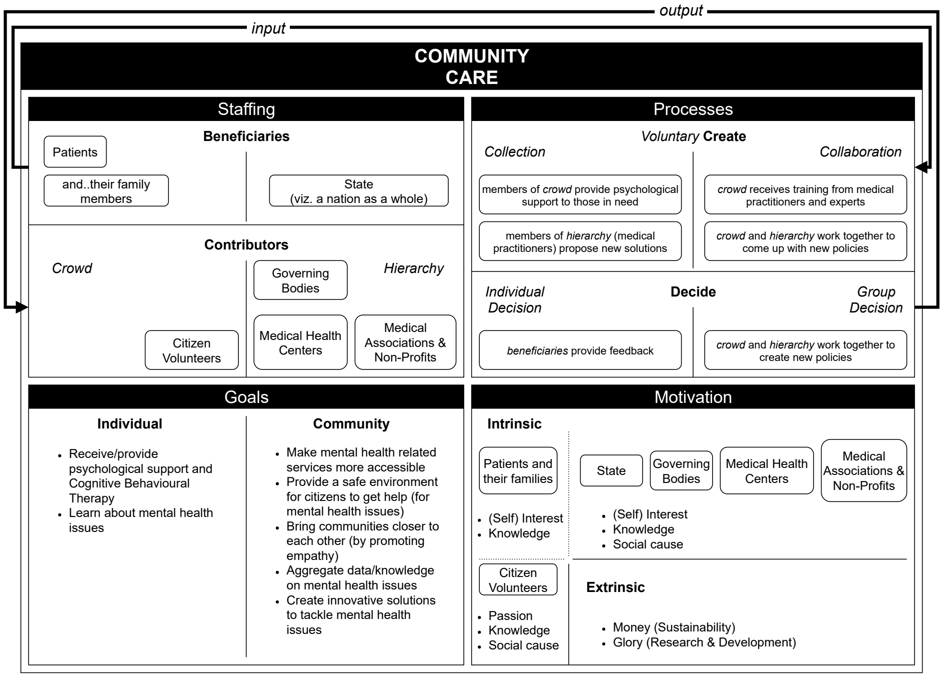


Figure 3 – Components of CommunityCare illustrated via our 'generic' CI model [IV].

the intelligence of employees of an organization in order to solve BI-related problems.

7 Related Work

In this section, we provide an overview of the most cited and most relevant CI models and frameworks from different domains published in the field of ICT.

For [I], we analyzed three CI models as discussed in the sequel.

The first model of CI in ICT was proposed by Boder in 2006 i.e., "A new model for CI in organizations" [14]. The author states that knowledge management is a key procedure for building new knowledge, for understanding particular questions, and for fostering innovation in an organization. Boder states that a major issue in knowledge management is the creation of organizational intelligence, as it is difficult to interconnect collective knowledge and skills from different sources and individuals into a single object. Taking inspiration from Nonaka's model of "The Knowledge-Creating Company" [77], Boder explains how, from a CI viewpoint, Nonaka's model can enable information and knowledge exchange (from tacit to explicit) within an organization [14]. Based on Nonaka's model, Boder introduces three building blocks for organizational CI that can enable creation of organizational knowledge: development of competencies, development of goals, and development of mechanisms. Boder explains competency development as the organization's need to develop complementary competencies among its employees. This is vital, as individuals with different competencies can use their implicit knowledge to come up with innovative solutions that might be impossible to develop in a homogeneous group. Through goal development, the independent goals of several organizational units should be aligned in order to allow the emergence of organizational intelligence. Finally, mechanical development means that mechanisms for interactions within organizations must be aligned, i.e., these mechanisms should be explicitly stated to the members of the organization, and they should foster trust and respect among the several groups within an organization. Although the model is full of important insight, we find that it is explicitly designed for organizational contexts and, therefore, we find that the model needed to adjusted if it were to be adapted to CI systems on the web (which are typically non-hierarchical).

The rise of the Internet has changed the way individuals and groups collaborate and interact; thanks to the Internet, millions of collectives can now collaborate with each other, share their knowledge and experiences, solve global challenges, and propose novel solutions to wicked problems [68]. For instance, through ClimateCoLab, which is an open problem-solving platform with more than 115,000 users, the members of the platform continuously generate new ideas to tackle climate-related problems [49]. Malone et al. examined more than 250 "web-enabled CI systems" [68], and, based on their analysis, proposed a novel model of CI systems. Using the analogy of biological genes, they described CI systems through their CI genome model that describes these systems by asking the questions "What is being done? Who is doing it? Why are they doing it? How is it being done?" [68, 67]. The authors proposed that all CI systems can be understood using four building blocks, namely staffing, incentive, goal, and structure/process. The authors further described that the staffing in CI systems can be either crowds or hierarchies. Incentives can be money, love, and glory. Goals of a CI systems can either be results of creation or decision, while, structures and processes can be of four types, namely collection, collaboration (i.e., creation), group decisions, and individual decisions. The authors further explained these building blocks in context of two CI platforms, i.e., Wikipedia and InnoCentive, and in terms of the Linux software community. Malone et al. also point out that in the future, more genes should be added to the current pool as may be needed to explain future CI systems [68, 67, 65].

Another model for CI was proposed in 2009 by Iandoli et al. in [48]. The authors also emphasize the importance of the web and how it enables mass collaboration in support of solving societal issues. In their work, Iandoli et al. present a novel model for collectives on online communities with focus in the field of management. Iandoli et al. argued that the various open issues of CI could be organized into two categories, namely "management of collective intelligence" and "design of collaborative tools" [48]. The authors proposed that, if collectives were to be modelled as organizations, they would need to inherit five distinct properties, including: clearly defined objectives or goals, critical mass (i.e., adequate number of participants), predefined sets of processes, rules and regulations (for community member to follow during interactions among themselves and with the system), and lastly, roles and responsibilities based on the tasks at hand [47, 48]. Iandoli et al. also pointed out that, even if virtual communities were successfully modelled as organizations, they would still encounter three critical issues. These are, that the communities would have to tackle biased knowledge, the members of a community would have to be motivated from time to time, and the roles and responsibilities of the community members would have to be updated from time to time, to ensure that interactions within the community remain smooth and stable. A critical issue pointed out by the authors was that CI platforms must have "trust and reputation appraisal systems" if they are about tackling the above mentioned issues [47].

For publication [III], we analyzed nine more models of CI as discussed in the sequel.

In order to use CI for complex problems, it is required to built a framework for engineering-oriented CI systems. Based on published scientific literature, Vergados et al. proposed a "resource allocation framework for CI system engineering" [114] in 2010. The authors stated that the proposed framework can promote CI in web-based platforms. Vergados et al. described their framework as having three main components: “human community, machine intelligence, and system information” [114]. The authors argued that the proposed framework could be utilized to develop several distinct forms of CI systems, and that these systems would still have some common underlying attributes. These attributes would include “set of possible individual actions, system state, community and individual objectives, expected community member action functions, future system state functions, objective functions, resource allocation algorithms, critical mass and motivation” [114]. A brief overview of these attributes can be found in [III]. It is important to note here, that Vergados et al. evaluated their framework by utilizing a simulation approach wherein they analyzed the quality of Wikipedia articles in order to understand how the quality of the analyzed articles could be improved by using the above-mentioned characteristics. We argue that given that the authors only examined one CI system, their findings can not be considered general, and, therefore, are not applicable to other types of CI systems.

In 2010, Schut proposed guidelines for the development of a new CI model [96]. To propose these guidelines, the author conducted several studies and discovered crucial contributions that clearly differentiated CI systems from other kinds of ICT systems, for example, information retrieval systems. Based on the available scientific literature, the author identified two categories of properties, namely properties that enable CI (i.e., adaptivity, individual behaviour, system behaviour, and interactions among the collective members) and CI-defining properties such as global–local, complex systems, redundancy and emergence. Based on these properties the author proposed a “systematic approach for designing CI system models” [96]. This systematic modelling methodology is distributed into three levels: system design, model design and models; and is further divided into three categories, namely generic modelling, specific modelling, and computer modelling). Furthermore, to evaluate the guidelines, the author conducted two case studies: the Chinese Whispering Room study (which Schut created by combining the well-known Chinese Whispers game and the ‘Chinese Room Argument’ thought experiment proposed by Searle [97]) and the Braitenberg collectivae scenario (based on Braitenberg vehicles [18]) [96]. Schut’s guidelines provide novel insights that go beyond conventionally discussed properties of CI, and have been the first to discuss adaptivity and emergence as key properties required to enable CI in ICT.

In 2010, Gregg proposed seven key requirements for designing CI applications, namely [41, 42]): “task-specific representation”, “data is the key”, “users add value", “facilitate data aggregation", “facilitate data access”, “facilitate access for all devices”, and “the perpetual beta”. The author proposed these requirements based on prior work of O’Reilly [79]. To explain how these requirements could be utilized to develop new CI systems, Gregg developed a web-based CI application called “DDtrac” [41, 42]. The application was designed to provide education and therapy to children with special needs. To evaluate the proposed requirements, the author observed the application and its users over a duration of 18 months. At the end of study, Gregg concluded that the application was successfully able to achieve its core objective, and, therefore, the proposed requirements could be used to create other CI systems. We find that, although the requirements identified in the work are critical for development of CI systems, they are not complete, as the evaluation has been carried out with a single domain-specific application.

To identify the common characteristics of CI systems, Lykourentzou et al. investigated the existing literature of CI and identified few common characteristics for CI systems [63]. They found that all CI systems have the following characteristics [63]: “set of possible individual actions, system state, community and individual objectives, expected user action function, future system state function, objective function, critical mass, task and workload allocation and motivation". Many of these characteristics are based on the work conducted by Vergados et al., [114]. Building on Vergados et al.'s work, Lykourentzou et al. proposed that CI systems can be divided into two categories, namely active systems and passive systems [63]; the active systems are further divide into collaborative systems, competitive systems, and hybrid systems. The authors then described the various types of CI systems based on the proposed characteristics; these CI systems (and initiatives) included Wikipedia and Open Source Software Development Communities (as examples of active collaborative CI systems), the DARPA Network Challenge (as an example of active competitive CI systems), and vehicular ad-hoc networks (VANETs) (as an example of passive CI systems). The authors also found that critical mass is a key factor for CI systems and that it requires further investigation to understand how many individuals or collectives are necessary to make sound decisions or to solve problems [63].

In 2012 Georgi et al. [37], proposed a comprehensive model based on scientific literature. The authors pointed out that research in CI is very limited and that most research that has been conducted so far is domain-specific. To this end, the author conducted a literature review and identified three models for CI in the existing literature. They then combined these models (the CI genome model [68], "mitigating biases in decision tasks" [15], and Lykourentzou et al.'s CI model [64]) and identified five characteristics that could be utilized to describe CI systems in general. These characteristics are [37]: "objective of the task", "size of the contribution", "form of input", "form of output", and "stakeholder". However, the authors did not evaluate the proposed characteristics.

In their study, Salminen conducted an SLR of scholarly articles providing case studies on three CI platforms (Quirky, OpenIDEO, and Threadless) [94]. The author conducted additional literature reviews and participated on the three CI platforms over a duration of one month. Based on these studies, Salminen investigated the role of CI in innovation and identified existing CI frameworks or models. The author concluded that CI systems can be described as having three levels: micro, emergence and macro. These levels can be further divided into themes, that represent different elements of theoretical CI frameworks [94]. In their work, Salminen raised a critical issue found on CI systems. The author found that users often created multiple fake accounts to increase or decrease the voting counts (i.e., up-/down-votes) on CI platforms, and that this can allow members of the platform to modify the CI systems outcomes. However, the author did not propose a solution to tackle this issue.

Skarzauskiene et al. analyzed different available CI frameworks and proposed a new conceptual framework. The authors described the proposed framework as having three levels: capacity Level, emergence Level and social maturity level [101]. The authors also conducted qualitative and quantitative studies on the various types of actions carried out by users on web groups in Lithuania. Combining the findings of the study with the three-level framework, the authors proposed a novel *CI potential index*.

Based on James Surowiecki's work "The Wisdom of Crowds" [110], Matzler et al. proposed that to harness the CI inside an organization, it is essential to follow four steps: "create cognitive diversity, promote independence, access decentralized knowledge, and effectively aggregate knowledge" [69]. Matzler et al. also stated that wikis and blogs are insufficient to harness the wisdom of crowd within organizations, and therefore, it is important to follow the aforementioned four steps. To explain these steps, the author conducted two case studies. In the first, they investigated "how diversity can drive innovation" [44], while in the second, they explored "the CEO's role in business model reinvention" [40]. The authors also studied Wikipedia's peer review system [69] and the PreMortem exercise [55] and concluded that the proposed steps adequately enable CI in organizational settings.

In 2018, Nguyen et al. proposed a novel CI framework [76]; influenced by Bonabeau's work on Decisions 2.0., Bonabeau defined their contribution as "a new era of decision-making in which the traditional decision-making process is supported using the wisdom of crowds through collaboration and CI" [15]. Nguyen et al. supported Bonabeau's claim and proposed four criteria for collectives. Using the criteria (namely diversity, independence, decentralization, and aggregation), the authors proposed a novel CI framework that included three characteristics: collective, aggregation methods, and collective performance measures [76].

For [IV], we looked at several crowd-oriented web-based platforms. CI platforms today have become an important venue of today's problem solving, decision making, and learning [65, 74], and this has lead to the development various kinds of CI systems. In [IV], we studied CI platforms that focus on assisting individuals suffering from mental health related issues. We found that platforms such as Kooth [88], iFightDepression [50], InnoWell [92], MoodPath [20] and Talkspace [93] offer therapies and guidance for individuals suffering from mental health issues. These platform also support family members of individuals who need therapies or other forms of psychological health support. Although these platforms are widely popular and provide several features to help individuals, their doctors and their family members, each of them has its drawbacks; for instance, all of these platforms are privately owned, and therefore often expensive for most individuals, also, the services provided by these platforms are mostly limited to specific cities, states or countries. Furthermore, these platforms only provide personalized services rather than providing overall help, therapies and counselling in an open environment.

8 Implications for Further Research

In future research, we will continue to work on the CI systems and deepen research in the following areas:

- I would like to investigate the role of critical mass in decision making and problem-solving.

- I would like to investigate why sometimes crowds show self-destructive behaviour ('Mad Crowd').

- I would like to investigate the literature of CI in the field of animal behaviour, to understand what models have been proposed by researchers in order to understand self-organization behaviour and cascading behaviour. I would like to implement the same models in human behaviour to analyze the differences in their self-organization and emergence behaviour.

- I would like to explore crowd behaviour on social media and understand the challenges crowds face and why, and how these challenges can be tackled.

The above-mentioned areas of research will widen the scope of CI not only in ICT, but also in other domains such as social sciences, group dynamics and social media. They will also assist other researchers to understand what kind of challenges members of a collective can face from others within a system such as: online bullying, information cascading and echo chambers.

9 Conclusion

This work focuses on the critical issue of lack of a domain-independent and reproducible framework for collective intelligence systems in ICT. We address these issues by developing and evaluating a novel generic CI framework, by combining insights from the vast literature in the field. We achieve this, by first conducting an exploratory literature study, followed by an exhaustive meta-study of more than 9,000 scholarly articles. From the literature, we identify key components of existing CI systems and frameworks, and combine them to create a novel framework. The proposed framework is further explained through a new CI model and additional requisites that enable CI. Finally, we evaluate the generic CI framework through a comparative case study and multiple use-case scenarios.

List of Figures

List of Tables

References

[1] [last accessed 20th Oct. 2021] https://thegovlab.org/.

[2] [last accessed 20th Oct. 2021] https://thelivinglib.org/frameworks-for-collective-intelligence-a-systematic-literature-review/.

[3] [last accessed 25th Oct. 2021] https://info.vtaiwan.tw/.

[4] [last accessed 25th Oct. 2021] https://trello.com/b/cdterrc6/collective-intelligence-projects.

[5] [last accessed 25th Oct. 2021] https://www.kaggle.com/.

[6] [last accessed 25th Oct. 2021] https://www.oxfordlearnersdictionaries.com/.

[7] [last accessed 28th Oct. 2021] http://osiris-smartsilvereconomy.eu/.

[8] [last accessed 28th Oct. 2021] https://www.blackrock.com/aladdin.

[9] P. Aragón, A. Kaltenbrunner, A. Calleja-López, A. Pereira, A. Monterde, X. E. Barandiaran, and V. Gómez. Deliberative platform design: The case study of the online discussions in decidim barcelona. In G. L. Ciampaglia, A. Mashhadi, and T. Yasseri, editors, *Social Informatics*, pages 277–287. Springer International Publishing, 2017.

[10] J. B. Bak-Coleman, M. Alfano, W. Barfuss, C. T. Bergstrom, M. A. Centeno, I. D. Couzin, J. F. Donges, M. Galesic, A. S. Gersick, J. Jacquet, A. B. Kao, R. E. Moran, P. Romanczuk, D. I. Rubenstein, K. J. Tombak, J. J. Van Bavel, and E. U. Weber. Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, 118(27), 2021.

[11] S. Beer. Fanfare for Effective Management – Cybernetic Praxis in Government. The 3rd Richard Goodman Memorial Lecture, Delivered at Brighton Polytechnic, Moulsecoomb, Brighton, 14th February 1973.

[12] S. Beer. *Brain Of The Firm*. Allen Lane, the Penguin Press, 1972.

[13] S. Beer. *The Heart of the Enterprise*. John Wiley & Sons, 1979.

[14] A. Boder. Collective intelligence: A keystone in knowledge management. *Journal of Knowledge Management*, 10(1):81–93, 2006.

[15] E. Bonabeau. Decisions 2.0: The power of collective intelligence. *MIT Sloan Management Review*, 50(2):45–52, 2009.

[16] E. Bonabeau and C. Meyer. Swarm intelligence – a whole new way to think about business. *Harvard Business Review*, 79(5), 2001.

[17] D. C. Brabham. Moving the crowd at threadless. *Information, Communication & Society*, 13(8):1122–1145, 2010.

[18] V. Braitenberg. *Vehicles: Experiments in Synthetic Psychology*. MIT Press, 1986.

[19] W. Buckley. Society as a complex adaptive system. In W. Buckley, editor, *Modern Systems Research for the Behavioral Scientist*. Aldine, 1968.

[20] S. Burchert, A. Kerber, J. Zimmermann, and C. Knaevelsrud. Screening accuracy of a 14-day smartphone ambulatory assessment of depression symptoms and mood dynamics in a general population sample: Comparison with the PHQ-9 depression screening. *PLoS ONE*, 16(1):1–25, 2021.

[21] S. A. Butt and D. Draheim. Ethical challenges of ICT for the silver economy. In L. Terán, J. Pincay, and E. Portmann, editors, *Proceedings of ICEDEG'201 – the 8th International Conference on eDemocracy & eGovernment*, pages 152–155. IEEE, 2021.

[22] S. A. Butt, E. A. Elhadjamorb, I. Pappel, E. Õunapuu, and D. Draheim. A knowledge map for ICT integration in the silver economy. *Procedia Computer Science*, 181:693–701, 2019. Proceedings of CENTERIS'2021 – the 12th International Conference on Enterprise Information Systems.

[23] S. A. Butt, K. Rava, T. Kangilaski, I. Pappel, and D. Draheim. Designing a digital collaborative platform for the silver economy: inception and conceptualization. In L. Terán, J. Pincay, and E. Portmann, editors, *Proceedings of ICEDEG'201 – the 8th International Conference on eDemocracy & eGovernment*, pages 47–54. IEEE, 2021.

[24] C. Certomà. *Digital Social Innovation: Spatial Imaginaries and Technological Resistances in Urban Governance*. Springer, 2021.

[25] C. Certomà, F. Corsini, and M. Frey. Hyperconnected, receptive and do-it-yourself city. An investigation into the European "imaginary" of crowdsourcing for urban governance. *Technology in Society*, 61:1–14, 2020.

[26] J. Chamberlain, B. Turpin, M. Ali, K. Chatsiou, and K. O'Callaghan. Designing for collective intelligence and community resilience on social networks. *Human Computation*, 8(2):15–32, 2021.

[27] R. Chaves, D. Schneider, C. Motta, A. Correia, H. Paredes, B. Caetano, and J. M. de Souza. Crowdsourcing urban narratives for a post-pandemic world. In *Proceedings of CSCWD'2021 – 24th IEEE International Conference on Computer Supported Cooperative Work in Design*, pages 894–900. IEEE, 2021.

[28] I. Dinant, J. Floch, T. Vilarinho, and M. Oliveira. Designing a digital social innovation platform: From case studies to concepts. In I. Kompatsiaris, J. Cave, A. Satsiou, G. Carle, A. Passani, E. Kontopoulos, S. Diplaris, and D. McMillan, editors, *Proceedings of INSCI'2017 – the 4th International Conference on Internet Science*, pages 101–118. Springer, 2017.

[29] D. Draheim. Smart business process management. In L. Fischer, editor, *2011 BPM and Workflow Handbook, Digital Edition*, pages 207–223. Future Strategies, Workflow Management Coalition, 2012.

[30] D. Draheim. The present and future of large-scale systems modeling and engineering (FDSE'2016 Keynote). In T. K. Dang, R. Wagner, J. Küng, N. Thoai, M. Takizawa, and E. Neuhold, editors, *Proceedings of FDSE'2016 – the 6xt International Conference on Future Data and Security Engineering*, Lecture Notes in Computer Science 10018, pages 355–370. Springer, 2016.

[31] D. Draheim. Collective intelligence systems from an organizational perspective (iiWAS'2019 Keynote). In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2019 – the 21st*

International Conference on Information Integration and Web-based Applications and Services, pages 3–4. ACM, 2019.

[32] D. Draheim. On the radical de- and re-construction of today's enterprise applications (CENTERIS'2019 Keynote). *Procedia Computer Science*, 164:120–122, 2019. Proceedings of CENTERIS'2019 – the 10th International Conference on Enterprise Information Systems.

[33] D. Draheim, M. Felderer, and V. Pekar. Weaving social software features into enterprise resource planning systems. In F. Piazolo and M. Felderer, editors, *Novel Methods and Technologies for Enterprise Information Systems*, Lecture Notes in Information Systems and Organisation 8, pages 223–237. Springer, 2014.

[34] D. Draheim, T. Kopetzky, and J. Küng. How to make mobile bpm robust and intelligent. In L. Fischer, editor, *Intelligent BPM – 2013 BPM and Workflow Handbook*, pages 107–116. Future Strategies, Workflow Management Coalition, 2013.

[35] M. Fuge, K. Tee, A. Agogino, and N. Maton. Analysis of collaborative design networks: A case study of OpenIDEO. *Journal of Computing and Information Science in Engineering*, 14(2), 2014.

[36] V. Furtado, L. Ayres, M. de Oliveira, E. Vasconcelos, C. Caminha, J. D'Orleans, and M. Belchior. Collective intelligence in law enforcement – the WikiCrimes system. *Information Sciences*, 180(1):4–17, 2010.

[37] S. Georgi and R. Jung. Collective intelligence model: How to describe collective intelligence. In *Advances in Intelligent and Soft Computing*, volume 113, pages 53–64. Springer, 2012.

[38] D. Goodrick. *Comparative Case Studies*. SAGE Publications, 2020.

[39] L. S. Gottfredson. Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1):13–23, 1994. *first published in the Wall Street Journal, December 13, 1994*.

[40] V. Govindarajan and C. Trimble. The CEO's role in business model reinvention. *Harvard Business Review*, 89(1-2):108–14, 2011.

[41] D. Gregg. Developing a collective intelligence application for special education. *Decision Support Systems*, 47(4):455–465, 2009.

[42] D. G. Gregg. Designing for collective intelligence. *Commununications of the ACM*, 53(4):134–138, 2010.

[43] A. R. Hevner, S. T. March, J. Park, and S. Ram. Design science in information systems research. *MIS quarterly*, 28(1):75–105, 2004.

[44] S. A. Hewlett, M. Marshall, and L. Sherbin. How diversity can drive innovation. *Harvard business review*, 91(12):30–30, 2013.

[45] K. A. Hintikka. Web 2.0 and the collective intelligence. In *Proceedings of MindTrek'08 – the 12th International Conference on Entertainment and Media in the Ubiquitous Era*, pages 163–166. ACM, 2008.

[46] J. H. Holland. Studying complex adaptive systems. *Journal of Systems Science and Complexity*, 19(1):1–8, 2006.

[47] L. Iandoli. Leveraging the power of collective intelligence through IT-enabled global collaboration. *Journal of Global Information Technology Management*, 12(3):1–6, 2009.

[48] L. Iandoli, M. Klein, and G. Zollo. Enabling on-line deliberation and collective decision-making through large-scale argumentation. *International Journal of Decision Support System Technology*, 1(1):69–92, 2009.

[49] J. Introne, R. Laubacher, G. Olson, and T. Malone. The Climate CoLab: Large scale model-based collaborative planning. In *Proceedings of CTS'2011 – the 2011 International Conference on Collaboration Technologies and Systems*, pages 40–47, 2011.

[50] A. Justicia, M. Elices, A. I. Cebria, D. J. Palao, J. Gorosabel, D. Puigdemont, J. de Diego-Adeliño, A. Gabilondo, A. Iruin, U. Hegerl, et al. Rationale and methods of the iFightDepression study: A double-blind, randomized controlled trial evaluating the efficacy of an internet-based self-management tool for moderate to mild depression. *BMC psychiatry*, 17(1):1–7, 2017.

[51] T. Kameda, T. Wisdom, W. Toyokawa, and K. Inukai. Is consensus-seeking unique to humans? A selective review of animal group decision-making and its implications for (human) social psychology. *Group Processes & Intergroup Relations*, 15(5):673–689, 2012.

[52] B. Kiran and R. Sharma. Designing and driving crowdsourcing contests in large public service organizations – lessons from Deutsche Bahn and Indian Railways: Using the proposed six-stage framework for crowdsourcing contests, public service organizations can leverage internal and external knowledge to cocreate effective solutions. *Research Technology Management*, 64(3):48–57, 2021.

[53] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman. Systematic literature reviews in software engineering – a systematic literature review. *Information and Software Technology*, 51(1):7–15, 2009.

[54] B. A. Kitchenham and S. Charters. Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, Keele University, 2007.

[55] G. A. Klein. *The power of intuition: How to use your gut feelings to make better decisions at work*. Crown Business, 2004.

[56] E. Kosmidis, P. Syropoulou, S. Tekes, P. Schneider, E. Spyromitros-Xioufis, M. Riga, P. Charitidis, A. Moumtzidou, S. Papadopoulos, S. Vrochidis, et al. hackAIR: Towards raising awareness about air quality in Europe by developing a collective online platform. *ISPRS International Journal of Geo-Information*, 7(5):187, 2018.

[57] S. Kroop, T. Schäfer, and B. Kieslinger. The CAPTOR project: Joint efforts reducing ozone pollution. *Frontiers in Environmental Science*, 4, 2016.

[58] M. B. Ksenia Koroleva, K. D. Diogo Vitorino, and J. Novak. Developing a collective awareness platform for urban sustainability challenges: Case study of the POWER project. *European Journal of Sustainable Development*, 8(4):214–226, 2019.

[59] B. Latour. *Reassembling the Social: An Introduction to Actor–Network Theory*. Oxford University Press, 2005.

[60] M. Lazzaro, P. Bàrberi, G. Calabrò, E. Toli, Y. Ioannidis, et al. The Horizon 2020 CAPSELLA project: Collective awareness platforms for environmentally-sound land management based on data technologies and agrobiodiversity. In *Proceedings of the 12th European International Farming Systems Association (IFSA) Symposium – Social and Technological Transformation of Farming Systems: Diverging and Converging Pathways*, pages 1–8. IFSA Europe, 2016.

[61] P. Levy. *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Perseus Books, 1997.

[62] S. Luo, H. Xia, T. Yoshida, and Z. Wang. Toward collective intelligence of online communities: A primitive conceptual model. *Journal of Systems Science and Systems Engineering*, 18(2):203–221, 2009.

[63] I. Lykourentzou, D. J. Vergados, E. Kapetanios, and V. Loumos. Collective intelligence systems: Classification and modeling. *Journal of Emerging Technologies in Web Intelligence*, 3(3):217–226, 2011.

[64] I. Lykourentzou, D. J. Vergados, and V. Loumos. Collective intelligence system engineering. In *Proceedings of MEDES'09 – the 1st International Conference on Management of Emergent Digital EcoSystems*, pages 134–140. ACM, 2009.

[65] T. Malone and M. Bernstein. *Handbook of Collective Intelligence*. MIT Press, 2015.

[66] T. W. Malone and M. Klein. Harnessing collective intelligence to address global climate change. *Innovations: Technology, Governance, Globalization*, 2(3):15–26, 2007.

[67] T. W. Malone, R. Laubacher, and C. Dellarocas. The collective intelligence genome. *IEEE Engineering Management Review*, 38(3):38, 2010.

[68] T. W. Malone, R. Laubacher, and C. N. Dellarocas. Harnessing crowds: Mapping the genome of collective intelligence. Technical Report Working Paper 4732-09, MIT Sloan School, 2009.

[69] K. Matzler, A. Strobl, and F. Bailom. Leadership and the wisdom of crowds: how to tap into the collective intelligence of an organization. *Strategy & Leadership*, 44(1):30–35, 2016.

[70] K. McBride and D. Draheim. On complex adaptive systems and electronic government – a proposed theoretical approach for electronic government studies. *The Electronic Journal of e-Government*, 18(1):43–53, 2020.

[71] A. Michalitsi-Psarrou, C. Ntanos, and J. Psarras. A collective awareness platform for missing children investigation and rescue. In *Proceedings of the 13th International Conference on ICT, Society, and Human Beings*, pages 27–33, 2020.

[72] J. H. Miller and S. E. Page. *Complex Adaptive Systems : An Introduction to Computational Models of Social Life*. Princeton University Press, 2007.

[73] N. Morelli, I. Mulder, J. S. Pedersen, T. Jaskiewicz, A. De Götzen, M. Arguillar, et al. Open data as a new commons. empowering citizens to make meaningful use of a new resource. In *Proceedings of INSCI'2017 – the 4th International Conference on Internet Science*, pages 212–221. Springer, 2017.

[74] G. Mulgan. *Big mind: How collective intelligence can change our world*. Princeton University Press, 2018.

[75] D. Nalmpantis, A. Roukouni, E. Genitsaris, A. Stamelou, and A. Naniopoulos. Evaluation of innovative ideas for public transport proposed by citizens using multi-criteria decision analysis (MCDA). *European Transport Research Review*, 11(1):1–16, 2019.

[76] V. D. Nguyen and N. T. Nguyen. Intelligent collectives: Theory, applications, and research challenges. *Cybernetics and Systems*, 49(5-6):261–279, 2018.

[77] I. Nonaka and T. Hirotaka. *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, 1995.

[78] L. O'Bryan, M. Beier, and E. Salas. How approaches to animal swarm intelligence can improve the study of collective intelligence in human teams. *Journal of Intelligence*, 8(1):9, 2020.

[79] T. O'Reilly. What is Web 2.0: Design patterns and business models for the next generation of software. Technical Report MPRA Paper 4578, University Library of Munich, Germany, 2007.

[80] V. Pattanaik, A. Norta, M. Felderer, and D. Draheim. Systematic support for full knowledge management lifecycle by advanced semantic annotation across information system boundaries. In J. Mendling and H. Mouratidis, editors, *Proceedings of CAiSE Forum 2018 – Information Systems in the Big Data Era*, Lecture Notes in Business Information Processing 317, pages 66–73. Springer, 2018.

[81] V. Pattanaik, I. Sharvadze, and D. Draheim. Framework for peer-to-peer data sharing over web browsers. In T. K. Dang, J. Küng, M. Takizawa, and S. H. Bui, editors, *Proceedings of FDSE'2019 – the 6xt International Conference on Future Data and Security Engineering*, Lecture Notes in Computer Science 11814, pages 207–225. Springer, 2019.

[82] V. Pattanaik, I. Sharvadze, and D. Draheim. A peer-to-peer data sharing framework for web browsers – analysis and evaluation. *Springer Nature Computer Science*, 1(214):1–10, 2020.

[83] V. Pattanaik, S. Suran, and D. Draheim. Enabling social information exchange via dynamically robust annotations. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2019 – the 21st International Conference on Information Integration and Web-based Applications and Services*, pages 176–184. ACM, 2019.

[84] V. Pattanaik, S. Suran, and S. Prabakaran. Inducing human-like motion in robots. In *Proceedings of I-CARE 2014 – the 6th IBM Collaborative Academia Research Exchange Conference*. ACM Press, 2014.

[85] S. A. Peious, S. Suran, V. Pattanaik, and D. Draheim. Enabling sensemaking and trust in communities. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2021 – the 23rd International Conference on Information Integration and Web-based Applications and Services*, pages 1–9. ACM, 2021.

[86] J. Penney, S. McKune, L. Gill, and R. J. Deibert. Advancing human-rights-by-design in the dual-use technology industry. *Journal of International Affairs*, 71(2):103–110, 2018.

[87] R. R. Pimplikar, K. Mukherjee, G. R. Parija, H. Vishwakarma, R. Narayanam, S. Ahuja, R. D. Vallam, R. Chaudhuri, and J. Mondal. Cogniculture: Towards a better human-machine co-evolution. *CoRR*, abs/1712.03724, 2017.

[88] J. Prescott, T. Hanley, and K. Ujhelyi Gomez. Why do young people use online forums for mental health and emotional support? Benefits and challenges. *British Journal of Guidance & Counselling*, 47(3):317–327, 2019.

[89] R. Radu, S. Hernández-Ortega, O. Borrega, A. Palmeri, D. Athanasiou, N. Brooke, I. Chapí, A. Le Corvec, M. Guglieri, A. Perera-Lluna, et al. Global collaborative social network (Share4Rare) to promote citizen science in rare disease research: Platform development study. *JMIR Formative Research*, 5(3):1–12, 2021.

[90] E. S. Raymond. *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. O'Reilly Media, 1999.

[91] P. C. Rogers, S. W. Liddle, P. Chan, A. Doxey, and B. Isom. Web 2.0 learning platform: Harnessing collective intelligence. *Turkish Online Journal of Distance Education*, 8(3):16–33, 2007.

[92] S. C. Rowe, T. A. Davenport, M. A. Easton, T. A. Jackson, J. Melsness, A. Ottavio, J. Sinclair, and I. B. Hickie AM. Co-designing the InnoWell platform to deliver the right mental health care first time to regional youth. *Australian Journal of Rural Health*, 28(2):190–194, 2020.

[93] A. Rutkin. Therapist in my pocket. *New Scientist*, 227(3038):20, 2015.

[94] J. Salminen. *The role of collective intelligence in crowdsourcing innovation*. PhD dissertation, Lappeenranta University of Technology, 2015.

[95] D. Schoder, P. A. Gloor, and P. T. Metaxas. Social media and collective intelligence – ongoing and future research streams. *KI – Künstliche Intelligenz*, 27(1):9–15, 2013.

[96] M. C. Schut. On model design for simulation of collective intelligence. *Information Sciences*, 180(1):132–155, 2010.

[97] J. R. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3):417—-424, 1980.

[98] T. Segaran. *Programming Collective Intelligence : Building Smart Web 2.0 Applications*. O'Reilly, 2007.

[99] P. Singh and N. Shadbolt. Linked data in crowdsourcing purposive social network. In *Proceedings of WWW'13 – the 22nd International Conference on World Wide Web*, pages 913–918. ACM, 2013.

[100] V. Singh, G. Singh, and S. Pande. Emergence, self-organization and collective intelligence – modeling the dynamics of complex collectives in social and organizational settings. In *Proceedings of UKSim'2013 – 15th International Conference on Computer Modelling and Simulation*, pages 182–189. IEEE, 2013.

[101] A. Skarzauskiene and M. Maciuliene. Modelling the index of collective intelligence in online community projects. In *Proceedings of ICCWS'2015 – the 10th International Conference on Cyber Warfare and Security*, pages 313–319. Academic Conferences International, 2015.

[102] H. Snyder. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104:333–339, 2019.

[103] D. J. Sumpter. The principles of collective animal behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361:5–22, 2006.

[104] D. J. T. Sumpter. *Collective Animal Behavior*. Princeton University Press, 2010.

[105] S. Suran, V. Pattanaik, and D. Draheim. CommunityCare: Tackling mental health issues with the help of community. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2020 – the 22nd International Conference on Information Integration and Web-based Applications and Services*, pages 377–382. ACM, 2020.

[106] S. Suran, V. Pattanaik, and D. Draheim. Frameworks for collective intelligence: A systematic literature review. *ACM Computing Surveys*, 53(1):1–36, 2020.

[107] S. Suran, V. Pattanaik, and D. Malathi. Discovering shortest path between points in cerebrovascular system. In *Proceedings of I-CARE'2014 – the 6th IBM Collaborative Academia Research Exchange Conference*. ACM Press, 2014.

[108] S. Suran, V. Pattanaik, M. Singh, P. Gupta, and P. Gupta. Brain imaging procedures and surgery techniques: Past, present and future. *International Journal of Bio-Science and Bio-Technology*, 9(3):23–34, 2017.

[109] S. Suran, V. Pattanaik, S. B. Yahia, and D. Draheim. Exploratory analysis of collective intelligence projects developed within the EU Horizon 2020 framework. In N. T. Nguyen, R. Chbeir, E. Exposito, P. Aniorté, and B. Trawiński, editors, *Proceedings of ICCCI 2019 – the 11th International Conference on Computational Collective Intelligence*, Lecture Notes in Artificial Intelligence 11684, pages 285–296. Springer, 2019.

[110] J. Surowiecki. *The Wisdom of Crowds*. Anchor, 2005.

[111] The Economist. Instant economocis. The Economist October 23rd-29th, page 13, 2021.

[112] The Economist. The real-time revolution. The Economist October 23rd-29th, pages 22–24, 2021.

[113] H. Tyagi, S. Suran, and V. Pattanaik. Weather – temperature pattern prediction and anomaly identification using artificial neural network. *International Journal of Computer Applications*, 140(3):15–21, 2016.

[114] D. J. Vergados, I. Lykourentzou, and E. Kapetanios. A resource allocation framework for collective intelligence system engineering. In *Proceedings of MEDES'10 – the 2nd International Conference on Management of Emergent Digital EcoSystems*, pages 182–188. ACM, 2010.

[115] T. Weninger. An exploration of submissions and discussions in social news: Mining collective intelligence of Reddit. *Social Network Analysis and Mining*, 4(1):173, 2014.

[116] R. K. Yin. *Case Study Research: Design and Methods, 5th edition*. SAGE Publications, 2009.

Acknowledgements

First and foremost, I am extremely grateful to my supervisor Dirk Draheim for his invaluable advice, continuous support, and patience during my PhD study. His immense knowledge and plentiful experience have encouraged me through the duration of my academic research. I would also like to thank my co-supervisors Alex Norta and Ingrid Pappel for their steady consultancy. I am grateful to my alma mater Tallinn University of Technology, for having made this work possible.

Deep from my heart, I would like to express my gratitude to my parents, my husband, and my friends. Without their tremendous understanding and encouragement in the past few years, it would not have been possible for me to complete my study.

Abstract
A Generic Framework for Collective Intelligence Systems

The concept of collective intelligence (CI), i.e., collaborative problem-solving and decision-making, has been a keen interest of philosophers ever since the Aristotelian era. Since the advent of the Internet in 1989, more and more interest has been shown in the concept, in academia as well as in practice, and notions from CI have been adopted and extended into a wide range of scientific domains ranging from sociology, psychology, biology, management, economics to computer science among many others. In this thesis, we focus on CI in Information and Communications Technology, and therefore, understand CI as having three main components: "individuals (with data/information/knowledge); coordination and collaboration activities (according to a predefined set of rules); and a means/platform for real-time communication (viz., hardware/software)". Thanks to advancements in ICT technologies, CI systems can be found in almost walks of life today, ranging from platforms that enable carpooling services, to services that assist people in saving food, to support for learning through open encyclopedias, to solving global challenges such as climate change and pandemics. Empowered by social media technologies, today's CI platforms enable effective and efficient mobilization and utilization of the skills and knowledge of crowds on the web – by allowing collaboration and coordination on massive scales.

However, even with the plethora of CI solutions available on the web, the development of CI systems still remains an exhaustive and costly venture. Literature suggests that this is because, there is a fundamental gap in our understanding of CI systems in general, and that to fill this gap we need to come up with generic yet comprehensive frameworks that can explain CI systems irrespective of their domains. To this end, in this thesis we explore the state of the art in CI research and propose a novel, generic CI framework, that is useful for CI researchers, CI experts as well as developers and practitioners involved in CI project.

The thesis contributes to the current body of research on CI by four distinct contributions. (i) It addresses the issue of reproducibility crisis and lack of systematic knowledge about the underlying architectural principles of CI systems (i.e., their frameworks and models) in ICT. To achieve this, we investigated several state-of-the-art CI systems and platforms. Based on our observations, we argue that current CI systems are developed without a sufficient understanding of CI concepts and are rather based on intuition and the system owners' requirements. We also concluded that none of the available CI models are able to be adequately describes today's CI systems. (ii) By conducting an exhaustive literature review of more than 9,000 scholarly articles published since 2000, we provided the first exhaustive study of CI-related scientific articles. From these, we identified 12 frameworks designed to support development of CI systems. Through our work, we examined each of these frameworks and provided a comparative of the novelties and pitfall of these frameworks. (iii) Building on this analysis, we elaborated a novel 'generic' CI framework that exhibits the components of current state-of-the-art CI frameworks. The proposed framework is explicitly designed to tackle the identified pitfalls, while it is at the same time both generic and comprehensive. (iv) To validate the genericness of the proposed framework, we utilized several evaluation methodologies. First, we conducted a comparative case study of six open and emerging CI platforms. Next, we explained the components of the proposed framework through two use case scenarios, each designed as a crowd-oriented web-based platform. Through this evaluation we conclude that the proposed framework can be used to describe any CI platform, and therefore provides a

complete view of CI in ICT.

As part of future research, we would like to investigate currently less-known aspects of CI, including the role of critical mass in CI, potential challenges that social media presents to CI and reasons for CI turning into self-destructive behaviours.

Kokkuvõte
Kollektiivse intelligentsuse süsteemide üldine raamistik

Kollektiivse intelligentsuse (KI) kontseptsioon ehk koostööl põhinev probleemide lahendamine ja otsuste tegemine on filosoofe köitnud juba Aristotelese ajast peale. Alates interneti loomisest 1989. aastal on nii akadeemilistes ringkondades kui ka praktikas hakatud üha enam huvi tundma kollektiivse intelligentsuse (KI) vastu ning KI-mõisteid on üle võetud ja laiendatud paljudes teadusvaldkondades, sealhulgas sotsioloogias, psühholoogias bioloogias, juhtimises, majandusteaduses ja arvutiteaduses. Doktoritöös keskenduti info- ja kommunikatsioonitehnoloogia (IKT) valdkonnale. Selles kontekstis on KI-l kolm põhikomponenti: üksikisikud (andmete/teabe/teadmistega), koordineerimine ja koostöötegevus (vastavalt eelnevalt määratletud reeglistikule) ning vahendid/platvorm reaalajas suhtlemiseks (st riistvara/tarkvara). Tänu IKT arengule leiab tänapäeval KI-süsteeme peaaegu kõigis eluvaldkondades. Näiteks platvormid, mis võimaldavad sõidujagamisteenuseid, teenused, mis aitavad inimestel toitu säästa, avatud entsüklopeediad, mis toetavad õppimist ja samuti ülemaailmsete probleemide lahendamine, näiteks kliimamuutus ja pandeemia. Tänapäeva KI-platvormidega saab veebis sotsiaalmeedia tehnoloogiate toel makrotasandil rahvahulkade oskuseid ja teadmisi tõhusalt ja tulemuslikult mobiliseerida.

Siiski, kuigi veebis on hulgaliselt KI-lahendusi, on KI-süsteemide arendamine ikka veel mahukas ja kulukas ettevõtmine. Erialase kirjanduse põhjal on see on tingitud sellest, et üldine arusaam KI-süsteemidest on puudulik. Selle lünga täitmiseks tuleb leida üldised, kuid samas põhjalikud raamistikud, mis suudavad kirjeldada KI-süsteeme sõltumata nende valdkonnast. Seepärast on käesoleva töö eesmärk uurida KI-uuringute seisu ja pakkuda välja uudne KI-üldraamistik, mis on kasulik nii KI-uurijatele, KI-ekspertidele kui ka KI-projektis osalevatele arendajatele ja praktikutele.

Doktoritöö täiendab praegust KI-alast teadustööd peamiselt neljast aspektist: (i) see käsitleb KI-süsteemide (st nende raamistike ja mudelite) aluseks olevate ülesehituspõhimõtete (st nende raamistike ja mudelite) taastatavuse probleemi ja süstemaatiliste teadmiste puudumist IKTs. Selleks uuriti mitmeid tänapäevaseid KI-süsteeme ja -platvorme. Kirjanduse uurimisest järeldati, et praegused KI-süsteemid on välja töötatud piisava arusaamata infosüsteemide mõistetest ning põhinevad pigem intuitsioonil ja süsteemi omanike nõudmistel. Samuti järeldati, et ükski olemasolevatest KI-mudelitest ei kirjelda adekvaatselt tänapäevaseid KI-süsteeme, (ii) viidi läbi esimene KI-teemaline kirjanduse ülevaade, milleks kasutati enam kui 9000 teadusartiklit, mis on avaldatud alates 2000. aastast. Nende hulgast leiti 12 raamistikku, mis on mõeldud KI-süsteemide arendamise toetamiseks. Töös uuriti kõiki neid raamistikke ning võrreldi nende raamistike uuendusi ja kitsaskohti, (iii) sellele analüüsile tuginedes töötati välja uudne KI-süsteemide üldraamistik, mis ühendab endas praeguste tipptasemel KI-raamistike komponendid. Välja töötatud raamistik on samal ajal nii üldine kui ka põhjalik ja see on loodud konkreetselt selleks, et lahendada tuvastatud kitsaskohti, (iv) väljapakutud raamistiku üldisuse kehtivuse kinnitamiseks kasutati mitut hindamismeetodit. Esmalt viidi läbi kuue avatud ja areneva KI-platvormi võrdlev juhtumiuuring. Seejärel kirjeldati pakutud raamistiku elemente kahe konkreetse olukorra põhjal, mis olid mõlemad mõeldud rahvahulgale suunatud veebipõhised platvormid. Hindamise põhjal järeldati, et välja töötatud raamistikku saab kasutada mis tahes info- ja kommunikatsioonitehnoloogia platvormi kirjeldamiseks ja seega annab see täieliku ülevaate kollektiivsest intelligentsusest info- ja kommunikatsioonitehnoloogia valdkonnas.

Edaspidi tuleks uurida KI vähemtuntud aspekte, sealhulgas kriitilise massi rolli, sotsiaalmeedia võimalikke väljakutseid ja enesehävituslikuks käitumiseks muutumise põhjuseid.

Appendix 1

[I]

S. Suran, V. Pattanaik, S. B. Yahia, and D. Draheim. Exploratory analysis of collective intelligence projects developed within the EU Horizon 2020 framework. In N. T. Nguyen, R. Chbeir, E. Exposito, P. Aniorté, and B. Trawiński, editors, *Proceedings of ICCCI 2019 – the 11th International Conference on Computational Collective Intelligence*, Lecture Notes in Artificial Intelligence 11684, pages 285–296. Springer, 2019

Exploratory Analysis of Collective Intelligence Projects Developed Within the EU-Horizon 2020 Framework

Shweta Suran(✉), Vishwajeet Pattanaik, Sadok Ben Yahia, and Dirk Draheim

Department of Software Science, Tallinn University of Technology, Tallinn, Estonia
{shweta,vishwajeet.pattanaik,sadok.ben,dirk.draheim}@taltech.ee

Abstract. Over recent years, Collective Intelligence (CI) and crowdsourcing platforms have become a vital resource for learning, problem-solving, decision making and predictions. Unfortunately, the only generic model for developing CI systems (i.e., the CI Genome model) was published nearly a decade ago. Most articles that discuss this model only use examples of older CI projects, thereby raising the question 'Can the genome model comprehensively describe recent CI platforms? If not, what new genes could be proposed to improve the model'? In this article, we answer this question by conducting an analysis of 10 CI projects developed after 2015. We first analyze these projects with respect to the genome model, and then identify three new components namely: Beneficiaries, Knowledge and Social Cause, and Collaboration-based Contest; that could help us improve the genome model, thereby improving our understanding of more recent CI initiatives.

Keywords: Collective intelligence · Wisdom of crowds · Crowdsourcing · Genome model · Exploratory analysis

1 Introduction

The idea of collective intelligence (CI) or 'wisdom of crowd', has been a keen interest among researchers at least ever since the 1780 s; when Marquis de Condorcet proposed the "Jury Theorem" [6]. Since then, numerous authors, researchers and philosophers from different research areas have provided a wide range of insights into the social and biological aspects of CI. The first formal definition of CI in Information and Communications Technology (ICT) however, was provided by Pierre Lévy in 1994 (i.e., after the inception of the World Wide Web). Lévy described CI as "a form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skill." [7]; followed by the several others, among which the most cited definitions are the ones proposed by Glenn (in 2013 [5]) and Malone et al. (in 2015 [10]). Interestingly, both Glenn's and Malone's work were influenced by the rising interest in the Social Web, which gained popularity in the early 2000s.

N. T. Nguyen et al. (Eds.): ICCCI 2019, LNAI 11684, pp. 285–296, 2019.
https://doi.org/10.1007/978-3-030-28374-2_25

Since then, numerous research institutes and organizations have developed and deployed several CI platforms that have enabled coordination, collaboration and competition among individuals [7] over the Web. Some commonly discussed applications of such platforms include: prediction markets [11], crowdsourcing [13], and open innovation [4]. This interest in applications of CI has allowed for distinct research in a variety of fields such as biology [2], management [1], computer science [5] and citizen science [15]. And this has, therefore, lead to the development of distinct CI frameworks and models like the ones proposed by Luo et al. [8], Malone et al. [11], Lykourentzou et al. [9] and Nguyen et al. [12]. Most of these models are designed using insights from specific domains, and are based on domain-specific requirements; presented as completely distinct entities. Unfortunately, this abundance of knowledge has not yet lead to the development of a comprehensive generic model for CI. Among the available models, the most cited is the "the CI genome" model proposed by Malone et al. [11].

Malone's genome model describes the components of CI systems by answering the questions about "Who? Why? What? How?" [11]. Deriving from the answers of these questions, the authors propose the 'genes' of CI systems and explain that the combination of these genes (i.e., genome) is what classifies one CI system from another. Although widely accepted, the model, however, lacks granularity and needs to be developed further [11]. Furthermore, the CI platforms (Linux, Threadless, Wikipedia and InnoCentive) used to describe these genes are some of the most commonly discussed examples and were developed at least two decades ago. This implies that the comprehensiveness of the genome model could be conclusively evaluated by examining newer CI platforms using the model. And, the results of the analysis could provide us with new insights, which could help us come up with new genes (i.e., components).

With these hypotheses in mind, we conduct an exploratory analysis of 10 CI projects developed within the EU's Horizon 2020 research framework. Each of the chosen projects were developed in or after 2015 (some ongoing) and were selected based on the availability of their deliverables, technical reports and publications. To ensure the reproducibility of our research, we only selected the projects whose related documents were openly available on the Web. Furthermore, to gain a deeper understanding of the inner functioning of these platforms, we also registered and participated in each of the platforms. Finally, based on the findings of the analysis, we present a set of components that could help us improve the CI genome model. The remainder of the paper is structured as follows. In Sect. 2, we briefly describe Malone's genome model for CI. In Sect. 3, we describe the previously mentioned 10 CI platforms. Section 4 discusses the findings of the analysis and Sect. 5 concludes the paper.

2 The Collective Intelligence Genome

The genome model defines CI systems using the analogy of biological 'genes,' i.e., its building blocks. And to better explain and classify these building blocks, Malone et al. use four questions "*Who* is performing the task? *Why* are they doing it?" and "*What* is being accomplished? *How* is it being done?" [11]. The answers to these questions namely *staffing*, *motivation*, *goal* and *structure/process* are classified into different 'genes' and "the full combination of *genes* [...] can be viewed as the *genome*" [11] of any particular CI system. It is this 'genome' of each CI system that distinguishes it from other similar systems.

To better explain the genome model, we describe the components of the CI platform *Kaggle* 'Home for Data Science' using the 'genes' proposed by Malone et al. [11]. Launched only a decade ago, the *Kaggle* platform is designed to gather solutions to complex data science problems by "tapping CI of worldwide data science community" [3].

- *Who genes:* These genes answer the question "Who undertakes the activity?" [11] and are classified into two basic genes 'hierarchy' and 'crowd'. *Hierarchy* describes an individual or a group who have been assigned some tasks by the host (the organization that built the CI system or the one that raised the challenge). For instance, in *Kaggle*, an organization that hosts (i.e., provides the dataset and evaluates the solutions) the competition is the *Hierarchy*. Such organizations work in collaboration with *Kaggle*, which allows them to host competitions on the platform.
  Crowd on the other hand, describes an individual who is not assigned specific tasks (by an authoritative figure) and may choose to contribute according to their own will. In *Kaggle*, individuals or groups of participants who participate in competitions, share new datasets or provide new solutions and feedback, are all part of the *Crowd*.
- *Why genes:* These genes answer the question "Why do people take part in the activity?" [11]. These genes are classified into three basic genes, namely 'money', 'love' and 'glory'. *Money* or financial gain can be an important motivator for both individuals employed by the host and the members of the Crowd, who are willing to share new ideas and innovative solutions. *Love* or interest/passion are intrinsic motivators that can encourage an individual to contribute to the system even in the absence of financial motivators. Finally, individuals interested in earning recognition or reputation are motivated by the *Glory* gene.
  Kaggle motivates its users using all three *Why genes*. Participants who submit solutions in competitions are motivated by *money*, while others are either motivated by the 'Kaggle Rankings' in Competitions, Kernels and Discussion, i.e., *glory*; or are motivated by their *love* and interest in learning about data science and machine learning.
- *What genes:* These genes answer the question "What is being done?" [11] and are classified into two genes 'create' and 'decide'. *Create* describes activities where an individual or a group contributes to the system by generating new ideas, artefacts or solutions. *Decide* describes the activities where members of the platform evaluate others' contributions and select the best ones.
  In *Kaggle*, participants *Create* by submitting new datasets, pieces of codes and answers to questions (in discussions); and *Decide* the best topics of discussions, answers and kernels (i.e., codes) by up-voting others' contributions.

– *How genes:* These genes answer the question "How is it being done" [11]. Malone et al. classify this genes as 'collection' (independent-create activities), 'collaboration' (dependent-create activities), 'individual decisions' (independent-decide activities) and 'group decision' (dependent-decide activities). *Collection* can be described as the creation of new ideas, solutions and artefacts independently contributed by individual members of the Crowd; a sub-type of this gene is *Contest*, in which submitted contributions are evaluated and the best ones are rewarded. In contrast to the 'collection' gene, *Collaboration* occurs when members of the Crowd work together to come up with new things. Lastly, *Group decisions* and *Individual decisions* occur when members of the Crowd make some choice or decision; that can have an effect on the group as a whole, or on just the individual (respectively).
  In *Kaggle*, the *Collection* gene occurs when participants contribute to the platform in form of data-sets, kernels and questions/answers (in discussions). While, kernels submitted in *Kaggle Competitions* belong to the *Contest* gene. And, the overall aggregated knowledge (in the form of codes, suggestions and feedback) about any given dataset or challenge can be viewed as a *Collaboration*. Similarly, *Individual decisions* are represented by the individual up-votes participants give to a dataset, kernel or question/answer. Whereas, the 'Hotness' of the datasets, kernels or questions/answers (calculated based on the overall up-votes) and represents the *Group decision* of the community.

3 Exploratory Analysis

As mentioned in Sect. 1 we now attempt to evaluate the comprehensiveness of the CI genome model by examining several ongoing/recent CI projects with respect to the different 'genes' proposed by Malone et al. [11]. To this end, we decided to investigate CI and crowdsourcing projects developed within the EU's Horizon 2020 Research Framework, considering that the deliverables and technical reports of such research projects are openly accessible on the Web. We opted to investigate EU funded projects over other popular business-oriented CI projects like GoldCorp and Threadless, since business-oriented platforms are typically owned by private organizations and therefore the technical descriptions of such platforms are not available in published literature. Business-oriented platforms that are discussed in scientific articles are generally published as case-studies and focus more on the theoretical aspects of the projects and not the technical aspects.

As an additional selection criterion, we decide to choose only those projects that were launched in the last five years. And, had alteast 3000–5000 registered users. Based on the availability of the projects' documentation and access to the projects' website, we decided to examine 10 CI platforms listed in Table 1.

To gain a clear understanding of these platforms, we first examined the available reports (from the projects' website) and then registered as a participant on each of the platforms; after which we studied the different components, activities, motivations and goals of these platforms, over the duration of a month

Table 1. List of selected Collected intelligence projects

PID	Project title	Start date	End date	URL
P1	CIPTEC	May' 2015	April' 2018	ciptec.eu
P2	POWER	Dec' 2015	Nov' 2019	power-h2020.eu
P3	Crowd4Roads	Jan' 2016	Dec' 2018	c4rs.eu
P4	Open4Citizens	Jan' 2016	June' 2018	open4citizens.eu
P5	Saving Food 2.0	Jan' 2016	Dec' 2017	savingfood.eu
P6	CAPTOR	Jan' 2016	Dec' 2018	captor-project.eu
P7	COMRADES	Jan' 2016	Dec' 2018	comrades-project.eu
P8	SOCRATIC	Jan' 2016	Feb' 2018	socratic.eu
P9	ChildRescue	Jan' 2018	Dec' 2020	childrescue.eu
P10	Share4Rare	Jan' 2018	Dec' 2020	share4rare.org

(February' 2019 - March' 2019). Participating in these platforms, helped us gain a better understanding of the inner functioning of each of these projects.

Combining our finding from the related documents and based on our experiences while using the platforms, we first present a brief summary on the selected CI projects; followed by the list of different components of the platforms classified according to the genome model.

- *CIPTEC:* CIPTEC aims to bring new innovative solutions in the form of services or business models, which can help improve urban public transport (PT). The platform gathers and analyzes customer & market intelligence to understand the demand for urban PT. Based on its finding CIPTEC then uses crowdsourcing and co-creation workshops to come with innovative ideas to solve the previously identified PT challenges and issues.
  To gather innovative ideas from its users (i.e., *Crowd*), CIPTEC uses competitive crowdsourcing (i.e., *Contest*); whereas collaborative crowdsourcing (i.e., *Collaboration*) is used to review, rate and discuss the submitted ideas. Members of the CIPTEC (*Crowd*), motivated by *Love* can choose to rate and provide feedback to the submitted ideas (*Individual Decisions*). Based on the users' feedback the best ideas are selected for the co-creation workshop, where the 15 most up-voted ideas (i.e., *Group decision*) are discussed and ranked by CIPTEC's expert advisory board (i.e., *Hierarchy*). After which, the top two ideas with the highest ranking/scores are rewarded with *Money* and free entrance to public transport (for a period of time).
  In addition to its platform, the CIPTEC project also provides a free online toolkit to improve coordination in the planning of solutions and measures in urban PT. The tool can be used by citizens as well as by the transport authorities and policymakers; and allows its users to provide feedback for new implemented ideas in terms of feasibility, ability and novelty.
- *POWER:* The main goal of POWER is to spread awareness and take actions on four water challenges namely: reduction of water consumption (in Milton

Keynes, UK), water quality (in Sabadell, Spain), flood risk (in Leicester, UK) and water conservation (in Jerusalem, Israel). POWER provides a channel for interaction, idea contests and knowledge sharing among citizens, municipal authorities, research organizations and policymakers. The motive of POWER idea contests is to generate new innovative ideas for climate change (specifically problems related to floods and droughts) and to engage communities in order to solve water sustainability issues.

Members of the *Crowd* can provide contributions in the form of sustainability apps and educational/awareness-building projects. The types of users, their motivations and the idea submission-selection process of POWER is similar to the CIPTEC project. The contributors of 10 most up-voted ideas (in any given POWER competition) are invited to a conference, where the community (i.e., *Hierarchy* including municipal authorities, policy makers and research organizations) select the best contributions; and assist the contributors in developing their ideas by providing both guidance and funding. As an intrinsic motivator, the POWER platform also provides its users with a community and personal progress score that indicates their preparedness on problem awareness, know-how and readiness to act to save water.

- *Crowd4Roads:* The Crowd4Roads project combines two initiatives, i.e., trip sharing and crowd sensing; and attempts to solve two well-known sustainability issues of road transport: low occupancy rate in cars and delay in road maintenance. The project attempts to resolve these issues by combining two transport related services, namely *BlaBlaCar* (a ride-sharing service) and *SmartRoadSense* (a crowdsourced road-quality sensing tool).

  BlaBlaCar is a trusted trip sharing (i.e., carpooling) initiative that allows people travelling between pairs of cities to connect and share rides. Car owners travelling from one city to another can offer available seats in their cars, to people who would like to join the trip in exchange for fuel expenses. SmartRoadSense, on the other hand, is a crowd sensing system that exploits the accelerometers of car-mounted smartphones to collect data about road roughness; and helps policymakers in taking well-informed road maintenance decisions.

  Crowd4Roads plans on integrating the carpooling services of BlaBlaCar into its SmartRoadSense app. Doing so the platform intends to provide citizens (i.e., *Crowd*) with a platform where users can contribute information about road quality (i.e., *Collection*) and policymaker could then use this collective knowledge (i.e., *Collaboration*) to make better-informed decisions about road maintenance (i.e., *Group decision*). Finally, the *Individual decisions* of the users on the platform can be seen in their choice of using the BlaBlaCar carpooling service. The users of the Crowd4Roads project can either be motivated by their *Love* for sustainable transport or by *Glory* and geo-coins (i.e., virtual *Money*) within the system's gamification mechanism.
- *Open4Citizens:* The aim of the Open4Citizens project is to empower citizens by allowing them to work with programmers, researchers and policymakers, with the goal to make better use of the open data provided by different governing bodies and organizations; thereby helping in improving the

quality of life in urban areas, specifically in Copenhagen (Denmark), Karlstad (Sweden), Rotterdam (Netherlands), Milano (Italy) and Barcelona (Spain). To achieve this aim, Open4Citizens organized multiple hackathons in each of the previously mentioned cities.

Each of these hackathon events was divided into three sub-events, namely pre-hackathon, second hackathon and post-hackathon. During the first stage (pre-hackathon) citizen groups and communities (i.e., *Crowd*) present challenges/issues which could be solved by open data (motivated by *Love*); to public officers and stakeholders (i.e., *Hierarchy*). The public authorities then take a *Group decision* and select the most important issues they would like to get solutions for through the hackathon. After this, the public authorities, together with the open-data owners, decide the specifications and framing of the dataset that would be provided to the participants of the hackathon. During the main hackathon (i.e., *Contest*) event citizens, programmers and researchers (i.e., *Crowd*) work together and co-create solutions (i.e., *Collaboration*) that are then evaluated by a panel of judges (employed by the governing bodies) and the best solutions are rewarded in cash (i.e., *Money*).

In the post-hackathon event, the solutions provided by the participants are made available on the 'OpenDataLab platform: ODL' (designed by the Open4-Citizens team). The goal of this platform is to provide citizens, students, researchers and policymakers with the consolidated knowledge and tools acquired from the hackathon events; and to raise awareness among citizens about the advantages and uses of open data, through experimental tools.

– *Saving Food 2.0:* The SavingFood project is a collective awareness platform that attempts to tackle food waste in the UK, Greece, Hungary and Belgium. The aim of the project is to raise awareness about food waste among citizens and to provide a solution by which organizations like supermarkets, restaurants and catering services can donate extra food to the people in need. The platform classifies its participants (i.e., *Crowd*) into three groups: donors, recipients and volunteers. The donors are organizations that prepare and store food in large quantities on a daily basis and would like to donate excess food items that might go to waste. The recipients are non-profit organizations like Boroume, Feedback and HFA who provide food to the vulnerable groups and people in need. Finally, the volunteers are citizens who would like to help in the social cause by transporting the food from the donors to the recipients. The collective activities of all three participant groups can be viewed as *Collaboration* as groups of participants come together and help in preventing the wastage of food.

  The SavingFood project provides its participants with a Crowdsourced map that bridges the gap between donors and recipients. Both donors and recipients can mark their locations or events on the map within the platform; thereby providing a *Collection* of addresses of possible collection or drop points for excess food. Participants of the platform can also provide *Individual decisions* in the form of ratings, to each other. The system aggregates the ratings and contributions of participants to reward them with individual scores and badges, thereby motivating them not just by *Love* but also by

Glory. In addition to the badges, the system also provides the donors with a 'Food Waste Cost Calculation Tool' that motivates them to donate excess food by calculating the economic and environmental cost of the food waste.

- *CAPTOR:* The CAPTOR project aims to monitor ozone pollution levels in three European regions: Barcelonès-Vallès Oriental-Osona (Catalonia, Spain), Pianura Padana (Po Valley, Italy) and Burgenland, Steiermark and Niederösterreich (Austria). The project empowers citizens and governing bodies by providing them with real-time ozone pollution data and helps them in making well-informed decisions for behavioural and policy changes. The project also attempts to raise awareness about elevated ozone levels, and it's effects on the day-to-day life of citizens. The developers of the CAPTOR project have designed two main artefacts that support this cause. First, an ozone sensor that is developed, updated and serviced by the CAPTOR team; and second, the CAPTOR tools (AirTact and CaptorAIR) that allow citizens to visualize the ozone data on an interactive map (on the Web). Citizens who wish to host an ozone sensor at their residence can participate in the project as volunteers (i.e., *Crowd*), and are only motivated by *Love*. Once established, these ozone sensors capture the pollution data in the local vicinity and send the data to the CAPTOR repositories (i.e., *Collection*). This collected data is then aggregated and presented on the project's website (i.e., *Collaboration*). Using this knowledge citizens, communities and, non-governmental organizations (NGO)s and governing bodies (i.e., *Hierarchy*) can then propose and introduce new policies (i.e., *Group decision*) to help reduce the ozone pollution in the city. Finally, the CAPTOR team, in collaboration with researchers and NGOs, also conduct workshops where they motivate and educate citizens to take part in the project initiatives.
- *COMRADES:* The COMRADES project aims to provide an open source platform that can assist communities by helping them "reconnect, respond to, and recover from crisis situations"[1]. The project does so by providing communities with two open source services, namely CREES Services (Crisis Event Extraction Service) and Rumour veracity classifier. The services allow its users to analyze tweet and shorts messages from citizens and communities experiencing or witness any kind of crisis; by either validating the truthfulness of a given text or by determining key components (like event type and information discussed) of a crisis related tweet.

  The users of the COMRADES project are categorized into three main groups: first, the communities affected by the crisis (i.e., beneficiaries); second, the reporters (i.e., *Crowd*) or the intrinsically motivated (i.e., *Love*) citizens and communities (i.e., *Crowd*) who share details of the crisis (i.e., *Collection*) on platforms like Twitter; and finally, volunteers and technical communities (i.e., *Hierarchy*) who host and use the COMRADES services to analyze and interpret incoming tweets and messages, and then provide this aggregated information (i.e., *Collaboration*) to the professional responders (i.e., *Hierarchy*) like: members of civil protection authorities, emergency services, government

[1] https://comrades-project.eu/.

agencies, local and international NGOs and UN agencies. These professional responders then take *collaborative* actions to assist, rescue, and support the affected.

- *SOCRATIC:* The main goal of the SOCRATIC project is to provide a collaborative space for citizens and organizations so that both can work together to identify, share and develop innovative solutions to achieve the three sustainable goals set by the United Nations[2].
  The SOCRATIC platform offers a set of tools and services that support 'Social Innovation Project Life Cycle'. The workflow for creation and selection of new ideas is exactly the same as to ones seen in the CIPTEC project. The key users of the system include individual or groups of challenge solvers (i.e., *Crowd*), challenge owners, coordinators and platform owners (i.e., *Hierarchy*), and beneficiaries. The challenge owners and platform owners raise challenges, solutions to which are proposed by challenge solvers who *compete* to achieve financial rewards. And finally, individuals and coordinators choose the best contributions.
- *ChildRescue:* The ChildRescue project aims to "effectively reduce the primary period between the moment a child is reported missing and the one when it is found"[3]. The project intends to attain three primary goals: to develop reliable scientific methodologies that can assist in missing investigation, to develop of a ChildRescue platform that is integrated to platforms of relevant government bodies and finally, to educate and familiarize citizens with the ChildRescue platform.
  The ChildRescue app classifies the process of finding missing children into four phases. During the first phase, law enforcement agencies, other governing bodies and ChildRescue administrators (i.e., *Hierarchy*) *collect* the details of missing migrant children and share it on the ChildRescue platform. The uploaded information is then examined by NGOs, caretakers and search & rescue teams (i.e., *Hierarchy* motivated by *Love*), who attempt to gather more details about the children. This aggregated information (i.e., *Collaboration*) is then viewed by members of society (i.e., *Crowd* motivated by *Love*) and by search & rescue teams who help in tracking down the whereabouts of the missing children. The details of the whereabouts of the children are then shared on the platform, after which the law enforcement agencies and other government bodies rescue the children in a safe and respectful manner.
- *Share4Rare:* The Share4Rare (S4R) platform provide a safe environment where patients, their family members and carers can share their experiences with clinicians and researchers; and can help in making other patients' lives better by contributing in ongoing research. The project is piloted towards two rare conditions: rare paediatric tumours and neuromuscular disorders; and intends to gather clinical data from 2019 to 2020.
  The project would have two types of communities: the awareness community (i.e., *Crowd*) that would constitute of patients with rare conditions and

[2] https://www.socratic.eu/about-us-3/.
[3] https://www.childrescue.eu/.

their carers; and the empowerment community (i.e., *Hierarchy*) which would constitute of the medical practitioners, experts and researchers. Both intrinsically (i.e., *Love*) motivated communities would share their knowledge and experiences over the platform (i.e., *Collection*); and combining this knowledge (i.e., *Collaboration*) the empowerment community would propose new research and guidelines that could help support/improve the condition of the patients.

4 Additional Components for the Genome Model

Based on our exhaustive exploratory analysis described in the previous section, we are able to classify different components of the studied CI platforms based on the genome model. However, as we expected, during our investigation we found that each of the CI platforms had some components that could not be described using 'genes' [11]. These components include: *Beneficiaries*, *Knowledge and Social Cause*, and *Collaboration-based Contest*, and are described as follows:

– *Beneficiaries:* As mentioned in Sect. 2 the genome model classifies the *Who* genes as *Hierarchy* and *Crowd* [11]; however, we found that each of the above described CI platforms also had a third member of the staffing namely, the *Beneficiaries*. These *Beneficiaries* could be both organizations and end-users; while, the *beneficiary* organizations use the solutions and knowledge generated by CI to develop new products and services; the end-users aim to utilize the services and knowledge production on the platform for individual benefits, including their interest to learn about innovative ideas and ongoing research. Such members of the platform, do not participate in any *create* or *decide* activities but still aim to gain from the platform in one form or another. Example of such *Beneficiaries* can be seen in *Crowd4Roads* where users interested in road quality data can simply check the same at *SmartRoadSense* website; whereas, in *Saving Food 2.0* its the vulnerable groups that are provided with the food; and in *CAPTOR* its the citizens who just want to keep an eye on the ozone pollution levels in their vicinity, but are not willing or cannot afford to establish a CAPTOR sensor at their place of residence.
– *Knowledge and Social Cause: Knowledge* and *Social Cause* are another set of possible 'genes' that we found in each of the CI projects. Although these motivators are similar to *Love* which is defined as intrinsic enjoyment, we realized that an individuals thirst for *Knowledge* and enthusiasm to support a *Social Cause* could be distinctly identified when compared to mere interest or passion. However, to understand the exact difference between these motivators and *Love*, we would have to investigate theories of motivation from a social science perspective.
– *Collaboration-based Contest:* Finally, we found that many of the CI platforms allowed groups of individuals to participate in *Contests* as 'teams'; thereby allowing not just collection-based contests, but also collaboration-based contests. However, as we stated earlier, the genome model suggests that only

creation activities conducted by independent members of the *Crowd* can be described using the *Contest* gene. This is a vital finding, because although literature suggests that 'collaboration-based contests' are an important part of CI [14]. However, none of the published CI frameworks and models explicitly define it as a component. Therefore, we are convinced that adding a new *How* gene that describes creation activities by dependent members of the *Crowd* viz. 'collaboration-based contests' could help improve the comprehensiveness of the genome model.

5 Conclusion

To summarize, this article attempts to validate the comprehensiveness of the genome model for CI and presents a set of components that could be included in the genome model to improve its granularity. To this end, we first provide a brief description of the genome model and then explain the different genes using the example of a crowd oriented data science platform 'Kaggle'. We then identified a set of recent or ongoing CI platforms that were developed within the EU's Horizon 2020 research framework. And, selected 10 CI projects based on the availability of the project documentation and accessibility of the project's website. We exhaustively analyzed the related documents of each of these platforms; and then, for a deeper understanding of the projects' inner functionalities, we examined the platforms as participants. Based on our findings, we compared the different components of these platforms to the 'genes' of the genome model and identified three components which could not be described using the model. We believe that this article is just a first step towards developing a novel comprehensive generic CI model. And, we hypothesize that analyzing and aggregating other CI models with the genome model could help us achieve this goal.

References

1. Boder, A.: Collective intelligence: a keystone in knowledge management. J. Knowl. Manage. **10**(1), 81–93 (2006). https://doi.org/10.1108/13673270610650120
2. Bonabeau, E., Meyer, C.: Swarm intelligence: a whole new way to think about business. Harv. Bus. Rev. **79**(5), 106–115 (2001)
3. Brackbill, D.: The network structure of collective innovation. Ph.D. dissertation, University of Pennsylvania (2017)
4. Bugshan, H.: Open innovation using web 2.0 technologies. J. Enterp. Inform. Manage. **28**(4), 595–607 (2015). https://doi.org/10.1108/JEIM-09-2014-099
5. Glenn, J.C.: Collective intelligence and an application by the millennium project. World Futures Rev. **5**(3), 235–243 (2013). https://doi.org/10.1177/1946756713497331
6. Ladha, K.K.: The condorcet jury theorem, free speech, and correlated votes. Am. J. Polit. Sci. **36**(3), 617 (1992). https://doi.org/10.2307/2111584
7. Lévy, P.: Collective Intelligence: Mankind's Emerging World in Cyberspace. Perseus Books, Cambridge (1997)

8. Luo, S., Xia, H., Yoshida, T., Wang, Z.: Toward collective intelligence of online communities: a primitive conceptual model. J. Syst. Sci. Syst. Eng. **18**(2), 203–221 (2009). https://doi.org/10.1007/s11518-009-5095-0
9. Lykourentzou, I., Vergados, D.J., Kapetanios, E., Loumos, V.: Collective intelligence systems: classification and modeling. J. Emerg. Technol. Web Intell. **3**(3) (2011). https://doi.org/10.4304/jetwi.3.3.217-226
10. Malone, T.W., Bernstein, M.S.: Handbook of Collective Intelligence. The MIT Press, Cambridge (2015)
11. Malone, T.W., Laubacher, R., Dellarocas, C.N.: Harnessing crowds: mapping the genome of collective intelligence. SSRN Electron. J. (2009). https://doi.org/10.2139/ssrn.1381502
12. Nguyen, V.D., Nguyen, N.T.: Intelligent collectives: theory, applications, and research challenges. Cybern. Syst. **49**(5–6), 261–279 (2018). https://doi.org/10.1080/01969722.2017.1418254
13. Salminen, J.: The role of collective intelligence in crowdsourcing innovation. Ph.D. dissertation, Lappeenranta University of Technology (2015)
14. Surowiecki, J.: The Wisdom of Crowds. Anchor (2005)
15. Tinati, R., Simperl, E., Luczak-Rösch, M., Kleek, M.V., Shadbolt, N.: Collective intelligence in citizen science - a study of performers and talkers. Comput. Res. Reposit. (CoRR) (2014). https://eprints.soton.ac.uk/363608/

Appendix 2

[II]

V. Pattanaik, S. Suran, and D. Draheim. Enabling social information exchange via dynamically robust annotations. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2019 – the 21st International Conference on Information Integration and Web-based Applications and Services*, pages 176–184. ACM, 2019

Enabling Social Information Exchange via Dynamically Robust Annotations

Vishwajeet Pattanaik
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
vishwajeet.pattanaik@taltech.ee

Shweta Suran
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
shweta@taltech.ee

Dirk Draheim
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
dirk.draheim@taltech.ee

ABSTRACT

With the emergence of new web paradigms, we currently see a tremendous increase in interest in different applications of the social web. However, this rising interest in social platforms has also led to the rise of numerous new challenges, especially issues like fake-news, filter-bubble, and web-page-decay. Motivated by these issues, we propose a novel DOM-oriented edit distance anchoring approach that enables stable tracking of ephemeral web content. We argue that such a stable anchoring approach could indeed foster the creation of a browser-based crowdsourcing information system that could help us tackle rising issues on the web. Building on this hypothesis, we present a new web annotation tool called Tippanee, that is designed around the proposed anchoring approach; and provides its users with a collaborative environment where web users could help in improving the quality of textual content on the web by annotating, archiving, linking, sharing and semantically describing content on-the-fly.

CCS CONCEPTS

• **Information systems → Collaborative and social computing systems and tools**; **Crowdsourcing**; **Web interfaces**; *Social tagging*; Data extraction and integration; • **Human-centered computing → Web-based interaction**; • **Applied computing → Annotation**; • **Theory of computation** → *Pattern matching*.

KEYWORDS

Crowdsourcing, digital reference, edit distance, information exchange, online community, social web, tree matching, web annotation

ACM Reference Format:
Vishwajeet Pattanaik, Shweta Suran, and Dirk Draheim. 2019. Enabling Social Information Exchange via Dynamically Robust Annotations. In *The 21st International Conference on Information Integration and Web-based Applications & Services (iiWAS2019), December 2–4, 2019, Munich, Germany.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3366030.3366060

iiWAS2019, December 2–4, 2019, Munich, Germany

ACM ISBN 978-1-4503-7179-7/19/12…$15.00
https://doi.org/10.1145/3366030.3366060

1 INTRODUCTION

Since the inception of the World Wide Web (WWW) three decades ago, the Web has transformed from a mere medium for broadcast to a dynamic social environment, emerging as a pre-eminent mechanism for global communication, education, political-economic-cultural exchange and more [4]. This rising interest in the Web and its applications has led to a paradigm shift [19] in the ways we solve challenges (e.g., [16]), exchange knowledge/information (e.g., [13, 20]) and innovate (e.g., [12]); specifically by enabling Collective Intelligence through the Social Web [30].

Unfortunately, for a medium that has become so significant to our day-to-day lives, the Web has started to "wane in the face of a 'nasty storm' of issues" [32] such as clickbait, fake news, and misinformation [17]. This emerging trend of fake news is further fueled by the formation of so-called social media filter-bubbles [26], thereby causing ideological polarization [11] among social media users [2]. In addition to these issues, the Web's expansiveness and ephemerality are also becoming increasingly worrisome [9, 28]. Studies have illustrated that the contents of most web pages now change within days, and the decay rate of web documents has dropped to nearly two years [25]. All of this is disconcerting, not only from end-users' perspective but also for content creators, developers, scholars, and researchers; since these digital artifacts not only represent our web experience, but they also hold our collective knowledge [28].

As the literature suggests, a viable approach to tackle some of these challenges could be to leverage the 'wisdom of the crowd' [31, 36]. However, in order to design a multifaceted crowdsourcing information system [14] to save the Web, the system must fulfill three additional requirements. First, the system must act as "a conversation layer over the entire web" [1]. Second, the system should treat web content as "first-class Web citizens" [8]; and finally, the algorithms through which the system interacts with web content shouldn't be susceptible to the Web's ephemerality. Interestingly enough, a class of systems that fulfill the first two requirements perfectly and has gained tremendous interest in recent WWW literature is "web annotation". However, the state-of-art web annotation platforms (like *Hypothesis*) are either unable to cope up with ever-changing web content [1] or, explicitly rely on website owners/developers to be integrated onto the websites (in case of W3C's Web Annotation Standards).

To address these challenges and to build upon our previous research [10, 27, 33], we set out to design an artifact that provides a twofold contribution. We first present a novel anchoring approach

[1] *Hypothesis* article: 'To enable a conversation over the world's knowledge' | https://web.hypothes.is/about/

that can robustly reattach annotations on both static and arbitrarily generated web pages; and then explain how enabling users to interact with the Web by means of stable annotations could support end-user-oriented archiving, linking, sharing and describing (semantically) of web content, on-the-fly. With our work, we set out an alternative approach to achieve the vision of a "humanist format for re-usable documents and media" [24] by facilitating the amalgamation of state-of-web technologies with 'wisdom of crowd'.

To evaluate the robustness of the proposed anchoring approach, we integrated the algorithm into a lightweight, user-friendly web annotation tool named Tippanee [2]. And, to validate our conjecture that crowdsourcing could help us resolve the aforementioned issues with the Web, we designed Tippanee as a social web annotation platform where users can create, store, link, visualize and share textual annotations while browsing the Web. Additionally, we added features that allow Tippanee users to add semantic descriptions to annotated contents, transclude annotations onto other web pages, and view orphaned annotations in their original state, even when the annotated content is altered or completely removed. As a preliminary evaluation of our research, we evaluated Tippanee's anchoring approach by annotating more than 650 random web pages and presented the tool to a small pool of end-users as part of our experiments.

2 RELATED WORK

Since our research is primarily related to web annotation tools and techniques, we first describe the state-of-art web-based text annotation tools. We then discuss the anchoring approach used by the Hypothesis web annotation tool. And, finally, describe a well-known pattern matching algorithm that exists at the center of our proposed anchoring algorithm.

2.1 Web Annotation Tools

Annotating, i.e., "the act of creating associations between distinct pieces of information" [29] has been recognized as a fundamental notion of hypertext systems since the inception of the WWW [21]. Due to its importance in the WWW community, several web annotation systems have been designed and deployed over the years, the most recent ones being Pundit, Genius and Hypothesis. Among these, *Hypothesis* has the most community support as it is a free and open-source platform. Based on the open-source JavaScript (JS) library Annotator.js, *Hypothesis* allows its users to add sentence-level annotations over web pages; and supports open critique and collaborative note-taking, by allowing users to highlight text, add/read and share annotations within private groups or in public.

Unlike the *Hypothesis* web annotation tool that is designed for end-users, the W3C's web annotation recommendations are oriented towards website owners and developers. These recommendation include the Web Annotation *Data Model*, *Vocabulary* and *Protocol* and are successor of the W3C Open Annotation Data Model [29]. The model specification describes a structured format through which annotations can be shared and reused across different hardware and software platforms. The underlying data model constitutes of three parts, namely: *a body*, *a target* and *a relation*.

While both the *Hypothesis* and W3C's web annotation recommendations are widely considered the state-of-art in web annotations, however, both systems have specific drawbacks that prevent them from being used as crowdsourcing information systems for the Web. The *Hypothesis* annotation tool relies on a string-matching-based approach (called *Fuzzy Anchoring*) to reattach annotations. This means that change in annotated texts can cause incorrect reattachment of annotations or can orphan (i.e., when the annotation is no longer attachable) the annotations altogether. This is critical as frequent changes in the annotated content can render the anchoring approach useless. In an empirical study conducted by Aturban et al. [1], the authors found that 27% annotations on *Hypothesis* were already orphaned, and another 61% were at risk of being orphaned if the live web page changed. The drawback of the W3C's web annotation recommendations, on the other hand, is that the choice to utilize these recommendations still lies with the website owners and developers, who are often bound by the technologies used in their platforms and may, therefore, choose not to adopt these recommendations.

2.2 Fuzzy Anchoring

To reattach its annotations *Hypothesis* utilizes a combination of multiple approaches relying on three selectors (namely, *RangeSelector*, *TextPositionSelector* and *TextQuoteSelector*) and four strategies (namely, *From Range Selector*, *From Position Selector*, *Context-first Fuzzy Matching* and *Selector-only Fuzzy Matching*) [3]. The first two strategies of the approach are based on XPath matching and are only usable in cases when there is no textual change in the document. Whereas, strategies 3 and 4 are useful when the underlying textual content is changed. These strategies utilize *approximate string matching* (fuzzy text search and comparison) algorithm, which is a combination of the Bitap matching algorithm [35] and Myers diff algorithm [22]. Since the approach is based on the 'robust anchoring approach' [3, 5], the selectors and the strategies of the approach rely entirely on 'keyword anchoring'.

As mentioned previously, since the *Hypothesis* approach relies primarily on string or keyword matching, textual changes in annotated web pages can cause the annotations to be attached at incorrect locations. For example, we created an annotation "5,710,618" on the Wikipedia homepage on September 13, 2018 using *Hypothesis*. When we viewed the annotation again on November 4, 2018, we found that the annotation was now being reattached to "710 articles" (see Figure 1). And then again, when viewed on January 26, 2019 we found that the annotation was already orphaned. Similar examples of incorrectly attached *Hypothesis* annotations (as viewed on January 28, 2019) on the Wikipedia home page [4] include (*annotated text* vs. *annotation reattached at text*):

- stars *vs.* seasons
- on, a league r *vs.* Hockey League for
- Albert Bridge *vs.* A left winger

[2]URL to Tippanee's source code, available on GitHub: https://github.com/victor013/tippanee-chrome-extension

[3]*Hypothesis* article: 'Fuzzy Anchoring' | https://web.hypothes.is/blog/fuzzy-anchoring/

[4]*Hypothesis* annotations on Wikipedia main page: https://hyp.is/4ZuRcLAJEeiNkQvaWuwvng/en.wikipedia.org/wiki/Main_Page

Figure 1: (a) Annotated text as seen on Hypothesis Chrome extension *vs.* (b) the reattached annotation highlighted in blue

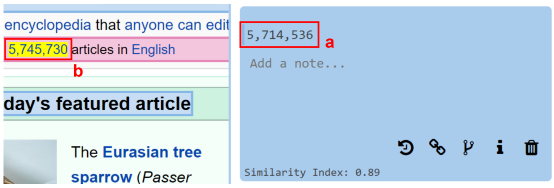


Figure 2: (a) Annotated text as seen on Tippanee Chrome extension *vs.* (b) the reattached annotation highlighted in yellow

We hypothesize that since the fuzzy anchoring approach relies so heavily on keywords, small annotations (i.e., maximum 3-4 words long) have a higher probability of being orphaned compared to long sentence-level annotations. However, further empirical studies would be required to validate the same. Incorrect reattachment of annotation (like the ones shown earlier) is unacceptable because similar changes to web pages could change the context of the annotation or could orphan the annotations altogether. While a few annotations might not be enough to evaluate the robustness of an annotation algorithm, however predicting when and how a textual web annotation would be orphaned can be a daunting task, as such predictions would depend on the half-life of the web page and the length of the annotated text. Continuing with our Wikipedia example "5,710,618", we reproduced the same annotation onto our web annotation tool and found that our proposed anchoring approach was able to reattach the annotation successfully and at the correct location on the web page; even when the annotated content was changed (as shown in Figure 2).

2.3 Edit Distance

The Levenshtein Distance or, Edit Distance algorithm [18] is used to evaluate the difference between two sequences (a, b). The distance represents the number of single-character edits (insertions, deletions or substitutions) required to change one sequence into the other using the following equation:

$$\mathbf{lev}_{a,b}(i,j) = \begin{cases} \mathbf{max}(i,j) & if\ min(i,j) = 0, \\ \boldsymbol{min} \begin{cases} \mathbf{lev}_{a,b}(i-1,j)+1 \\ \mathbf{lev}_{a,b}(i,j-1)+1 \\ \mathbf{lev}_{a,b}(i-1,j-1)+1_{(a_i \neq b_j)} \end{cases} & otherwise \end{cases} \tag{1}$$

Similar to Levenshtein Distance, the Tree Edit Distance (TED) algorithm [37] is used to evaluate the similarity between ordered labeled trees. Two trees are considered similar if their tree edit distance is below a predefined threshold depending upon the chosen path strategy. Although most available TED solutions are quite efficient, they cannot be used to compare HTML Document Object Model (DOM) trees because DOM elements are not necessarily labeled. We work around this problem by generating unique labels for annotated DOM elements. These unique labels are comprised of DOM attributes such as element name, id, class, etc. Once all nodes in a DOM element are uniquely labeled, they are then arranged in prefix notation. This information is then stored as anchor data. When reattaching annotations, web pages DOM elements undergo the same process. The achieved results are then compared to the stored anchor data using the Edit Distance approach.

3 SYSTEM OVERVIEW

3.1 Generating Anchors

To facilitate robust reattachment of annotations it is critical to uniquely identify annotated DOM elements. However, this cannot be done using only XPaths, as changes in the document structure can alter the XPath of annotated element rendering the stored XPath invalid. To counter this challenge, Tippanee's anchoring algorithm uniquely identify anchors by preserving the context and layout of the annotated elements. When a user annotates an element using the annotation tool, the system stores the DOM properties of the annotated element based on following strategy: (1) traverse the annotated DOM element sequentially (in prefix notation); (2) if the DOM element is a node, store its available properties i.e., *id*, *nodeName*, *className*, *alt*, *dataset*, *href* and *src*. And, if the DOM element is *#text*, store its *nodeName* and *nodeValue* and; (3) calculate the depth of each element and store it as *nodeDepth*.

The above-mentioned properties of the annotated elements are then stored as an array of JavaScript Object Notation (JSON) objects. Depending on the selected content, the system then uses the following strategies for anchoring different types of selections:

Figure 3: Illustration of anchored element *(highlighted in blue)* on Wikipedia homepage *[above]* and it's HTML DOM as viewed on Chrome developer tool *[below]*

- *Annotating an element*: If the user annotates a complete DOM element, the anchoring algorithm selects the target DOM and transforms it into the above-mentioned JSON objects.
- *Annotating a #text node within an element*: If the user annotates some *#text* node within an element (but not all the text), the anchoring algorithm selects the target DOM and transforms it into a JSON object. However, in this case, it adds an extra boolean property 'annotated' to the selected *#text* node.
- *Annotating text in two or more elements*: If the user annotates texts from multiple DOM elements, the algorithm selects the common ancestor element as the target DOM and then generates it's JSON object. Again, the 'annotated' property is added to the selected *#text* nodes.
- *Annotating substring within a #text node*: If the user annotates a substring from within a *#text* node, the system additionally stores the selected substring, and it's starting and ending offsets within the *#text* node.

For instance, if a user selects the text "5,569,955 articles" on the Wikipedia main page (see Figure 3), the generated JSON object array would look like so,

```
[
  { id: "articlecount", nodeDepth: 0, nodeName: "DIV" },

  { href: "https://en.wikipedia.org/wiki/Special:Statistics",
  nodeDepth: 1, nodeName: "A" },

  { annotated: true, endOffset: 0, nodeDepth: 2,
  nodeName: "#text", nodeValue: "5,569,955", startOffset: 0 },

  { annotated: true, endOffset: 4, nodeDepth: 1,
  nodeName: "#text", nodeValue: " articles in ", startOffset: 0 },

  { href: "https://en.wikipedia.org/wiki/English_language",
  nodeDepth: 1, nodeName: "A" },

  { nodeDepth: 2, nodeName: "#text", nodeValue: "English" }
]
```

3.2 Reattaching Annotations

To reattach annotations, the algorithm searches for possible target elements using *getElement* DOM methods. The returned DOM elements are then compared to the anchor's JSON object. The algorithm sequentially traverses the returned DOM element and its children. If the stored JSON object attributes match the attributes of the corresponding DOM element, it is considered a match. Whereas, elements that do not have the same attributes are considered mismatch.

In order to compare *#text* JSON objects to respective *#text* nodes, the algorithm uses approximate string matching. In the current version of Tippanee, we use a JS implementation of approximate string-matching available on GitHub [5]. The JS program matches the two texts and scores their similarity with a value between 0 and 1, with 1 being a perfect match. The DOM element with maximum matches and minimum mismatches is selected as a viable target. If the *similarity index* (*SimIndex*), i.e., the ratio between matches and anchor count of the identified elements exceeds the specified threshold (currently 0.5), the DOM element is considered as the final target for reattaching the annotation. In case no DOM element qualifies the minimum *SimIndex* threshold, the system assumes that the annotation is orphaned. Setting up a higher *SimIndex* threshold would mean that the system would only attach anchors that are perfect matches. Whereas, decreasing the threshold too low would cause selection of incorrect targets. Finally, if an annotation is orphaned, the annotation can still be reconstructed in its original form and can be viewed using Tippanee's *reconstruct anchor* feature.

While developing Tippanee's anchoring algorithm, we found that comparing DOM trees using conventional edit distance often leads to instances where multiple viable elements returned the same mismatch count. This led to us introduce a modified version of edit distance algorithm that analyses both mismatches and matches. Using both mismatches and matches, it becomes possible to identify a single element that has the maximum probability of being the target DOM element, making the approach more robust compared to *Hypothesis'* approach. In Figure 4 we present the detailed process of reattaching annotations, illustrated as a flowchart.

[5]GitHub link to fuzzyset.js repository: https://github.com/Glench/fuzzyset.js

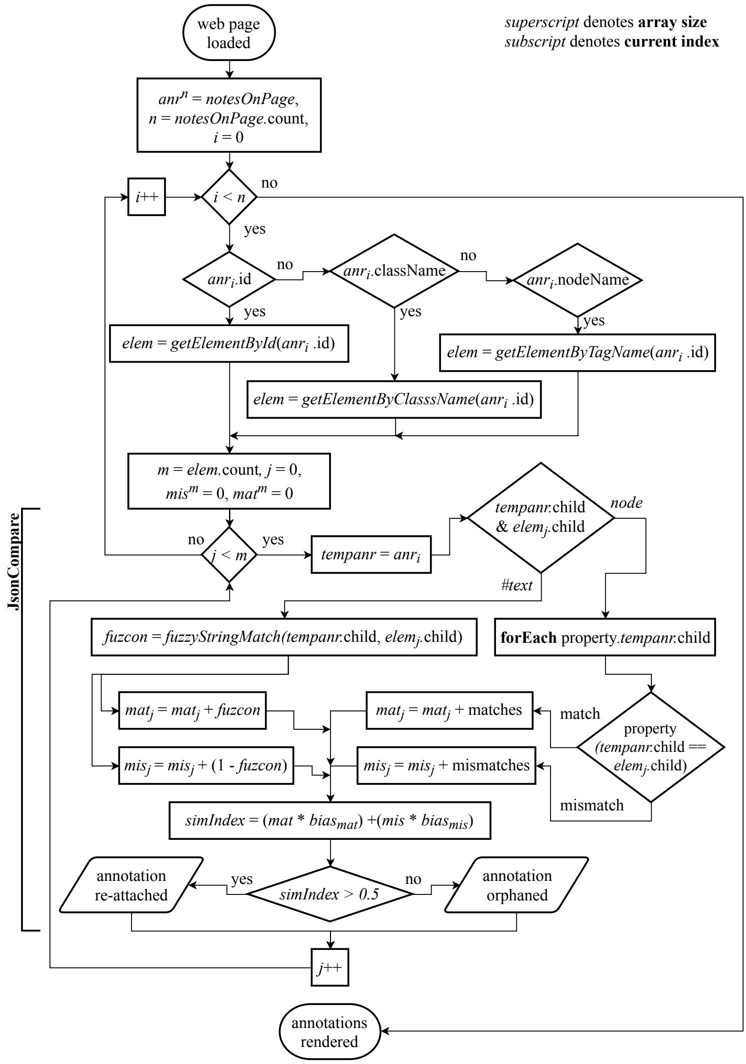


Figure 4: Tippanee's anchoring approach presented as a flowchart

3.3 Implementation

The Tippanee annotation tool is available as a browser extension on Google Chrome Web Store. The extension is designed using HTML, CSS, JS and jQuery [6], with additional JS libraries (vis.js [7] for visualization and fuzzyset.js [8] for fuzzy string matching). The Chrome extension supports both online and offline modes. When using *offline* mode, user annotations are stored locally within Google Chrome's local storage, and therefore the mode does not support annotation sharing. Whereas, in *online mode*, users can create and share annotations in groups, with the server-side code currently deployed on Google's Firebase [9]. Irrespective of the modes, the computational processes for generating and reattaching anchors are always carried out exclusively on the client-side.

[6] jQuery homepage: https://jquery.com/
[7] GitHub link to visjs-network repository: https://github.com/visjs/vis-network
[8] GitHub link to fuzzyset.js repository: https://github.com/Glench/fuzzyset.js
[9] Google Firebase homepage: https://firebase.google.com/

Apart from the conventional annotation tool features, Tippanee also includes some novel components that we believe would allow the system to be utilized as a crowdsourcing information system for the Web. These features include:

- *Reconstructing anchors*: When transforming annotated texts into anchors, Tippanee's anchoring approach stores partial DOM information of the annotated elements. This DOM information can then be utilized to reconstruct the annotations to its original state from when the annotation was first created. This is especially useful in the case of orphaned annotations, as the reconstructed anchor can help users better understand the relevance of the annotated content with respect to its surrounding text. Additionally, the same DOM information can also be utilized to support end-user-oriented information retrieval and web archival.
- *Linking and visualizing annotations*: This feature was inspired by the Linked Open Data Cloud visualization. The feature allows users to link created annotations and then visualize the same as a graph or 'web of annotations'. Such graphs are an effective method of visually extending user-generated annotations and can facilitate meta-analysis and organization of ideas (e.g., [38]). The feature can also help enhance users' web experience by presenting annotations and their associations in an easily understandable format.
- *Transclusions*: The system supports transclusions by allowing users to import and view annotations from other web pages onto the current web page. This is made possible by Tippanee's transclusion server that uses the proposed anchoring approach to scrape requested web documents. After the annotated element is identified within the scraped document, the server pushes the DOM element onto the current web page. However, unlike conventional transclusion services, the pushed DOM element is only a part of the real-time copy of the requested web document.
- *Adding semantic descriptions*: As an added feature, Tippanee allows users to describe annotated text using Schema.org vocabulary. Users can currently only add, review, and change the stored semantic metadata, or they can use this metadata to search through stored annotations. However, in the future we would like to use this feature in support of full knowledge management life-cycle by allowing creation and exchange of new ontologies for organizations (e.g., [27]) and over the World Wide Web. Additionally, we would like to utilize user-generated semantic metadata for improving information access and search on the Web (e.g., [23]).
- *Similarity index*: The concept of SimIndex is a vital part of Tippanee's annotation reattachment procedure. Apart from its role in the anchoring approach, the SimIndex of the attached annotations can also be exploited by end-users to track changes in annotated content. Users can track content changes on platforms such as news articles, Q&A portals, statistics and prices on e-commerce websites. In future, this feature could also support crowd-sourced fake news detection (e.g., [31]) and fact-checking (e.g., [7]).

4 EXPERIMENTAL EVALUATION

As part of the evaluation to validate the stability of the proposed algorithm, we created several web annotations using Tippanee over a duration of three months. To ensure that the conducted experiments are not biased, we replicated a random set of 735 *Hypothesis* annotations created by random users. We acquired these annotations using the 'Annotation viewing and export' tool [10] made available by *Hypothesis* Labs. The tool allows users to view and export publicly shared *Hypothesis* annotations searchable by user, group, URL and tag. We manually replicated the selected annotations on Tippanee's Chrome extension, by first browsing to the annotated web page; and then selecting and annotating the exact text as done in the respective *Hypothesis* annotations. During this process, we also found a few orphaned annotations, which we simply decided to ignore for this experiment. Once the 735 *Hypothesis* annotations were replicated in Tippanee, we left them aside for a month.

After a month, we revisited each of the annotated web pages and extracted the SimIndex(s) of the annotations using Tippanee. We were able to achieve a total of 675 SimIndex(s). Out of these, only 58 (8.59%) annotations had SimIndex(s) below 0.5, i.e., the annotations were now orphaned; while 611 (91.41%) annotations were successfully attached. Among the attached annotations 538 (88.0%) annotations had SimIndex(s) more than 0.9. Figure 6 presents a comparison of the number of annotations with respect to achieved similarity indexes (scaled from 0 - 1.0 to 0 - 100). The SimIndex threshold for the current version of Tippanee was set to 0.5, for this experiment. Our initial experiments clearly indicate a significant improvement of 12.41% over *Hypothesis*' anchoring approach (which had 79% successful reattachments during Aturban et al.'s (2015) experiments); however, to demonstrate the robustness of our proposed approach in a reproducible manner, we plan to develop a first web annotation testbench that would simulate varying levels of change in a set of Alexa's top-ranking websites [11]. The said testbench would be crucial to decisively validate Tippanee's anchoring approach, as it would allow us to mimic web page decay in real-world web pages in a controlled environment; without the need to wait for a live web page to change or decay.

4.1 User Evaluation

Additionally, to evaluate Tippanee's features and to study user behavior when creating annotations on the Web, we presented our work to two different groups of users (12 web developers and 13 students). The groups comprised of both men and women between the ages 25 and 45; and were chosen in a way that they represented a mix of both technical and non-technical users. The groups were given a quick demonstration of Tippanee's user interface and its features, after which they were asked to use the Chrome extension for a duration of at least 7 days. The users were requested to create and share annotations on web pages they visited regularly. After two weeks, the users were provided with a focused questionnaire enquiring about their experiences with the tool.

[10] Annotation viewing and export tool: https://jonudell.info/h/facet/
[11] Alexa's top-ranking websites: https://www.alexa.com/topsites/category

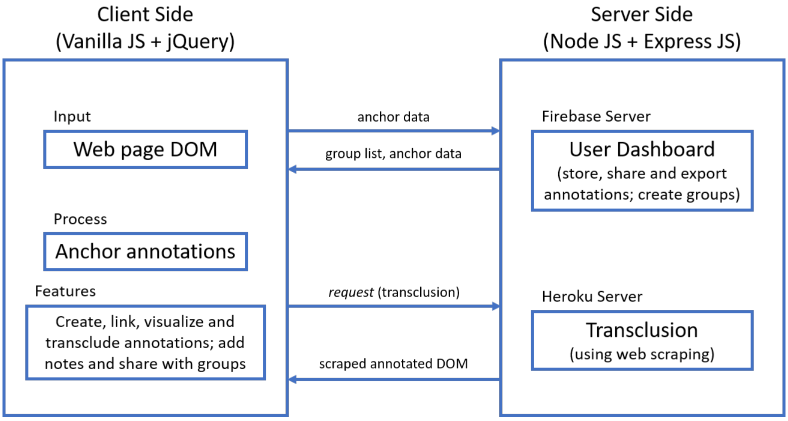


Figure 5: Illustration of Tippanee's system flow

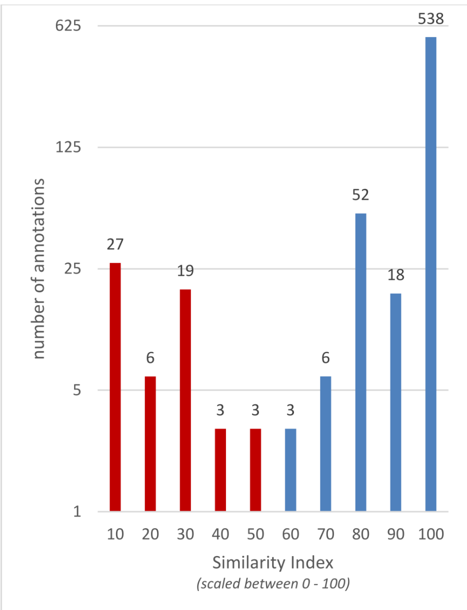


Figure 6: Similarity indexes of 675 annotations created on Tippanee Chrome extension, scaled between 0 - 100; with orphaned annotations colored in orange and reattached annotations colored in blue.

As expected, most users suggested that they found the extension quite useful, especially for information retrieval and sharing. Interestingly, quite a few users suggested that they also used the tool for social interactions and expression of opinions. Finally, users also reported that they created most annotations on educational websites, Q&A portals, news websites, and search engines. Although this preliminary study was carried out only for two weeks, it was valuable for our research as it provided us insights into how users interact on social-annotation platforms and helped us come up with elaborate ideas for future quantitative-quantitative evaluations.

5 DISCUSSION

In this section, we present the limitations of our work and elaborate on the legal aspects of web scraping, and it's relevance with respect to our approach. We also briefly discuss the next steps of our work before we conclude the paper.

5.1 Limitations

We understand that the experiments conducted during the evaluation of Tippanee might not be adequate to demonstrate the robustness of the proposed approach and that further evaluation is required for the same. However, since not all web pages are designed the same, neither is their page half-life; therefore it is difficult to determine which anchoring approach is the most robust. As a part of our future work, we plan on evaluating our proposed algorithm against the Fuzzy anchoring approach by creating a test bench with varying combinations of changes in both web page content and structure.

With respect to our formulae for calculating similarity indexes, the current threshold of 0.5 is only an estimated guess, and might not be ideal. Identifying the ideal threshold would require more extensive and long-term evaluations. Additionally, adding biases to the match-mismatch count of textual nodes could help enhance the accuracy of the algorithm; identifying the ideal values of these biases, is again a part of our future work.

5.2 Legal Aspects

Web scraping, caching and archiving has often been viewed as a severe ethical issue on the World Wide Web. Despite the fact that over the last couple of years the Open Data Movement has received tremendous support from governments, research institutes, laboratories, and libraries. Nevertheless, the idea of a truly open web seems to be a far-fetched notion. Several national libraries around the world have been supporting the concept of digital preservation and archiving, but most of these initiatives are hindered or affected by legal concerns, i.e., copyright infringement and piracy acts [34] (e.g., [6, 15]).

While designing the Tippanee platform, we studied these legal obligations in order to realize the scope of our system. Although web scraping, caching or archiving requires explicit permission from the web page owners, we argue that our system does not violate any piracy or copyright concerns, because of the following reasons: first, the system stores the context of the annotated content only to ensure that users do not lose their annotations and to improve the usability of orphaned anchors. In order words, content is saved primarily for the interest of the end-user. Second, the saved content is utilized solely for reattaching anchors and has no other commercial purpose whatsoever. Third, the context is stored exactly as it was published by the owner. And, the source of the content is clearly stated in the anchor's JSON object, meaning that the content is always linked to the source and the link is visible to the end-user. Lastly, enabling users to add semantic markup to annotations and creating links between annotations, helps in enhancing the user's web experience by encouraging semantic content creation and participation, therefore reinforcing and improving the social aspects of the Web.

6 FUTURE WORK AND CONCLUSION

The proposed anchoring approach provides a proof-of-concept for our hypothesis that using DOM information for generating anchors can facilitate more stable and robust annotations. Such DOM information can be especially useful to support end-user-oriented web archiving and information sharing processes. Finally, the social and semantic aspects of Tippanee can be utilized to support knowledge creation and sharing on the Web. As our next steps, we would like to optimize Tippanee's anchoring approach and add further features to support collaborative critiquing and knowledge sharing activities. And then, develop Tippanee into a lightweight, browser-based crowdsourcing information system that would allow users to contribute to web content by means of stable, interlinked and semantically rich annotations.

To conclude, in this paper, we have presented a novel DOM-oriented edit distance approach for enabling stable annotation and preservation of dynamically evolving web content. We presented evidence revealing that such a DOM-oriented approach could encourage smooth knowledge and information exchange over today's ephemeral web. And, we discussed how collaborative annotation tools like Tippanee could help tackle key challenges on the Web. Given the rising interest in collaborative platforms, annotation technologies and the importance of user's web experience, it is just a consequent step to team together all of these notions. Albeit the combination of these notions opens a wide design space for a whole class of next-generation collaborative annotation platforms, the key challenge for making such systems a success is in providing robust anchoring mechanisms.

REFERENCES

[1] Mohamed Aturban, Michael L. Nelson, and Michele C. Weigle. 2015. Quantifying Orphaned Annotations in Hypothes.is. In *Research and Advanced Technology for Digital Libraries: 19th International Conference on Theory and Practice of Digital Libraries, TPDL 2015, Poznań, Poland, September 14-18, 2015, Proceedings*. Springer International Publishing, Cham, 15–27. https://doi.org/10.1007/978-3-319-24592-8_2

[2] Vian Bakir and Andrew McStay. 2018. Fake News and The Economy of Emotions. *Digital Journalism* 6, 2 (2018), 154–175. https://doi.org/10.1080/21670811.2017.1345645

[3] David M Bargeron, Alice Jane Bernheim Brush, and Anoop Gupta. 2010. Robust anchoring of annotations to content. US Patent 7,747,943.

[4] Tim Berners-Lee. 2010. Long live the web. *Scientific American* 303, 6 (2010), 80–85.

[5] A.J. Brush and David Bargeron. 2001. *Robustly Anchoring Annotations Using Keywords*. Technical Report. Microsoft. 19 pages. https://www.microsoft.com/en-us/research/publication/robustly-anchoring-annotations-using-keywords/

[6] Jhonny Antonio Pabón Cadavid. 2014. Copyright Challenges of Legal Deposit and Web Archiving in the National Library of Singapore. *Alexandria* 25, 1-2 (2014), 1–19. https://doi.org/10.7227/ALX.0017 arXiv:https://doi.org/10.7227/ALX.0017

[7] Sylvie Cazalens, Philippe Lamarre, Julien Leblay, Ioana Manolescu, and Xavier Tannier. 2018. A Content Management Perspective on Fact-Checking. In *Companion Proceedings of the The Web Conference 2018 (WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 565–574. https://doi.org/10.1145/3184558.3188727

[8] Paolo Ciccarese, Stian Soiland-Reyes, and Tim Clark. 2013. Web Annotation as a First-Class Object. *IEEE Internet Computing* 17, 6 (Nov 2013), 71–75. https://doi.org/10.1109/MIC.2013.123

[9] Miguel Costa, Daniel Gomes, and Mário J. Silva. 2017. The evolution of web archiving. *International Journal on Digital Libraries* 18, 3 (01 Sep 2017), 191–205. https://doi.org/10.1007/s00799-016-0171-9

[10] Dirk Draheim, Michael Felderer, and Viktor Pekar. 2014. Weaving Social Software Features into Enterprise Resource Planning Systems. In *Novel Methods and Technologies for Enterprise Information Systems*, Felix Piazolo and Michael Felderer (Eds.). Springer International Publishing, Cham, 223–237. https://doi.org/10.1007/978-3-319-07055-1_18

[11] Ivan Dylko, Igor Dolgov, William Hoffman, Nicholas Eckhart, Maria Molina, and Omar Aaziz. 2017. The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. *Computers in Human Behavior* 73 (Aug 2017), 181–190. https://doi.org/10.1016/j.chb.2017.03.031

[12] Bonabeau Eric. 2009. Decisions 2.0: the Power of Collective Intelligence. *MIT Sloan Management Review* 50, 2 (2009), 45. http://proquest.umi.com/pqdlink?did=1625861001{&}Fmt=7{&}clientId=9268{&}RQT=309{&}VName=PQD

[13] Shixuan Fu, Gert-Jan de Vreede, Xusen Cheng, Isabella Seeber, Ronald Maier, and Barbara Weber. 2017. Convergence of Crowdsourcing Ideas : A Cognitive Load Perspective. In *Thirty eighth International Conference on Information Systems*. Association for Information Systems, 1–11. http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1300{&}context=icis2017

[14] David Geiger, Michael Rosemann, Erwin Fielt, and Martin Schader. 2012. Crowdsourcing Information Systems-Definition, Typology, and Design. In *Thirty Third International Conference on Information Systems*. Association for Information Systems, 1–11. https://ub-madoc.bib.uni-mannheim.de/32631/

[15] Lachlan Glanville. 2010. Web archiving: ethical and legal issues affecting programmes in Australia and the Netherlands. *The Australian Library Journal* 59, 3 (2010), 128–134. https://doi.org/10.1080/00049670.2010.10735999 arXiv:http://dx.doi.org/10.1080/00049670.2010.10735999

[16] Joshua Introne, Robert Laubacher, Gary Olson, and Thomas Malone. 2013. Solving Wicked Social Problems with Socio-computational Systems. *KI - Künstliche Intelligenz* 27, 1 (Feb 2013), 45–52. https://doi.org/10.1007/s13218-012-0231-2

[17] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. The science of fake news. *Science* 359, 6380 (Mar 2018), 1094–1096. https://doi.org/10.1126/science.aao2998

[18] V. I. Levenshtein. 1966. Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady* 10 (Feb. 1966), 707. http://adsabs.harvard.edu/abs/1966SPhD...10..707L

[19] Pierre Lévy. 2010. From social computing to reflexive collective intelligence: The IEML research program. *Information Sciences* 180, 1 (Jan 2010), 71–94. https://doi.org/10.1016/j.ins.2009.08.001

[20] T.W. Malone (Ed.) and M.S. Bernstein (Ed.). 2015. *Handbook of Collective Intelligence*. MIT Press. 232 pages. https://books.google.com/books?id=iR3iCgAAQBAJ

[21] Catherine C. Marshall. 1998. Toward an Ecology of Hypertext Annotation. In *Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia : Links, Objects, Time and Space—structure in Hypermedia Systems: Links, Objects, Time and Space—structure in Hypermedia Systems (HYPERTEXT '98)*. ACM, New York, NY, USA, 40–49. https://doi.org/10.1145/276627.276632

[22] Eugene W. Myers. 1986. An O(ND) difference algorithm and its variations. *Algorithmica* 1, 1-4 (1986), 251–266. https://doi.org/10.1007/bf01840446

[23] Beate Navarro Bullock, Andreas Hotho, and Gerd Stumme. 2018. *Accessing Information with Tags: Search and Ranking*. Springer International Publishing, Cham, 310–343. https://doi.org/10.1007/978-3-319-90092-6_9

[24] Theodor H Nelson. 2007. Transliterature: a humanist format for re-usable documents and media. *Transliterature. org* (2007).

[25] Fatih Oguz and Wallace Koehler. 2016. URL decay at year 20: A research note. *Journal of the Association for Information Science and Technology* 67, 2 (2016), 477–479. https://doi.org/10.1002/asi.23561

[26] Eli Pariser. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Group, UK.

[27] Vishwajeet Pattanaik, Alex Norta, Michael Felderer, and Dirk Draheim. 2018. Systematic Support for Full Knowledge Management Lifecycle by Advanced Semantic Annotation Across Information System Boundaries. In *Information Systems in the Big Data Era*, Jan Mendling and Haralambos Mouratidis (Eds.). Springer International Publishing, Cham, 66–73. https://doi.org/10.1007/978-3-319-92901-9_7

[28] Jeffrey M. Perkel. 2015. The trouble with reference rot. *Nature* 521, 7550 (May 2015), 111–112. https://doi.org/10.1038/521111a

[29] Robert Sanderson, Paolo Ciccarese, and Herbert Van de Sompel. 2013. Designing the W3C Open Annotation Data Model. In *Proceedings of the 5th Annual ACM Web Science Conference (WebSci '13)*. ACM, New York, NY, USA, 366–375. https://doi.org/10.1145/2464464.2464474

[30] Detlef Schoder, Peter A. Gloor, and Panagiotis Takis Metaxas. 2013. Social Media and Collective IntelligenceâĂŤOngoing and Future Research Streams. *KI - Künstliche Intelligenz* 27, 1 (Feb 2013), 9–15. https://doi.org/10.1007/s13218-012-0228-x

[31] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake News Detection on Social Media: A Data Mining Perspective. *SIGKDD Explor. Newsl.* 19, 1 (Sept. 2017), 22–36. https://doi.org/10.1145/3137597.3137600

[32] Olivia Solon. 2017. Tim Berners-Lee on the future of the web: 'The system is failing'. , 2017 pages. http://www.nytimes.com/2010/04/27/world/27powerpoint.html

[33] Shweta Suran, Vishwajeet Pattanaik, Sadok Ben Yahia, and Dirk Draheim. 2019. Exploratory Analysis of Collective Intelligence Projects Developed Within the EU-Horizon 2020 Framework. In *Computational Collective Intelligence*, Ngoc Thanh Nguyen, Richard Chbeir, Ernesto Exposito, Philippe Aniorté, and Bogdan Trawiński (Eds.). Springer International Publishing, Cham, 285–296. https://doi.org/10.1007/978-3-030-28374-2_25

[34] Richard S. Vermut. 1996-1997. File Caching on the Internet: Technical Infringement or Safeguard for Efficient Network Operation. *Journal of Intellectual Property Law* 4 (1996-1997), 273.

[35] Sun Wu and Udi Manber. 1992. Fast Text Searching: Allowing Errors. *Commun. ACM* 35, 10 (Oct. 1992), 83–91. https://doi.org/10.1145/135239.135244

[36] Amy X. Zhang and Justin Cranshaw. 2018. Making Sense of Group Chat Through Collaborative Tagging and Summarization. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 196 (Nov. 2018), 27 pages. https://doi.org/10.1145/3274465

[37] K. Zhang and D. Shasha. 1989. Simple Fast Algorithms for the Editing Distance Between Trees and Related Problems. *SIAM J. Comput.* 18, 6 (Dec. 1989), 1245–1262. https://doi.org/10.1137/0218082

[38] J. Zhao, M. Glueck, S. Breslav, F. Chevalier, and A. Khan. 2017. Annotation Graphs: A Graph-Based Visualization for Meta-Analysis of Data Based on User-Authored Annotations. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 261–270. https://doi.org/10.1109/TVCG.2016.2598543

Appendix 3

[III]

S. Suran, V. Pattanaik, and D. Draheim. Frameworks for collective intelligence: A systematic literature review. *ACM Computing Surveys*, 53(1):1–36, 2020

Frameworks for Collective Intelligence: A Systematic Literature Review

SHWETA SURAN, VISHWAJEET PATTANAIK, and DIRK DRAHEIM, Tallinn University of Technology, Estonia

Over the last few years, Collective Intelligence (CI) platforms have become a vital resource for learning, problem solving, decision-making, and predictions. This rising interest in the topic has to led to the development of several models and frameworks available in published literature. Unfortunately, most of these models are built around domain-specific requirements, i.e., they are often based on the intuitions of their domain experts and developers. This has created a gap in our knowledge in the theoretical foundations of CI systems and models, in general. In this article, we attempt to fill this gap by conducting a systematic review of CI models and frameworks, identified from a collection of 9,418 scholarly articles published since 2000. Eventually, we contribute by aggregating the available knowledge from 12 CI models into one novel framework and present a generic model that describes CI systems irrespective of their domains. We add to the previously available CI models by providing a more granular view of how different components of CI systems interact. We evaluate the proposed model by examining it with respect to six popular, ongoing CI initiatives available on the Web.

CCS Concepts: • **General and reference → Surveys and overviews**; • **Information systems → Crowdsourcing**; **Collaborative and social computing systems and tools**; • **Human-centered computing → Collaborative and social computing theory, concepts and paradigms**; **Collaborative and social computing systems and tools**; *Human computer interaction (HCI)*;

Additional Key Words and Phrases: Collective intelligence, crowdsourcing, human computer interaction, Web 2.0, wisdom of crowds, systematic literature review

ACM Reference format:
Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2020. Frameworks for Collective Intelligence: A Systematic Literature Review. *ACM Comput. Surv.* 53, 1, Article 14 (February 2020), 36 pages.
https://doi.org/10.1145/3368986

1 INTRODUCTION

The concept of "Collective Intelligence" (CI) (i.e., collaborative problem solving and decision-making) has been a keen interest of researchers ever since the 18th century [41, 63]. Since this period, the different applications of CI and its associated concepts have extended throughout a wide spectrum of research domains ranging from sociology, psychology, biology, management, and economics to computer science, among many others [50]. In our work, we focus on CI in Information and Communications Technology (ICT), and therefore, we adhere to the widely accepted formal definition of CI in the ICT domain, proposed by Pierre Lévy in 1995 [43]. Lévy defined CI as

Authors' addresses: S. Suran, V. Pattanaik, and D. Draheim, Tallinn University of Technology, Akadeemia tee 15a, Tallinn, 12616, Estonia; emails: {shweta, vishwajeet.pattanaik, dirk.draheim}@taltech.ee.

0360-0300/2020/02-ART14 $15.00
https://doi.org/10.1145/3368986

a "form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills" [43]. Some of the CI platforms of the early period include WikiWikiWeb, Experts-Exchange, and Google [50]. Since then, advancements in ICT technologies like Web 2.0 [65, 71], Semantic Web [28, 44], and Crowdsourcing [7, 17] have enabled and drastically eased large-scale collaborations over the Internet, leading to the development of well-known CI platforms like WaterWiki [16, 62], Climate CoLab [34, 51], DDtrac [26, 27], WikiCrimes [19, 68], and Goldcorp [4], which facilitate knowledge sharing, problem-solving, and decision-making among individual users and groups, through web-based interactions and collaborations.

The success of these systems can be credited to their underlying architectures or frameworks (hereinafter referred to as "models"). Unfortunately, most of these models are often defined using system-specific elements, principles, attributes, requirements, or their combinations [39], and are based on specific problems [21]. Since each of these CI systems is designed for a specific problem or use-case, the models proposed for these systems are often presented as completely different entities. However, comparing these models shows that although each new CI system and model expands on our current understanding of CI, nevertheless many of these systems bear a few similarities [48]. Sadly, this abundance of diverse knowledge has not yet led to the development of a unified CI model [13, 56, 67] that can support the development of new CI systems based on systematic knowledge rather than intuition [39]. Also, many of the existing CI systems are proprietary and are therefore not available in scientific literature. And, systems that are described in scientific literature, focus more towards the theoretical foundations, usability, and future applications of collective intelligence [21], rather than focusing on the implementation [39]. This lack of well-defined and systematic knowledge about the architecture and principles of the underlying CI systems has led to a reproducibility crisis.

In order to achieve comprehensive knowledge of CI systems, it is imperative that we extensively investigate published scientific literature irrespective of the so-called proposed models. We are convinced that although different CI systems are defined in different ways, they must share more than just a few common characteristics. And, identifying these characteristics could help us to achieve a unified formal model for designing CI systems, irrespective of their application. To this end, we contribute by conducting a first of its kind Systematic Literature Review (SLR) of CI models. In this SLR, we extensively investigate the characteristics of 12 CI models, selected from a pool of 219 scientific publications. And, based on the results of our review, we develop a novel framework that can be utilized to understand existing CI systems. The proposed framework provides a generic model and a set of requisites that would enable creation of novel CI systems, regardless of their domains. This is achieved by exhaustively combining all attributes of the studied CI models into the proposed framework.

Additionally, to better explain the functioning of CI systems with respect to the proposed framework, we examine the different components of six ongoing CI projects: CAPSELLA, hackAIR, openIDEO, Climate CoLab, WikiCrimes, and Threadless.

In particular, we aim at answering the following research questions:

RQ1: What are the underlying models of existing CI systems? What are the common terminologies used to describe CI models? What are their components? And, how are these components associated to each other?

RQ2: Do any of the available CI models appropriately define all CI systems, irrespective of their applications? Can these models be used to create CI systems for novel challenges?

RQ3: If not, then can we somehow combine the available knowledge of CI models and systems to create a unified model that could define all CI systems?

The article is structured as follows. In Section 2, we describe the research methodology used for conducting this SLR. Section 3 presents a brief summary of the selected studies, followed by the aggregated list of terminologies used to describe CI systems in Section 4. In Section 5, we present a novel framework for CI systems and evaluate our generic theoretical CI model by means of comparative case studies in Section 6. Finally, Section 7 presents the threats to validity of the SLR and Section 8 concludes by summarizing the key findings of this article and provides insights for future research.

2 RESEARCH METHODOLOGY

To answer the research questions mentioned in Section 1 through a transparent and objective approach, we decided to conduct this review based on Kitchenham's "Guidelines for performing Systematic Literature Reviews in Software Engineering" [37]. A SLR summarizes, critically appraises, and identifies valid and applicable evidence in available research by using explicit methods to perform thorough literature search [9, 37, 66]. Based on Kitchenham's guidelines, we perform this SLR in five stages:

(1) Search Strategy
(2) Study Selection
(3) Study Quality Assessment
(4) Data Extraction
(5) Data Synthesis

2.1 Search Strategy

Based on the previously identified research questions, we selected a set of search terms. We then used the combination of these search terms to look for relevant research articles in different academic databases. After this, we applied the inclusion criteria on the identified articles and shortlisted the most relevant articles (which we refer to as "Primary Studies"). Following Kitchenham's guidelines, we then evaluated the primary studies using the quality assessment criteria. And finally, the selected studies were investigated in the data extraction and synthesis stages of the SLR.

2.1.1 Search Terms. As researchers often use the terms "Crowdsourcing" and "Wisdom of Crowds" as synonyms for CI [39, 64], we decided to use all three keywords as the primary search terms. And, for the secondary search terms, we used keywords such as model, framework, and others that are commonly used to describe ICT systems. In order to construct the search string we used the following guidelines provided by Kitchenham and Charters [37]:

(1) Derive search terms from research questions and from initial literature review.
(2) Identify synonyms for search terms from scientific literature.
(3) Use the Boolean "AND" and "OR" to link search terms and their synonyms.

The list of identified primary and secondary search terms, and the resulting search string are presented in Table 1.

2.1.2 Academic Databases. The resulting search string was used to search for pertinent articles in four different databases, namely, *ACM Digital Library*, *IEEE Xplore*, *ScienceDirect (Elsevier),* and *Springer.* The search was restricted to articles published since the year 2000; because, the first popular Web 2.0 based CI platform 'Goldcorp' and 'Threadless' were launched in the same year [18, 52, 83]. It was only after this period that CI systems became popular and were recognized as a significant area of research in ICT. In order to identify relevant books, technical reports, and theses, we also conducted a manual search on Google Scholar.

Table 1. Search Terms Identified Based on Research Questions

Primary Search Terms	collective intelligence, wisdom of crowds, crowdsourcing
Secondary Search Terms	model, framework, architecture, requirements, principles, attributes, properties
Search String	("collective intelligence" OR "wisdom of crowds" OR "crowdsourcing") AND ("model" OR "framework" OR "architecture" OR "requirements" OR "principles" OR "attributes" OR "properties")

Table 2. Search Results

Year	Database	Total Count
2000–2017	ACM Digital Library (proceedings, journals, newsletters, and magazines)	1,289
2001–2017	IEEE Xplore (conferences, early access articles, journals, magazines, and books)	1,214
2000–2017	SpringerLink (from sub-discipline "Information Systems Applications incl. Internet": chapters, conference papers, articles, and books)	3,196
2000–2017	ScienceDirect (reviews, research articles, and books)	4,096
1997–2018 *Manual Search*	Google Scholar (research articles, books, reports, and theses)	53
	Total	9,848
	Total (after screening)	9,418

2.1.3 Search Process. During the search process, we found that many of the databases indexed each others' articles, therefore the chances of getting redundant results were high. Thus, to avoid duplicate results, we manually selected different options (like Publication Type, Publisher, etc.) while searching through each database. In total 9,418 articles were identified after removing 430 redundant articles. Table 2 presents the number of relevant articles identified from each academic database.

2.2 Study Selection

To identify the articles relevant to our research questions, we applied a two-phase selection process. During this process, two researchers of this review independently analyzed the identified articles and selected the studies, which were most likely related to our research questions.

2.2.1 Selection Phase 1. In this phase, we studied the titles and abstracts of the identified articles and assessed them on the basis of the inclusion criteria listed in Table 3. After completion of this phase, 216 primary studies (PS) were selected. We then scanned the reference list of the selected primary studies to identify related articles that we might have missed during our initial search. We found three articles which passed our inclusion criteria, and therefore we added these articles to our list of primary studies; making a total of 219 articles (see Table 2 in Supplementary Material).

2.2.2 Selection Phase 2. In this phase, we applied the quality assessment criteria illustrated in Table 4 to the primary studies selected in Selection Phase 1. After completion of this selection

Table 3. Inclusion Criteria for Selection Phase 1

Criteria ID	Inclusion Criteria
IC1	The article explains the theoretical foundations of collective intelligence in computer science.
IC2	The article describes the role of collective intelligence in crowdsourcing and open innovation.
IC3	The article focuses on architecture/frameworks of CI systems.
IC4	The article describes CI systems/applications available on the Web.
IC5	The article focuses on knowledge generation and exchange in crowds.
IC6	The article is related to at least one aspect of our research questions.
IC7	The article should not compare collective intelligence with swarm intelligence and artificial intelligence.

Table 4. Quality Assessment Criteria for Selection Phase 2

Criteria ID	Quality Criteria Check-List
QC1	Are the research objectives clearly defined in the study?
QC2	Does the study propose a new framework, or provide technological details of an existing CI system?
QC3	Is the system architecture/framework/design/experiment clearly defined in the study?
QC4	Is the proposed CI architecture or framework compared to existing CI models or systems?
QC5	Does the study provide insights about the role, importance, and behavior of individuals in the proposed CI system or model?
QC6	Does the study propose novel solutions to crowd management issues in CI?

phase, 12 primary studies were finally selected. These 12 studies were then used for data extraction and data synthesis. We describe both stages further in the next sections.

2.3 Study Quality Assessment

The intention of this phase is to determine the relevance of selected studies while limiting bias in the study selection process. In this phase, all three researchers of this review independently assessed the primary studies by answering the questions presented in Table 4. For each primary study, the researchers answered the questions as "Yes," "Partly," or "No"; scoring each criteria as 1, 0.5, and 0, respectively. The individual scores for each question were then added to derive a total score for each primary study. The studies that scored 3 or higher were finally selected for the data synthesis stage. Any conflict of opinion about the process and results of the quality assessment measures were discussed among all three researchers to reach a consensus. The scores of the remaining 12 primary studies that satisfied the quality assessment criteria are presented in Table 5; followed by the title of the studies and the publication type presented in Table 6.

2.4 Data Extraction and Synthesis

The intention of data extraction stage is to identify the main contributions of the selected studies, and to present a summary of the work. Table 7 presents the data items extracted from the 12

Table 5. Quality Score of Selected Studies

Primary Study ID	QC1	QC2	QC3	QC4	QC5	QC6	Total Score	Selected Study ID
PS1	1	1	0.5	0	0.5	1	4	S1
PS19	1	0	0.5	0	1	1	3.5	S2
PS48	1	1	0.5	0.5	0.5	0	3.5	S3
PS73	1	1	1	1	1	0	5	S4
PS108	1	1	0.5	0	0.5	0	3	S5
PS138	1	0	1	1	0.5	0	3.5	S6
PS154	1	0	1	1	0	0	3	S7
PS155	1	0	1	1	0	0	3	S8
PS156	1	1	1	1	0	0	4	S9
PS173	1	1	1	1	1	0	5	S10
PS174	1	0.5	0.5	0	0.5	0.5	3	S11
PS204	1	1	1	1	0	1	5	S12

Table 6. List of Final Selected Studies

Study ID	Study Title	Publication Type
S1	"Intelligent Collectives: Theory, Applications and Research Challenges" [58]	Journal Article
S2	"Leadership and the Wisdom of Crowds: How to Tap into the Collective Intelligence of an Organization" [53]	Journal Article
S3	"Modelling the Index of Collective Intelligence in Online Community Projects" [73]	Conference Paper
S4	"The Role of Collective Intelligence in Crowdsourcing Innovation" [67]	PhD Thesis
S5	"Collective Intelligence Model: How to Describe Collective Intelligence" [21]	Conference Paper
S6	"Collective Intelligence Systems: Classification and Modeling" [48]	Journal Article
S7	"Designing for Collective Intelligence" [27]	Journal Article
S8	"On Model Design for Simulation of Collective Intelligence" [70]	Journal Article
S9	"A Resource Allocation Framework for Collective Intelligence System Engineering" [81]	Conference Paper
S10	"Harnessing Crowds: Mapping the Genome of Collective Intelligence" [52]	Journal Article
S11	"Leveraging the Power of Collective Intelligence through IT-enabled Global Collaboration" [32]	Journal Article
S12	"Collective Intelligence: A Keystone in Knowledge Management" [3]	Journal Article

selected studies and Section 3 presents a summary of the same. The contributions, i.e., models and elements of the Selected Studies, are presented in Table 16.

The goal of the data synthesis stage is to collate and summarize the contributions of the selected studies. In Section 4, we first catalog the definition types and classifications of the studied CI models; we then identify all unique and synonymous *characteristics*, *levels*, *requirements*, *properties*, and *building blocks* and classify them into 24 distinct attributes (presented in Table 16). Finally,

Table 7. Data Extracted from Selected Studies

Extracted Data Item	Description
Study Title	See Table 6.
Author(s)	See Table 2 in Supplementary Material.
Year	See Table 2 in Supplementary Material.
Publication Title	See Table 2 in Supplementary Material.
Publication Type	See Table 6.
Source/Publisher	See Table 2 in Supplementary Material.
Summary	See Section 3.

Table 8. Criteria for Collective to be Intelligent (as Presented by Nguyen et al. in S1) [58]

Criteria	Description
Diversity	Individuals must belong to diverse backgrounds, knowledge bases, and so forth.
Independence	Freedom for individuals to act according to their choice, without others influence.
Decentralization	Facilitate individualism and assure diversity in individuals.
Aggregation	Appropriate methods to integrate individual solutions [58].

based on these findings, we then answer the first two research questions (RQ1 and RQ2) in Section 4 and the final research question (RQ3) in Section 5.

3 SUMMARY OF SELECTED STUDIES

3.1 S1 (Van Du Nguyen et al. 2018)

The aim of this study is to define the criteria necessary for a collective to be intelligent. To do so, the study [58] presents a novel general CI framework based on crucial attributes of a collective.

Influenced by Bonabeau's [4] concept of Decisions 2.0, which is defined as "a new era of decision-making in which the traditional decision-making process is supported using the wisdom of crowds through collaboration and collective intelligence" [58], Nguyen et al. state that "collective intelligence is considered as the power of Decisions 2.0" [58]. Based on this premise, the study proposes a CI framework based on characteristics vital for an intelligent collective, as proposed by Surowiecki [77]. According to Nguyen et al., a collective must fulfill four criteria (presented in Table 8) [58] to be intelligent. And, based on these characteristics, the authors propose a general framework of CI (namely, *Collective*, *Aggregation Methods*, and *Collective Performance Measures*) [58] for wisdom of crowds.

3.1.1 Diversity. A collective must be diverse, as a heterogeneous group of individuals can provide new knowledge and diverse viewpoints to any given problem. Nguyen et al. further categorize diversity as"diversity in the composition of collective members" [58] and "diversity of individual predictions in a collective" [58]. To explain diversity, the authors used an example of weather forecasting; where accurate weather predictions are a difficult task even if relying on experts. The authors claim that such prediction problems could be solved more easily if multiple individuals were allowed to add extra information and provide different perspectives to solve the problem [58].

3.1.2 Independence. The individuals in a CI system must be allowed to provide their own inputs and their decisions should not be influenced by others [58, 77]. This is important, because information cascades can diminish the intelligence of the collective [1].

3.1.3 Decentralization. This criteria helps individuals act independently, while avoiding others' influence, and thus ensures diversity [58]. To explain this, the authors used the example of Linux, where solutions to specific problems are selected from a pool of solutions submitted by independent programmers from around the world.

3.1.4 Aggregation. This criteria provides the appropriate mechanism to integrate the opinions and solutions provided by the individuals [58]. Examples of such new aggregation methods include prediction markets [82] and social tagging [86].

3.2 S2 (Kurt Matzler et al. 2016)

The aim of this study is to present activities necessary to promote collective intelligence within organizations. The proposed activities are based on the work of Surowiecki [77] and are explained using case studies and real world examples [53].

In this study, Matzler et al. argue that although platforms such as wikis, blogs, prediction markets, and so forth, might be enough to harness the wisdom of the crowd from end users, such platforms are inadequate to support collective intelligence within organizations. The authors propose that in order to harness the power of collective intelligence within organizations, it is imperative that managers follow the following steps: "create cognitive diversity" [53], "promote independence" [53], "access decentralized knowledge" [53], and "effectively aggregate knowledge" [53].

3.2.1 Cognitive Diversity. To explain cognitive diversity, Matzler et al. refer to the work done by Page [61]. Page states that cognitive diversity can be explained as a combination of "diverse perspectives" [53], "diverse interpretations" [53], "diverse heuristics" [53], and "diverse predictive methods" [53]. Matzler et al. explain the relevance of these attributes in organizations by the means of two case studies; namely, "How diversity can drive innovation" [30] and "The CEO's role in business model reinvention" (a case study from Infosys Technologies Limited) [24].

3.2.2 Promote Independence. Matzler et al. emphasize the importance of this step, by explaining how lack of independence or peer pressure may force employees to convey incorrect or sugar-coated information to their managers, which may lead to biased decisions [53]. The authors suggest that managers should create an atmosphere of open dialog where all employees can share their honest opinions and ideas; the authors recommend techniques like the PreMortem exercise [38] to create such an independent environment.

3.2.3 Access Decentralized Knowledge. In regard to this step, Matzler et al. state how, in the past, knowledge was organized hierarchically in organizations; where as now, due to globalization, decentralization, and data ubiquity, knowledge within organizations is not limited to the organizations themselves [53]. In other words, when looking for novel solutions and ideas, organizations now rely heavily on participants via online contests, social media platforms, blogs, and wikis [74]. Matzler et al. argue that organizations could boost their internal collective intelligence by allowing their employees to tap into this decentralized knowledge aggregated from the social web [53]. Employees could then use this knowledge to come up with ideas and solutions to support the organization's growth, while being aligned with the organization's vision and mission.

3.2.4 Effectively Aggregate Knowledge. The final step for promoting collective intelligence within organizations is to effectively aggregate dispersed knowledge. In this study, Matzler et al. briefly discuss techniques (such as averaging individual opinions) that could be utilized to aggregate knowledge from different sources [53]. The authors further describe this step using the examples of predictive markets and peer review systems, which have been found to be effective knowledge aggregation techniques [29]. Lastly, Matzler et al. discuss the effectiveness of Wikipedia's peer review system [53], by comparing the accuracy of its knowledge base to that of Britannica, as investigated by Jim Giles [22].

Table 9. Levels for Assessing CI Potential (as Presented by Skarzauskiene et al. in S3) [73]

Level	Description
Capacity Level	Describes possible user actions, both as an individual and as a member of the community [48]. It also includes massive participant interactions [47] that promote knowledge creation and innovation [3].
Emergence Level	Describes a system state [48] that supports self-organizing, "emergent" behavior, "swarm effect" [47], and mechanic development [3].
Social Maturity Level	Describes the clarity of system goals [3] and community/individual objectives [48].

3.3 S3 (Aelita Skarzauskiene et al. 2015)

The aim of this study is to propose measures to quantify the minimum potential required by community projects, necessary to transform them into CI systems. The authors do so by investigating the trends in engagement and participation in online communities in Lithuania. Skarzauskiene et al. conduct both qualitative and quantitative research by extensively interviewing 20 individuals and by conducting a public opinion survey with 1,022 Lithuanian participants between the ages of 15 and 74 [73]. Finally, the authors propose three levels/measures a community project must fulfill in order to be considered as a CI system [73].

Before conducting qualitative and quantitative research, Skarzauskiene et al. briefly analyzed several CI frameworks proposed by researchers. Based on the literature, Skarzauskiene et al. proposed a conceptual framework for assessing the potential of CI [73]. The authors define the proposed conceptual framework in three levels, presented in Table 9 [73]. Using the proposed levels, combined with results of qualitative and quantitative analysis, the authors calculate a CI Potential Index, which they claim could assist developers and initiators of community projects by helping to assess the CI potential of such projects [73].

3.4 S4 (Juho Salminen 2015)

The aim of this doctoral thesis is to explore the role of collective intelligence in crowdsourcing innovations [67]. Salminen's work is motivated by the fuzzy nature of CI, which has led to different interpretations of the concept including "wisdom of crowds" [77] and "swarm intelligence" [5]. In his work, Salminen attempts to defuzzify the notion of collective intelligence by investigating its emergence as a complex-adaptive system [67].

To do so, the author conducted a systematic literature review of published case studies discussing three CI platforms (OpenIDEO, Quirky, and Threadless). He then observed user behavior on each of these platforms for over a month, and gathered relevant data including web clips, diary entries, and statistics. Salminen also conducted a literature review of available CI frameworks, based on which he proposed a new theoretical framework for CI. Finally, he evaluated the proposed framework based on his observations from the previously analyzed CI platforms. Salminen defines the proposed framework through three levels of abstraction [67]:

- *Micro*: "enabling factors of collective intelligence."
- *Emergence*: "from local to global."
- *Macro*: "output of the system and wisdom of crowds" [67].

Table 10 presents the elements of Salminen's proposed theoretical framework based on themes from literature. Apart from the proposed theoretical CI framework, Salminen also highlights the crucial issue of biased feedback. When observing the previously mentioned CI platforms, the author found that participants would often create multiple accounts to vote up their own ideas and

Table 10. Themes and Elements of the CI Theoretical Framework (as Presented by Salminen in S4) [67]

Level	Theme	Elements of the Theoretical Framework	References
Micro	Humans as social animals	Human capabilities for interaction*	
	Intelligence		
	Personal interaction capabilities		[12, 35, 85]
	Trust		[2]
	Motivation		
	Attention		
	Communities [67]		[8, 10]
Emergence	Complex adaptive systems	Agents, activities, feedback, emergence	[60]
	Self-organization	Agents, activities, feedback	[36]
	Emergence	Emergence	[14]
	Swarm intelligence	–	
	Stigmergy	Agents, activities, feedback, distributed memory	[78]
	Distributed memory [67]	Distributed memory	[6]
Macro	Decision-making	Output	
	Wisdom of crowds	Output	[77]
	Aggregation	–	
	Bias	–	
	Diversity	–	[31]
	Independence [67]	–	

* Same for all themes in micro level.

would demoralize their competitors by providing negative/incorrect feedback and down-votes. Salminen states that to prevent such issues, researchers must create measures to evaluate the accuracy of crowd decisions [67].

3.5 S5 (Sandro Georgi et al. 2012)

The aim of this study is to build a comprehensive model based on available literature while recognizing the characteristics that describe CI [21].

Georgi et al. draw attention to a very important issue in the field of CI, i.e., that research about the topic in general is very limited, as most available research is application- and type-specific. The authors state that although numerous scientific articles and reports have been published about CI platforms, frameworks, and models, only little research has been done on "how to describe collective intelligence in general" [21]. To fill this gap, the authors first studied the existing scientific literature and choose three models of CI, namely, "the collective intelligence genome" by Malone et al. [52], "mitigating biases in decision tasks" by Bonabeau [4], and "the collective intelligent system" by Lykourentzou et al. [49]. Combining these three models, the authors propose five novel characteristics and argue that these can appropriately describe collective intelligence. Table 11 presents these characteristics and their descriptions as stated by Georgi et al. [21].

3.6 S6 (Ioanna Lykourentzou et al. 2011)

This study aims to design a modeling process that can identify the common characteristics of CI systems. Additionally, the process helps to identify challenges that prevent the construction of a generic CI system [48].

Lykourentzou et al. claim that their work is the first attempt in classifying the common shared characteristics of CI systems. The authors state that although all CI systems may seem to be

Table 11. Characteristics that Define CI (as Presented by Georgi et al. in S5) [21]

Characteristics	Description
Objective of task	Can be described as the outcome that the CI intends to achieve. These objectives can be categorized as "create" (creation of knowledge or ideas or physical objects) and "decide" (correctness or best or most suitable, respectively).
Size of contribution	Represents the amount or volume of contribution, and can vary depending upon the complexity of the problem and form/structure of the CI.
Form of input	Can be presented in the form of rules or data/information (pictures, text, datasets, etc.), and can be categorized as instructions, challenge descriptions, or raw material.
Form of output	Can be of two types: knowledge (i.e., intangible) or products (i.e., tangible).
Stakeholder	Defines stakeholders of a CI system based on their roles. "Initiators" are those whose objective is to reach a desired goal. "Contributors" do the actual work and use their intelligence to provide solutions. Finally, "beneficiaries" are those who profit from the outcomes of such systems [21].

Table 12. Common Characteristics that Define a CI System (as Presented by Lykourentzou et al. in S6) [48]

Characteristics	Description
Set of possible individual actions	Set of actions that an individual is allowed to perform when contributing (in some form or another) within the system.
System state	Set of minimal variables that completely define the system.
Community and individual objectives	List of goals that a community or an individual intends to achieve by using the system.
Expected user action function	Effort expected from users, necessary to achieve individual/community goals.
Future system state function	Expected state of the system after some time, given the system's current state and user actions.
Objective function	Measures the extent to which individual/community goals of the system have been achieved [48, 81].

substantially different from each other, they all seem to share quite a few characteristics. After analyzing published literature on CI, Lykourentzou et al. proposed that all CI systems could be categorized as either "active" or "passive" systems. Additionally, "active" CI systems could further be classified into "collaborative," "competitive," or "hybrid" systems [48]. The authors suggest that in "passive" CI systems, groups of users would exhibit behavior of swarms, irrespective of whether the system requires such a behavior or not. Whereas in "active" systems, crowd behavior is created and coordinated by the system itself [48].

Lykourentzou et al. further state that based on this classification, CI systems have several common attributes (described in Table 12) [48]. The authors also highlight issues of "critical mass," "task and workload allocation," and "motivation" that should be considered when designing CI systems. Finally, Lykourentzou et al. model three types of CI systems (*Collaborative*: Wikipedia and open source software development communities; *Competitive*: Innocentive, BootB, DesginBay,

Table 13. Requirements for CI Applications (as Presented by Gregg in S7) [27]

Requirements	Description
Task-specific representation	CI applications should support task-specific views depending upon the application domain.
Data is the key	The effectiveness of CI applications is directly proportional to its data quality and quantity, and therefore should facilitate data collection and sharing among its users.
Users add value	CI applications should help users to improve the usefulness of data, by providing mechanisms that enable user-oriented addition, modification, or enhancement of data.
Facilitate data aggregation	Keeping the importance of data in mind, CI applications should be designed with necessary features that enable data aggregation throughout the duration of the systems' use.
Facilitate data access	CI applications should offer services and mechanism that facilitate reuse of data outside the application.
Facilitate access for all devices	CI applications should provide services that are usable not just with PCs and internet servers, but also portable devices like PDAs, smart-phones, and tablets.
The perpetual beta	New features must be added to CI applications from time to time, depending upon the community needs and requirements [27].

DARPA Network Challenge; *Passive*: vehicular ad-hoc networks) using the previously identified attributes [48].

3.7 S7 (Dawn G. Gregg 2010)

The aim of this study is to demonstrate the requirements for designing CI applications. Gregg states that a CI application harnesses the knowledge of its users by facilitating human interaction and decision-making, and therefore, new CI applications must center around the importance and use of user-defined data [27]. Inspired by the work of O'Reilly [79], Gregg proposes seven key requirements for CI applications (described in Table 13) [27].

To illustrate how these requirements could be used to design CI applications, the author developed the "DDtrac" application for children with special needs. The application was intended to support decision-making in special education and therapy. DDtrac is a web-based CI application and has two main objectives: first, the application facilitates communication between therapists and teachers so that they could share information about the needs of the children; second, the application allows data collection and provides tools for data analysis to understand a child's progress and to determine adjustments necessary for a better development of the child. The application was deployed for a duration of 18 months with one autistic student and his teachers and therapists. After the conclusion of the trial, all participants reported that the application successfully achieved both its core objectives and helped to improve the academic performance of the student [27].

3.8 S8 (Martijn C. Schut 2010)

This study aims to provide systematic guidelines and instructions for development of CI models, irrespective of the developer's domain. To come up with these guidelines the author first conducted a number of research studies and identified key contributions which distinguish CI systems from other ICT systems. Based on the literature, Schut compiled a list of properties of CI

Table 14. Properties of CI Systems (as Presented by Schut in S8) [70]

<table>
<tr><th>Types</th><th>Properties</th></tr>
<tr><td>Enabling CI properties</td><td>These properties enable the emergence of collective intelligence in a system.
Adaptivity refers to the capability of a system to change its behavior or structure depending upon the environment.
To understand system behavior, it is important to understand both individual actions and interactions among individuals as a whole. These interactions enable the flow of information within systems.
Individual or system behavior can be described fundamentally using rules. Such rules implicitly represent the relationship between system inputs and outputs [70].</td></tr>
<tr><td>Defining CI properties</td><td>If these properties exist in a system, it can be considered a CI system.
Global-Local are levels which distinguish between aggregation at system and individual level, respectively. This distinction is important for understanding adaptivity and emergence. Adaptivity can occur at local and/or global level, whereas emergence is achieved by going from local to global.
Complex systems must have elements of randomness in order to behave as self-organized critical systems.
Emergence is defined as the principle that “the whole is greater than the sum of its parts” [14], and occurs when moving from “the lowest abstraction level (individual) to the highest abstraction level (system)” [70].
Redundancy means that the system should allow its users to access/visualize available knowledge and information at different locations within the system’s user interface.
Redundant data can make the system robust, as data that are lost due to malfunctions could still be recovered from other sources [70].</td></tr>
</table>

systems [70] presented in Table 14. After this, the author investigated several strands of research such as complex adaptive systems, swarm intelligence, and others, that are often described as being synonymous or at least associated to collective intelligence [70]. Based on the findings, Schut finally proposed a “systematic approach for designing CI system models” [70] and illustrated the proposed methodology using two case studies, namely, the “Chinese Whispering Room” and the “Braitenberg collectivae” [70].

The CI system modeling approach proposed by Schut is divided into three phases, i.e., “system design,” “model design,” and “models” (which are further categorized into “generic,” “system,” and “computer” models) [70]. The components of the “system design” phase are inspired by examples from self-organization, multi-agent systems, and swarm intelligence; whereas the components of the “model design” phase are influenced by the work of van Gigch [80] on system modeling and of meta-modeling.

3.9 S9 (Dimitrios J. Vergados et al. 2010)

The aim of this study is to present a framework that can foster the emergence of CI in web community based platforms. Based on published research, Vergados et al. describe a generic CI system as having three main components, i.e., "human community," "machine intelligence," and "system information" [81].

Vergados et al. argue that although the proposed CI framework may lead to the development of completely different CI systems, all systems would share a number of common characteristics [49]. The authors describe these characteristics as follows [81]:

- **System attributes** (same as described in Table 12)
  - *Set of possible individual actions*
  - *System state*
  - *Community and Individual objectives* [48, 81]
- **Functions** (same as described in Table 12)
  - *Expected community member action functions*
  - *Future system state functions*
  - *Objective functions* [48, 81]
- **Other elements**
  - *Resource allocation algorithms*: These algorithms define the required user actions (depending upon the system state) necessary to reach user/system goals and to maximize the usefulness of the system.
  - *Critical mass*: This indicates "the minimum number of users necessary for the system to function effectively" [81].
  - *Motivation*: A vital factor, important to improve the quantity and quality of user participation in a CI system. [81]

Finally, Vergados et al. evaluate the proposed framework by means of simulation where they analyze how the quality of Wikipedia articles could be improved, if the system was based on the proposed concepts. The authors compare the performance of their approach against the current approach used in Wikipedia by using the mathematical functions of the proposed framework. The authors claim that, based on their framework, a CI-enabled Wikipedia community could significantly improve the quality of articles, while reducing the time required for these articles to reach satisfactory quality [81].

3.10 S10 (Thomas W. Malone et al. 2009)

This study aims to propose a new framework that explains the underlying model of CI systems. To do so, Malone et al. examined 250 web-enabled CI systems; and based on their findings, identified the building blocks or "genes" (analogy adopted from biology) of CI [52]. The authors then classified these building blocks, using two pairs of fundamental questions [52], i.e.,

- "Who is performing the task? Why are they doing it?"
- "What is being done? How is it being done?" [52]

The answers to these questions with respect to *staffing*, *incentive*, *goal*, and *structure/process* were then proposed as the "genes" of CI systems [52]. Malone et al. state that different CI systems could be modeled using the combination and recombination of these building blocks. A brief overview of these "genes" is presented in Table 15 [52].

To explain these genes further, Malone et al. examined four web-enabled CI systems: *Linux*, *Wikipedia*, *InnoCentive*, and *Threadless*. Finally, the authors claim that the "sequences of genes" of

Table 15. Building Blocks of CI (as Presented by Malone et al. in S10) [52]

	Genes	Description
Who?	Hierarchy	In this gene, tasks are assigned to individuals or groups by someone in authority (similar to traditional hierarchical organizations).
	Crowd	In this gene, individuals within the group can indulge in activities if they choose to do so; and there is no authoritative figure [52].
Why?	Money	Financial gain can be a big motivator for individuals in markets and organizations.
	Love	In many situations, emotional states such as love, affection, passion, or simply interest could be a great motivator for participants.
	Glory	Recognition by competitors, colleagues, or general public is another important motivator [52].
What?	Create	In this gene, participants create something like a T-shirt design, a piece of code, or an innovative solution to a given problem.
	Decide	In this gene, participants evaluate and select items from a set of options; primarily, submitted by other participants [52].
How?	*Create*	
	Collection	This gene occurs when participants create solutions independently. A sub-type of this gene is the *contest gene*, which occurs when one or many contributions are recognized as best and are rewarded.
	Collaboration	This gene occurs when participants create solutions as a group, and the proposed solutions are interrelated or interdependent.
	Decide	
	Group	This gene occurs when "members of a crowd make a decision that applies to the crowd as a whole" [52]. Important variants of this gene include *voting*, *consensus*, *averaging*, and *prediction markets*.
	Individual	This gene occurs when members of the crowd make their own independent decisions, which might be influenced by other members but are not necessarily identical. Two important variants of this gene are *markets* and *social media* [52].

each of these systems could be combined into *genomes* that could help us to understand these CI systems better [52].

3.11 S11 (Luca Iandoli 2009)

The aim of this study is to provide a model for management of CI, and to raise issues that must be considered when designing CI systems. Iandoli argues that although there are several open issues in CI, all of these issues could be organized into two macro-areas, i.e., "management of collective intelligence" [32] and "design of collaborative tools" [32].

Iandoli states that online/virtual communities could be viewed as organizations and, therefore, could also be modeled as such. Based on this hypothesis, Iandoli et al. proposed five characteristics of online/virtual communities "when modelled as organizations" [33]:

(1) "Clear goals and objectives" [32, 33] coherent with a predefined mission.
(2) "A large number of participants" [32, 33] who can offer their time and efforts to achieve the system goals (by knowledge sharing, creation, and consensus activities) in return for incentives.

(3) "A set of processes" [32, 33] that allow participants to develop, submit, or evaluate new ideas, artifacts, and decisions by collaborating with others.
(4) "Rules" [32, 33] that govern how participants interact with the system and each another.
(5) "Participant roles and responsibilities" [32, 33].

Iandoli further argues that even if virtual communities are modeled as organizations, such communities would still face major governance issues because of the many differences between virtual communities and real organizations. Three of these issues [32] are the following:

— *Attention governance*: This involves reducing the possibility of premature, incomplete, or biased decisions, caused due to the lack of correct and unbiased knowledge or due to peer pressure.
— *Participation governance*: The system must facilitate and support participation of large numbers of individuals from diverse backgrounds. Participants must be provided with suitable incentives to keep them inspired and motivated to share their information and knowledge, and to help achieve the system objectives in an unbiased fashion.
— *Community governance*: Appropriate rules must be established to enable smooth and stable interactions among participants and communities; the system should be organized hierarchically and individuals should be given clearly defined roles, responsibilities, and incentives [32].

Finally, Iandoli states that even if all the above-mentioned issues are resolved, there would still be two challenges, i.e., "designing proper visualization tools" [32] and "designing trust and reputation appraisal systems" [32] that would have to be dealt with, irrespective of the technologies used when designing such collaborative platforms [32].

3.12 S12 (Andre Boder 2006)

This study aims to establish a new model for CI in organizations. The model is inspired by Nonaka's work on "The Knowledge-Creating Company" [59] and provides insights that enable transformation from tacit to explicit knowledge within and among organizations, from a collective intelligence perspective [3].

Pertaining to literature on knowledge management in organizations, Boder argues that the process of how organizational elements (such as individuals, their expertise, formal and informal networks, methods of communication, and implicit cultural norms) interact to enable knowledge creation, represents a form of CI. Based on this argument, Boder presents the building blocks of organizational collective intelligence [3]:

— Block A (*Development of competencies*), i.e., the first block "is the development of competencies" [3]. Although difficult to realize, organizations should aim to develop complementary competencies. This could possibly be achieved by human resource managers, who should identify individuals with different competencies gained from different situations; and once such individuals are identified, knowledge managers should bring them together so that their competencies complement each other. Doing so, organizations could take advantage of individual competencies and therefore create new knowledge.
— Block B (*Goal development*). The second block "is the development of a common representation of the goals" [3]. Although each group or department within the organization could have its own goals and objective, these goals and their representations should be aligned with the organizations overall objectives and should be coherent.

— Block C (*Mechanic development*). The third block "is the development and alignment of processes into mechanics of interactions between entities involved" [3], i.e., *organizations*. The formal and informal norms of the organization must be stated explicitly; additionally, employees should respect each others expectations and should trust each others competencies, because such a culture would enable smooth articulation when dealing with new problems or challenges [3].

To illustrate how these building blocks could be used in the process of building CI, Boder breaks down these actions into six groups, and describes six generic tools that could be utilized to apply these actions [3]. He then uses these tools and actions to describe three scenarios: "the value chain" [3], "co-integration of key competencies to achieve a critical medical mission" [3], and "innovative problem-solving" [3]. Finally, the author concludes by stating that organizations must create novelty to survive and evolve. And this is only possible if organizations build collective intelligenceCI by combining the know-how of their employees and integrate organizational knowledge with partner organizations by "coordinating their respective value chains" [3].

4 DATA SYNTHESIS

In this section, we look at the different definitions and classifications of elements that describe a CI model, as proposed by the studies discussed in Section 3. Looking at all these elements, it is clear that different authors define CI systems using different terminologies such as *characteristics*, *levels*, *requirements*, *properties*, and *building blocks*; however, a deeper examination of these models proves that each of these definition types propose similar (if not the same) concepts. Similarly, many of the selected studies explain CI from different perspectives (such as CI in organizations, CI as self-organizing systems, and others); however, the characteristics of CI presented in these studies are very much alike. Table 16 presents the list of all characteristics proposed in the selected studies and classifies them into 24 *unique* attributes (described in Section 5) according to their definitions (described in Section 3). It is important to note here, that some of the selected studies have proposed combinations of characteristics from previous research; and therefore, are presented as combinations of attributes in Table 16.

Based on the findings of the data extraction and data synthesis stages, we now answer the first two research questions.

4.1 Research Question (RQ1)

What are the underlying models of existing CI systems? What are the common terminologies used to describe CI models? What are their components? And, how are these components associated to each other?

4.1.1 RQ1.1. What are the Underlying Models of CI Systems?

Literature shows that CI is a multidisciplinary field, drawing concepts and techniques from a number of different disciplines including computer science [23], organizations [25], social media [69], complexity sciences [70], and psychology [84]; therefore, different scholars have described CI from different perspectives. However, over the years only three definitions of CI have been widely adopted in ICT; two of which were proposed in this decade. The first formal definition of collective intelligence (in ICT) was proposed by Pierre Lévy (1997) [43], followed by Jerome C. Glenn (2013) [23] and Thomas W. Malone (2015) [50]. Although each of the definitions describes CI in its own distinct way, nevertheless, when examined together, the definitions express CI as having three main components, i.e., *individuals* (with data/information/knowledge); *coordination and collaboration activities* (according to a predefined set of rules); and *means/platform for real-time*

Table 16. Terminologies Used (in S1–S12) to Describe CI Systems and Attribute ID(s) of Their Respective Classifications

Study ID	Definition Type	Classification	Sub-classification	Attribute ID(s)
S1	Characteristics	Diversity		A1
		Independence		A2
		Decentralization		A1, A2
		Aggregation [58]		A16
S2	Steps	Cognitive diversity		A1
		Promote independence		A2
		Access decentralized knowledge		A17
		Effectively aggregate knowledge [53]		A16
S3	Levels	Capacity level	Set of possible individual actions	A18
			Massive participant interaction	A19
			Competencies development	A7
		Emergence level	System state	A20
			Self-organizing	A8
			Emergent behavior	A9
			Mechanic development	A10
		Social maturity level	Community and individual objectives	A11
			Goal development [73]	A12
S4	Levels	Micro-level	Humans as social animals	A13
			Personal interaction capabilities	A19
			Trust	A10
			Motivation	A3
			Attention	A19, A3
			Communities	A4
		Level of emergence	Complex adaptive systems	A8, A9
			Self-organization	A8
			Emergence	A9
			Swarm intelligence	A8, A9
			Stigmergy	A8, A9
			Distributed memory	A17
		Macro-level	Decision-making	A13
			Wisdom of crowd	A13
			Aggregation	A16
			Diversity	A1
			Independence [67]	A2

(Continued)

Table 16. Continued

Study ID	Definition Type	Classification	Sub-classification	Attribute ID(s)
S5	Characteristics	Objective of a task		A12
		Size of contribution		A5
		Form of input		A21
		Form of output		A21
		Stakeholder [21]		A4
S6	Characteristics	Set of possible individual actions		A18
		System state		A20
		Community and individual objectives		A11
		Critical mass		A5
		Task and workload allocation		A14
		Motivation [48]		A3
S7	Requirements	Task-specific representation		A22
		Data is the key		A23
		Users add value		A6
		Facilitate data aggregation		A16
		Facilitate data access		A17
		Facilitate access for all devices		A17
		The perpetual beta [27]		A8, A9
S8	Properties	Enabling CI properties	Adaptivity	A8, A9
			Interactions	A19
			Rules	A10, A21
		Defining CI properties	Global-local	A16
			Randomness	A8
			Emergence	A9
			Redundancy	A22
			Robustness [70]	A24
S9	Characteristics	System attributes	Community and individual objectives	A11
			Set of possible individual actions	A18
			System state	A20
		Other elements	Resource allocation algorithms	A14
			Critical mass	A5
			Motivation [81]	A3

(Continued)

Table 16. Continued

Study ID	Definition Type	Classification	Sub-classification	Attribute ID(s)
S10	Genes	Staffing	Crowd	A4
			Hierarchy	A4
		Incentive	Extrinsic motivation	A3
			Intrinsic motivation	A3
		Goal	Create	A13
			Decide	A13
		Structure/Process	Collection	A15
			Collaboration	A15
			Group Decision	A15
			Individual Decision [52]	A15
S11	Characteristics	Clear goals coherent with mission		A12
		Large number of motivated participants		A5, A3
		A set of processes		A15
		Rules		A10, A21
		Roles and responsibilities [32]		A18
S12	Building Blocks	Competencies development		A7
		Goal development		A12
		Mechanic development [3]		A10

communication (viz., hardware/software). When combined, these components enable intelligent behavior in groups or crowds.

Table 17 is the result of segregating all the characteristics defined in Section 3 in terms of the just discussed three main components of CI systems.

4.1.2 RQ1.2. What are the Common Terminologies Used to Describe CI Models?

As suggested in the selected studies, CI models have been described using terminologies such as *characteristics* (S1, S5, S6, S,9 and S11), *steps* (S2), *levels* (S3 and S4), *requirements* (S7), *properties* (S8), *genes* (S10), and *building blocks* (S12). And, each of these terminologies is further segregated into different classification and sub-classifications as described in Section 3. However, as mentioned in the previous sections, the terminologies used in these models describe similar concepts, and therefore can be classified into unique attributes as presented in Table 16.

4.1.3 RQ1.3. What are the Components of CI Models? And, How are These Components Associated to Each Other?

Typically the components of ICT systems are classified as data, hardware, software, information, procedures, and people. However, since the selected studies describe CI models by the means of their characteristics, these characteristics can be interpreted as the components of CI models. Based on the definitions of CI [23, 43, 50], we can segregate these characteristics/attributes and their relationship into the three main components of CI as described in Section 4.1.1 and presented in Table 17.

Table 17. Unique Attributes of CI (from S1to S12) Segregated According to the Components of CI

Component	Characteristics	Attr. ID	Study ID(s)
Individuals (with data, information, knowledge)	Diversity	A1	S1, S2, S4
	Independence	A2	S1, S2, S4
	Motivation	A3	S4, S6, S9, S10, S11
	Crowd	A4	S4, S5, S10
	Critical mass	A5	S5, S6, S9, S11
	Users add value	A6	S7
Coordination and collaboration activities (according to a predefined set of rules)	Competencies development	A7	S3, S12
	Self-organization	A8	S3, S4, S8
	Emergence	A9	S3, S4, S8
	Trust and respect	A10	S3, S4, S8, S11, S12
	Community and individual objectives	A11	S3, S6, S9
	Clear goals and objectives	A12	S3, S5, S11, S12
	Wisdom of crowd	A13	S4, S10
	Task and workload allocation	A14	S6, S9
	Set of processes	A15	S10, S11
Means for real-time communication (viz., hardware/software)	Aggregate knowledge	A16	S1, S2, S4, S7, S8
	Access to decentralized knowledge	A17	S2, S4, S7
	Roles and responsibilities	A18	S3, S6, S9, S11
	Massive interactions	A19	S3, S4, S8
	System state	A20	S3, S6, S9
	Predefined input/output types	A21	S5, S8, S11
	Task-specific representation	A22	S7, S8
	Data is key	A23	S7
	Robust	A24	S8

4.2 Research Question (RQ2)

Do any of the available CI models appropriately define all CI systems, irrespective of their applications? Can these models be used to create CI systems for novel challenges?

Comparing all characteristics of the studied CI models (see Table 16) with the components of CI systems (described in Section 4.1.1 and presented in Table 17), we see that none of the studied models have all 24 unique attributes, and therefore cannot define all CI systems completely. However, the existing models provide insights that can assist in planning when designing a CI system; and point out challenges that would have to be solved in order to achieve a robust and adaptive CI system. Most authors themselves state that their proposed CI models only describe collective intelligence in specific domains (S2, S3, S4, S9, S11, and S12), and suggest that further research and investigation is required to gain a better understanding of generic CI systems (S1, S3, S6, S8, S10, and S11). Therefore, although particular CI models can be used to define CI systems for specific domains, the same models might not be as useful when designing CI systems from other disciplines.

Furthermore, since the proposed models are evaluated using either quantitative/qualitative interviews (S3), or case studies (S4, S8), or examples from scenarios (S12), or simulations (S9), or applications/systems built based on the models (S6, S7, S10), it is not possible to identify a single model that can be used (in its current state) to design CI systems for novel challenges. For now, the most generic CI model available in literature is the one proposed by Malone et al. (S10) [21]; however, this highly cited and accepted model needs to be developed further for a deeper and more accurate understanding of CI [21, 52, 76].

5 A NOVEL FRAMEWORK FOR CI

Using the findings from the data extraction phase of the SLR, we now attempt to contribute to the available CI models by proposing a unified framework for CI by combining the 24 unique attributes (see Table 17) of CI models identified from studies S1–S12. The purpose of the proposed framework is to answer the final research question (RQ3) and provide additional insights and explanations that can help us better understand CI systems in general. In order to evaluate the proposed "generic" CI model, we will compare the model to multiple CI systems, each designed for a different objective and belonging to different disciplines (see Section 6).

RQ3. Can we somehow combine the available knowledge of CI models and systems to create a unified model that could define all CI systems?

We combine the knowledge of the CI models studied in this SLR, and propose a novel framework that describes CI systems in a fine-grained manner. We do so by comprehensively classifying all components of the studied CI models into 24 unique attributes (see Table 16), and then categorizing them into three sections:

— a "generic" model that defines all CI systems;
— additional requisites for CI systems; and
— CI as a complex adaptive system.

While taking inspiration from the building blocks for CI proposed by Malone et al. [52], combined with the findings from Section 4.1.1, we propose a model that describes CI systems by the means of *staff*, *process*, *goal*, and *motivation*. Designed as an extension to Malone's concept of building blocks, the proposed generic model segregates the originally proposed *genes* into more fine-grained *types*; introduces a new classification, namely, *interactions*; and suggests vital *properties* for the *staff* and *goal* building blocks of the generic model. Finally, remaining attributes that could not be accommodated into the building blocks are aggregated into the additional requisites category.

5.1 A Generic Model for CI Systems

As mentioned in Section 3, Malone's genome model for collective intelligence [52] is based on two pairs of questions: *"Who is performing the task? Why they are doing it?"* and *"What is being done? How is it being done?"* [52]. Based on these questions, the authors proposed the analogy of *genes* categorized as staffing, incentives, goal, and process. Each of these categories was subdivided into individual *genes* which, when combined, created the genome of CI systems. Drawing from the literature, we decided to move away from the concept of genes, and rather examine the proposed *genes* as *types*. Doing so, we realized that the available *genes* could be segregated into new types and sub-types. And, while some of the *genes* could be better understood as *interactions* between *types*, others could be explained as necessary *properties* inherent to these new *types*.

5.1.1 Who is Performing the Task? The *staff* in CI are the actors who perform different tasks within the system (as suggested in S10). As literature suggests, these actors or individuals must interact with each other based on certain rules depending upon the structure (hierarchical/non-hierarchical) of the system. And, when viewed as a collective, the *staff* of a CI system must exhibit a specific set of properties for the system to function effectively.

— *Types*: The actors or individuals in a collective are the first component of a CI system, and therefore play a vital role in describing how the system functions (A6). Typically, these actors (A4) can be segregated on the basis of their roles and responsibilities (A18) within the system. Drawing insights from S4, S5, and S10, we determine that actors in a CI system can be classified as follows:

- *Passive actors* or beneficiaries are individuals who aim to gain from the outputs produced by the CI system, but do not wish to contribute in the problem solving process. These beneficiaries could either be stakeholders who are financially motivated, or end-users who simply want to exploit the knowledge produced by the system (but do not wish to actively contribute). Examples of stakeholders in CI systems could be seen in the following projects: Threadless, InnoCentive [15], and GoldCorp [83]. Here, the host organizations crowdsource their problems (designing of T-shirts, research and development, and identifying ideal mining locations, respectively) to the general public, with the intention of using the produced knowledge or artifacts for their own advantage.
- *Active actors* or contributors are individuals who are involved in CI processes (defined in "How"); such actors use their knowledge and expertise and help to create innovative solutions to the given problem. Such contributors can be further divided into two categories, namely, *crowd* and *hierarchy*.
  * *Crowd* in a CI system comprises actors who actively contribute new knowledge, information, or artifacts to the system. Such actors are allowed to carry out a predefined set of actions, based on concrete sets of rules and regulations; however, there is no authoritative figure that has direct control over the actors' individual actions. Examples of crowd in CI systems can be seen in the following projects: Climate CoLab [34], WikiCrimes [19, 68], and WeKnowIt [42] where users contribute data and information about the weather, crimes, and disasters, and also help verify the authenticity of the accumulated knowledge. Whereas, in projects such as Threadless [15], members of the crowd contribute by creating new artifacts and deciding on the best.
  * *Hierarchy* in a CI system comprises administrators and experts who are responsible for allocating tasks to the crowd. While the administrators monitor crowd behavior in the system and make sure that the community and individual goals of the collective are achieved, the experts analyze and verify the contributions of the crowd. Additionally, in some cases the experts also help in identifying the best contributions or solutions. An ideal example of such a hierarchy can be seen in WikiCrimes: institutional agents, monitor agents, reputation agents, and others are responsible for different administrative activities within the system [19, 68].

- *Properties*: To ensure that a collective exhibits intelligent behavior, the collective of actors in a system must have a few crucial properties. According to S1, S2, and S4, a CI system must promote diversity and independence among its actors, as this can enable the creation of novel solutions exploiting knowledge from individuals familiar with multiple domains and with different experiences. Also, these actors should be allowed to act independently, as this can help to get rid of peer pressure, and therefore to reduce user-generated bias. Finally, to enable an effective collective intelligence, a collective must have critical mass or a minimum number of actors as suggested in S5, S6, S9, and S11.
  - *Diversity* (A1) in CI systems refers to the heterogeneous nature of actors, who belong to different age groups, genders, and educational, financial, and cultural backgrounds. This is important as, such diverse actors can provide diverse pieces of knowledge, perspectives, interpretations, and experiences; and this could lead to the creation of innovative solutions and better decisions. An example of the advantages of diversity in actors can be seen in InnoCentive [15]: organizations with small R&D groups crowdsource their problems to acquire new and innovative ideas.

   — *Independence* (A2) means that the opinions of one actor should not be influenced by the opinions of others. Independence among actors is vital, as it can help to avoid information cascades where users pass information that they assume to be true (without appropriate evidence or knowledge), and therefore make irrational choices and decisions [1, 46, 55].
   — *Critical mass* (A5) in collectives is defined as the minimum number of actors who must participate in system processes for the system to function effectively. Although studies suggest that critical mass is an imperative property that enables effective creation and constant exchange of diverse knowledge and information, the concept needs to be investigated further as critical mass in different CI systems can often depend upon the system goals and objectives.
— *Interactions*: Interactions in CI systems can either exist between two or more actors, or among actors and the contributions of others. Such interactions can be categorized as follows:
   — *Trust and respect* (A10) are two preconditions for cooperation. When dealing with new problems or challenges, actors in a collective must treat each other with respect and should trust each others' abilities and competencies, as doing so can enable smooth and efficient flow of knowledge and information within the system.
   — *SECI*: "Socialization, Externalization, Combination, and Internalization" [59] (A7) are the four components of Nonaka's model for knowledge creation in organizations [59]. Using these knowledge dimensions, organizations can convert their employees' tacit knowledge into explicit organizational knowledge and back. Since the SECI model was originally designed to promote sustainable innovation in organization, these concepts can also be utilized in CI systems to enable competency development in actors (as suggested in S12 and [11]).

  Finally, as suggested in S3, S4, and S8, a CI system must support such interactions in massive volumes (A19).

5.1.2 Why They are Doing It? Motivation (A3) in CI systems is essential to maintain user engagement and encourage participation. Depending upon on the objectives of a system, users in a CI system could be motivated by their desire to gain knowledge (as in Wikipedia [45]) by money and glory (as in Threadless [8] and InnoCentive [57]) or by social cause (as in hackAIR [54]). According to Malone et al. (S10), money, love, and glory can be considered high-level motivations for people participating in CI systems [52]; whereas, Vergados et al. (S9) categorize motivation as tangible, intrinsic, and self-fulfilling [81]. Combining the recommendations from S4, S6, S9, S10, and S11, we categorize motivation as *intrinsic* and *extrinsic*.

— *Intrinsic* motivations such as social cause, interest, passion, and self-fulfillment encourage actors in a collective to collaborate and contribute for the betterment of the community or its individuals. An example of such motivation can be seen in DDtrac: school teachers and therapists collaborate to understand a child's needs and determine necessary adjustments in teaching techniques for better development of children with special needs [26, 27].
— *Extrinsic* motivations are factors external to CI tasks that encourage actors to contribute in hopes of getting rewards. Such motivations can be either *tangible* like money and trophies or, *intangible* like fame and glory. In CI projects like Threadless [8], InnoCentive [57], and Goldcorp [83] participants are offered cash rewards and prizes for submitting ideas and designs; whereas, in WikiCrimes [19] participants gain a reputation based on the reliability their contributions.

5.1.3 What is Being Accomplished? Unlike Malone's gene model (S10) that attempts to answer the question of "What is being done?" (from an organizational perspective), we decided to focus on the question of "What is being accomplished?" for our proposed model. Based on the literature, we found that our question is a better fit, as it could appropriately define the different types of objectives/goals (of CI systems) presented as characteristics in several selected studies. In general, these *goals* can be defined as "observable and measurable desired results bound to one or more objectives, that have to be achieved by committed actors within a finite time-frame." Since collective intelligence initiatives are typically motivated by *community* or *individual* objectives (A11) as suggested by S3, S6, and S9, we segregate CI goals into the two aforementioned *types*. These types can be seen again in Threadless: individuals with a niche in T-shirt designing participate in competitions to present their contributions to the community, learn from others' feedback, and earn money; whereas, the community's goal is to bring new T-shirt designs to the marketplace by choosing and popularizing trending designs [8]. Additionally, drawing from the contributions of S3, S5, S11, and S12 the requisite properties of these CI system goals could be categorized as *well-defined* and *objective* (A12).

5.1.4 How is It Being Done? Malone et al. categorized the processes in CI systems as combinations of dependent-independent and create-decide activities, where the create-decide activities answered the question "What is being done?" [52]. In our proposed model, however, we describe CI *processes* (A15) as *types* of activities and *interactions*. As literature suggests, the activities can be either *create*, where actors come up with new ideas or design new artifacts; or it can be *decide*, where actors express their likes or dislikes for a particular subject or artifact. Since both of these activities can be either be done by individual actors or groups of actors;, these activities could also be viewed as *dependent* or *independent* interactions. To add more granularity to process types, create activities can be further classified into *contest* (S10) and *voluntary*. As the names suggest, *contest* create activities are carried out in competitive environments and are extrinsically motivated, whereas *voluntary* create activities are intrinsically motivated. It is the combination of these three types (decide, contest, and voluntary) and interactions (dependent and independent) that enables intelligence in collectives (A13).

— *Collection* (i.e., *create* plus *independent*): In such activities or processes, actors participate as individuals and their contribution to the system is a result of their independent work. An example of *collection through contest* can again be seen in Threadless: individuals compete for cash rewards by creating and submitting new T-shirt designs [8]. Whereas, in WikiCrimes, actors contribute through *voluntary collection* by reporting criminal activities they witness in their local vicinity [19].
— *Collaboration* (i.e., *create* plus *dependent*): Such activities are carried out by groups of actors or communities where multiple individuals work together as a single entity and create new ideas or products.

  As an instance for *voluntary collaboration*, we can again look at DDtrac: therapists and teachers work together to maximize the learning outcomes of students with special needs [27]. Similarly, in hackAIR, volunteers from NGOs conduct workshops to build citizen interest in the hackAIR platform and educate them on how they could become a part of the project's community and help to gather air quality data from their local vicinity [40]. Whereas, *collaboration in contests* is seen in openIDEO, where multiple participants work as a team and propose solutions to societal challenges, in hopes of getting financial rewards [67].
— *Individual decision* (i.e., *decide* plus *independent*): Such decisions are made by individuals acting as independent entities and can be different for different actors. However, in some

cases these decisions may be influenced by the information provided by other actors. For instance, in Threadless the members of the community independently vote for T-shirt designs submitted by the participants. Unfortunately, as suggested by Salminen in S4, in some cases participants tend to create multiple accounts with the intension to down-vote their competitors, thereby influencing other members and generating biased community feedback [67].

— *Group decision* (i.e., *decide* plus *dependent*): In such activities, decisions are made by multiple individuals as a group or a community, and the outcome of the decisions impacts the community as a whole. For instance, such consensus can be seen in Threadless: the employees of the organization review the T-shirt designs chosen by the community and finally decide which designs to produce and award [8].

5.1.5 Input and Output. The final component of a CI system is the *flow of information*, or form of input/output (A21), and can be explained as interactions between the "who" and the "how" of the system. The flow of information starts from the *actors* who are responsible for providing inputs like individual knowledge and experiences, data from sensors, or end-user opinions and feedback from social media platforms. The collected inputs are then processed using different activities in "how," and the results of these activities are then presented back to the *actors* who now take new decisions or produce new artifacts based on this newfound knowledge. Since this flow of information between the *actor*s and the *process*es of the CI system is so vital, we decided to add it to our generic CI model.

The aggregation of the aforementioned components is illustrated as the proposed "generic" model for CI systems, in Figure 1.

5.2 Additional Requisites

Although any CI system can be described as the combination of the above-mentioned components, there are a few additional requisites that must exist in a CI system for the system to work effectively.

— *System state* (A20). This can be expressed as the minimum set of variables that completely define a CI system. As discussed in S3, S6, and S9 the system state can include challenges/issues raised by the members of the community, the identified solutions, activities of the users, and the system resources. Since our proposed model defines CI systems as the combination of different processes, actors, motivations, and goals, unique combinations of the same can be used to express the system state of a CI system.

— *Data is the key* (A23). "Collective intelligence draws on user-generated content and sharing of information, knowledge and ideas" [69] and, therefore, data or information/knowledge provided by members of the collective is a vital component of a CI system. For a CI system to be able to reach its goals, the system must allow its users to collect, manipulate, and share large volumes of data; this can enable robust innovations and decisions.

— *Aggregate knowledge* (A16). Since the effectiveness of a collective intelligence relies primarily on user-generated data/information, CI systems must have mechanisms and processes that aggregate this data/information. These aggregation processes are important as information provided by the community can often come from a variety of sources and could be incorrect or biased [20]. Aggregating the information, however, could help resolve conflicting information and could therefore allow for better innovations and reliable decisions. Additionally, systems should also provide mechanisms that allow users to aggregate their knowledge by means of social tagging (for information retrieval), collaboration (for exchange of vocabularies), and task-specific representation.

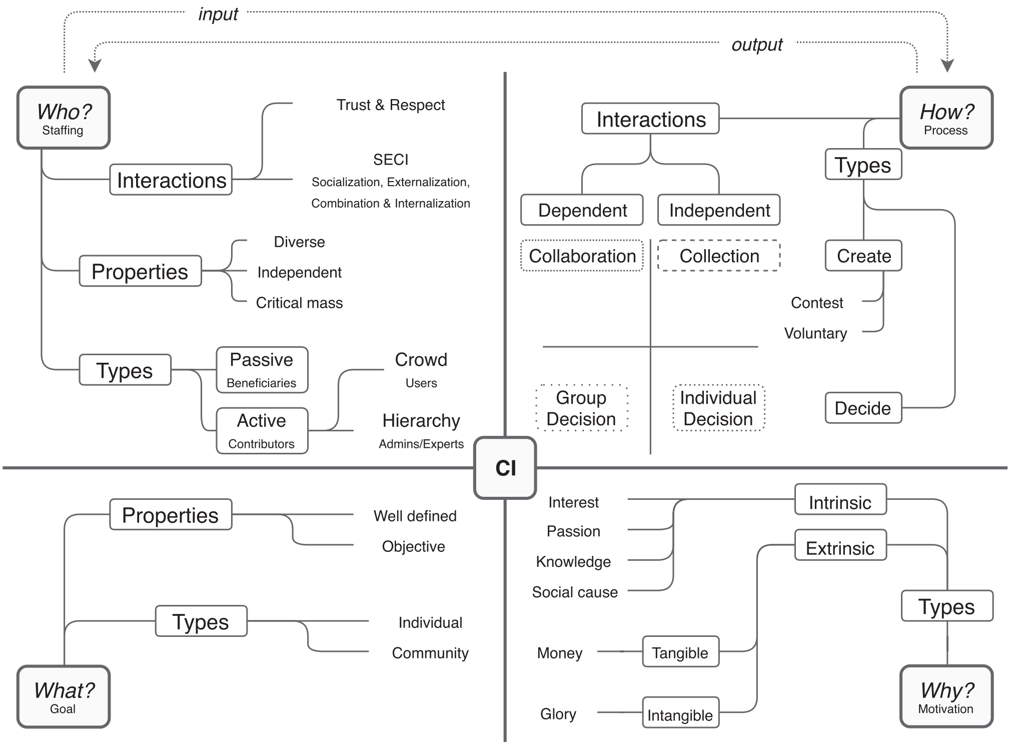


Fig. 1. Generic model for collective intelligence systems.

— *Access to decentralized knowledge* (A17). Thanks to the growing number of internet users, more and more people are able to communicate, collaborate, and share information on the Web. Keeping user interest in mind, it is important for CI systems to allow users from different parts of the world to participate and gain from the knowledge or artifacts generated by the system. To do so, the system must facilitate access across multiple devices like PCs, laptops, smart phones, servers, and others. Furthermore, CI systems should support open data and open innovation practices, and should allow data access to users even outside the system.
— *Task and workload allocation* (A14). Another important aspect that should be kept in mind when designing CI systems is the methods for coordination and resource allocation. When designing CI systems, the expected tasks of different actor types should be predefined; and based on these tasks, the rules and extent of interactions among actors and actor's access to the aggregated knowledge must be outlined. For instance, participants should be allowed to add new solutions and view solutions submitted by others; however, they should not be allowed to make changes to others' contributions without the contributor's consent. On the other hand, system administrators should have complete access to the data/information/knowledge produced within the system.
— *Task-specific representation* (A22). To support knowledge creation and enable fluid information exchange among actors from asynchronous groups, CI systems should provide task-specific representations like tables, charts, histograms, plots, and knowledge graphs. Additionally, depending on the task or problem, the system should allow its users to visualize the same knowledge/information in different forms.

— *Robust* (A24). Finally, since CI systems are designed as complex systems with multiple components, actors, users, and resources; it is important for such systems to be able to handle redundant and erroneous inputs. In addition to this, the system should also have appropriate mechanisms for data/information/knowledge backup and recovery, in case of a system crash or malfunction.

5.3 CI as Complex Adaptive System

CI systems are complex by nature [70] and should be able to adapt to their environments, making such platforms complex adaptive systems (as suggested in S4 and S8). However, for a system to be complex adaptive, the system must exhibit adaptivity, self-organization, and emergence [60, 67, 70, 75].

Adaptivity means that the system or its components should allow constant changes over the period of its existence, depending upon the needs of its collective [70]. System developers should regularly update and evolve the platform by bringing in new technologies and services, based on user feedback and requirements, under the condition that these requirements are aligned with the system goals and objectives.

Self-organizing (A8) [47, 67] means the systems should be able to organize and re-organize its internal structure without the need of an external control [36, 72]. This behavior could be facilitated by allowing the creation of communities, where each member of the community would have a reputation that they could gain by providing useful contributions (in the form of insights, knowledge, or artifacts) and through up-votes/stars given to them by other members of their community. Such a reputation model can help create a structure within these communities, and therefore further interactions between such communities can lead to self-organizing behavior within the system.

Emergence (A9) in a system occurs when simple interactions among low-level system components give rise to new and unexpected patterns or properties, disparate from the properties of the system as a whole (based on the definition of emergence proposed by Damper [14]). In adaptive and self-organizing systems, regular modifications to the system and ever-changing user behavior may lead to the creation of unforeseen patterns, properties, or outcomes, thereby exhibiting emergent behavior.

6 COMPARATIVE CASE STUDIES

In this section, we evaluate the proposed generic model from Section 5.1 by examining six CI platforms with respect to the aforementioned model. The CI platforms were chosen on the basis of the following criteria: the platforms should belong to different disciplines/domains, the systems should be available for use (during the time of study), the platforms should have been published/discussed in scientific literature, the deliverables of the platforms should be available online, and lastly, the platforms should be recent or ongoing.

Based on these criteria, we identified six CI platforms (see Table 1 in the Supplementary Material). To analyze the platforms, we created user profiles on each of the platforms and observed system processes for create and decide activities; over the duration of 6 months, i.e., starting January 2018 to the end of June 2018. During this period, we interacted with the system as passive users. We created projects/ideas to analyze the creation process; however, we never submitted the projects/ideas for evaluation. We observed submissions for other participants and feedback from active users, and analyzed how the system communities and hierarchies work synchronously to come up with new contributions and innovative ideas. Additionally, we studied the available technical reports, scientific publications, FAQs, and other useful resources for each of the platforms. Aggregating our observations, we found that different aspects each of the six CI platforms could

Table 18. List of Studied CI Platforms and Their "What"

CI Platform	Year	What*	Domain
CAPSELLA	2016–2018	*IG*: Learn about new ICT technologies that can help improve agronomic practices. *CG*: Develop new ICT solutions, software and applications, and promote start-ups that can provide such solutions for agrifood business and farmers.	Agrobiodiversity
hackAIR	2016–2018	*IG*: Learn about the concentration of air pollutants (especially particulate matter) in cities and its effect on the health of local residents. *CG*: Provides citizens with real-time information about air pollution levels in their local vicinity and enables conversations for possible improvements in air quality.	Air pollution
openIDEO	Ongoing since 2010	*IG*: Demonstrate their skills and expertise to solve complex challenges, and learn from others' work. *CG*: Tackle global challenges by developing innovative solutions using human-centric collaboration activities.	Innovation platform
Climate CoLab	Ongoing since 2009	*IG*: Participate in initiatives to help reach global climate goals. *CG*: Collaborate with other communities and experts, and help design/choose solutions to help identify sustainable growth initiatives.	Climate change
WikiCrimes	Ongoing since 2008	*IG*: Report criminal incidents. And keep track of crime rates in the local vicinity. *CG*: Assist governing bodies in validating reports of crimes provided by individuals. Help maintain a public record of all criminal activities.	Crime monitoring
Threadless	Ongoing since 2000	*IG*: Showcase their artistic ability by creating new T-shirt design. *CG*: Express community interest and select best T-shirt designs. Bring new and trending T-shirt designs to the marketplace.	Apparel design

* *IG*: Individual goals. *CG*: Community goals.

be described using our proposed generic model. Tables 18, 19, and 20 present the "What," "Who," and "Why-" "How" for each of the platforms, respectively.

6.1 What?

The goals of the six CI platforms can be summarized as follows:

The *CAPSELLA* project is designed to enable the creation of new ICT solutions for farmers and agricultural experts. The platform focuses on ICT contributions that facilitate the collection and exchange of data and experiences from individuals working in agriculture and bio-diversity.

Table 19. Comparative of CI Platforms—"Who"

CI Platform	Who: Active (Crowd)	Who: Active (Hierarchy)	Who: Passive/Beneficiaries	Open Data (*Yes/No*)
CAPSELLA	Farmers, food and seed communities	Agro-ecology, agri-food, ICT experts	Farmer communities, technology providers, other organizations	*Yes*
hackAIR	Citizens, open source communities	Environmental/ health/educational organizations, scientific communities	Enterprises, local governments	*Yes*
OpenIDEO	Participants, innovators, alliances	Experts, challenge sponsors, advisory board	Participants (who only wish to participate in workshops)	*No*
Climate CoLab	Participants, community members	Fellows, judges	Government bodies, business organizations, civil society, individual citizens, consumers	*No*
WikiCrimes	Citizens	Agents, news media, government agencies	Citizens, government agencies	*No*
Threadless	Designers, consumers	Organization (Threadless)	Consumers	*No*

Table 20. Comparative of CI Platforms—"Why" and "How"

CI Platform	Why: Intrinsic	Why: Extrinsic		How*: Create		How*
	Interest (I)/*Passion* (P)/ *Knowledge* (K)/*Social cause* (S)	*Tangible*	*Intangible*	*Contest* (CL/CB)	*Voluntary* (CL/CB)	Decide (ID/GD)
CAPSELLA	IKS	Money	-	CL	Both	GD
hackAIR	IPKS	hackAIR sensors	Points, badges	CL	Both	GD
OpenIDEO	IPKS	Money	Glory	CL	Both	Both
Climate CoLab	IPKS	Money	Points	Both	CL	Both
WikiCrimes	IKS	-	Reputation	-	Both	Both
Threadless	IPK	Money	Design Quotient	CL	CL	Both

* CL: *Collection.* CB: *Collaboration.* ID: *Individual decision.* GD: *Group decision.*

hackAIR is designed as a platform where citizens can collect and access information about air quality in different parts of the world. The system empowers citizens by providing openly available DIY sensor designs, tool-kits and tutorials, thereby enabling citizens to be a part of the data collection process.

Similar to hackAIR, the *openIDEO* and *Climate CoLab* platforms deal with climate change and other environmental/societal challenges. However, both of these platforms are designed to enable the creation of new and innovative solutions by means of collaboration. While the contributions in Climate CoLab are focused toward global climate change goals, the contributions in openIDEO are focused more toward open innovation practices for societal change.

The *WikiCrimes* platform allows residents to anonymously report criminal activities in their local vicinity. This is especially useful in countries where citizens are not willing to contact the law enforcement agencies due to fear or lack of trust. The platform also allows its users to track the frequency and scale of criminal activities in different areas, thereby helping users in making better decisions when visiting specific locations.

Finally, the *Threadless* platform is meant for e-commerce and focuses on retail of apparels. The platform enables artists and designers to showcase their talent by sharing their T-shirt designs with the community. The best designs are then made available for sale on the Threadless marketplace, thereby providing artists and designers with a means of income.

We further elaborate the goals of these CI platforms as *individual* and *community goals* and domains in Table 18.

6.2 Who?

Table 19 presents the different actors of the analyzed CI platforms, segregated into three categories, namely, *active (crowd)*, *active (hierarchy)*, and *passive/beneficiaries* based on our proposed generic model. The table also indicates whether the platforms provide open data for future research or not.

6.3 Why and How?

Table 20 illustrates how each of the analyzed CI platforms motivates different kinds of actors using different sets of intrinsic and extrinsic motivators, and how different kinds of actors carry out different create and decide activities based on their roles within the system.

After mapping our observations (from each of the platforms) to our generic model, we found some interesting relationships between *actor* types, their *motivations*, and their *activities*:

- *Decide* activities are typically *intrinsically* motivated.
- *Contest (create)* activities by individuals of the *active (crowd)* are always *extrinsically* motivated. Whereas, *voluntary (create and decide)* contributions by individuals of the *active (crowd)* are always *intrinsically* motivated.
- *Voluntary (create)* contributions can be of two types: as data or information contributed by *crowd*, as in CAPSELLA, hackAIR, and WikiCrimes, or as feedback and suggestions given by *crowd* and members of *hierarchy* to help improve participants' contributions like in OpenIDEO, Climate CoLab, and Threadless.

7 THREATS TO VALIDITY

The primary threats to the validity of this Systematic Literature Review include bias in search strategy, bias in selection process, and inaccuracies in data extraction.

The selection of studies relied on the search strategy, which included the selection of search terms and literature resources, and the search process. The search terms were selected based on both the research questions and an initial literature review; followed by a three-step process to construct the search string as described in Section 1. We then chose four prominent academic databases of computer science and used the formulated search string to identify relevant literature. Table 2 presents the number and types of research articles identified from each of the academic databases. To avoid bias in our search strategy and to identify relevant technical reports, books, and theses, we conducted a manual search on Google Scholar.

To avoid bias in the study selection process, we first reviewed the titles and abstracts of the identified studies and then selected only those studies that fulfilled the inclusion criteria. We then studied these selected articles and manually checked their references to make sure that we did not miss any relevant articles during the search process. Finally, the selected studies were then evaluated based on the quality assessment criteria. As a result of the study selection phase, we were able to identify the most relevant studies with respect to our research questions.

To eliminate inaccuracies in data extraction, each primary study was independently studied by all researchers and any disparities in findings were resolved through discussions. During the process, we found two pairs of studies, i.e., S1, S2 and S6, S9, which shared a couple of similarities. The first pair (S1, S2) described the characteristics of CI systems using similar classifications, while the second pair (S6, S9) was written by the same authors. By consensus, we decided to keep both pairs in our selected studies, as S1 and S2 described CI systems from different perspectives, whereas S6 and S9 provided different contributions.

8 CONCLUSION

The objective of this article was to analyze different collective intelligence models described in the scientific literature and to identify a generic model that could be utilized to design new CI platforms. To this end, we conducted a Systematic Literature Review, in which we identified 9,418 articles on collective intelligence models. Out of these articles, we selected 12 studies based on an exhaustive selection process. We then critically analyzed these selected studied and found that none of the models provided a generic view of CI systems, as each of the models was designed based on specific perspectives. And, the models that could potentially be used to design domain independent CI systems lacked granularity and needed to be researched further. So, to fill this research gap, we aggregated the components of the CI models described in the selected studies and proposed a unified framework for understanding CI systems. The proposed framework describes CI systems in three parts. First, a generic model, which describes CI systems as a combination of goals, staff, motivation, and processes, which are further described as types, interactions, and properties. Second, a list of requisites necessary for CI systems to work effectively. And third, guidelines that could enable complex adaptive behavior in CI platforms.

To evaluate if the proposed model could define CI systems from different domains, we selected a set of ongoing CI projects and observed user activities within the platform, over a duration of 6 months. After this, we systematically organized our observations and segregated them according to the different components of our proposed generic model. We found that our model successfully described the components of each of the CI platforms and revealed some interesting relations between the types of actors, their activities, and motivations. The evaluation of the proposed model also gave us the opportunity to present our unified CI framework by means of examples (i.e., six ongoing CI initiatives). It was imperative that we describe the components of these CI platforms in terms of the proposed CI model, so that both researchers and system designers/developers in the field could utilize our novel model to design and develop new CI systems. The 24 unique attributes that describe the proposed framework could provide initial insights to system designers and developers, and could be beneficial during the requirement elicitation process when developing new CI systems. We recognize that we need to further examine the proposed framework by comparing it to a larger set of CI platforms, as doing so would help us gain a deeper understanding about how the proposed framework could be used to design new CI systems. Additionally, we would like to evaluate the proposed framework by conducting qualitative interviews with domain experts and researchers working on upcoming CI initiatives. And finally, we would also like to investigate different trust and reputation models that could be utilized to reduce user bias within CI platforms, thereby enhancing user experience and enabling a smooth exchange of knowledge and information within communities.

REFERENCES

[1] Lisa R. Anderson and Charles A. Holt. 1997. Information cascades in the laboratory. *The American Economic Review* 87, 5 (1997), 847–862. http://www.jstor.org/stable/2951328.

[2] Kirsimarja Blomqvist. 1997. The many faces of trust. *Scandinavian Journal of Management* 13, 3 (1997), 271–286. DOI: https://doi.org/10.1016/S0956-5221(97)84644-1

[3] Andre Boder. 2006. Collective intelligence: A keystone in knowledge management. *Journal of Knowledge Management* 10, 1 (2006), 81–93. DOI: https://doi.org/10.1108/13673270610650120

[4] Eric Bonabeau. 2009. Decisions 2.0: The power of collective intelligence. *MIT Sloan Management Review* 50, 2 (Winter 2009), 45–52.

[5] E. Bonabeau and C. Meyer. 2001. Swarm intelligence—A whole new way to think about business. *Harvard Business Review* 79, 5 (May 2001), 106+.

[6] T. Bosse, C. M. Jonker, M. C. Schut, and J. Treur. 2006. Collective representational content for shared extended mind. *Cognitive Systems Research* 7, 2–3 (2006), 151–174. DOI: https://doi.org/10.1016/j.cogsys.2005.11.007

[7] Daren C. Brabham. 2008. Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence* 14, 1 (2008), 75–90. DOI: https://doi.org/10.1177/1354856507084420

[8] Daren C. Brabham. 2010. Moving the crowd at Threadless. *Information, Communication & Society* 13, 8 (2010), 1122–1145. DOI: https://doi.org/10.1080/13691181003624090

[9] David Budgen and Pearl Brereton. 2006. Performing systematic literature reviews in software engineering. In *Proceedings of the 28th International Conference on Software Engineering (ICSE'06)*. ACM, New York, NY, 1051–1052. DOI: https://doi.org/10.1145/1134285.1134500

[10] Romina Cachia, Ramón Compañó, and Olivier Da Costa. 2007. Grasping the potential of online social networks for foresight. *Technological Forecasting and Social Change* 74, 8 (2007), 1179–1203. DOI: https://doi.org/10.1016/j.techfore.2007.05.006

[11] Kyung Jin Cha, Yang Sok Kim, Byeonghwa Park, and Choong Kwon Lee. 2015. Knowledge management technologies for collaborative intelligence: A study of case company in Korea. *International Journal of Distributed Sensor Networks* 11, 9 (2015), 368273. DOI: https://doi.org/10.1155/2015/368273

[12] Cary Cherniss. 2010. Emotional intelligence: Toward clarification of a concept. *Industrial and Organizational Psychology* 3, 2 (2010), 110–126. DOI: https://doi.org/10.1111/j.1754-9434.2010.01231.x

[13] Prerna Chikersal, Maria Tomprou, Young Ji Kim, Anita Williams Woolley, and Laura Dabbish. 2017. Deep structures of collaboration: Physiological correlates of collective intelligence and group satisfaction. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW'17)*. ACM, New York, NY, 873–888. DOI: https://doi.org/10.1145/2998181.2998250

[14] R. I. Damper. 2000. Editorial for the special issue on 'Emergent Properties of Complex Systems': Emergence and levels of abstraction. *International Journal of Systems Science* 31, 7 (2000), 811–818. DOI: https://doi.org/10.1080/002077200406543

[15] Miguel de Castro Neto and Ana Espírtio Santo. 2012. Emerging collective intelligence business models. In *MCIS 2012 Proceedings*. Mediterranean Conference on Information Systems. https://aisel.aisnet.org/mcis2012/14.

[16] Tony Diggle. 2013. Water: How collective intelligence initiatives can address this challenge. *Foresight* 15, 5 (2013), 342–353. DOI: https://doi.org/10.1108/FS-05-2012-0032

[17] Anhai Doan, Raghu Ramakrishnan, and Alon Y. Halevy. 2011. Crowdsourcing systems on the World-Wide Web. *Communications of the ACM* 54, 4 (2011), 86. DOI: https://doi.org/10.1145/1924421.1924442

[18] Ketry Gorete Farias dos Passos and Edna Lúcia da Silva. 2012. Effect of collective intelligence in organizations. *Transinformação* 24, 2 (May 2012), 127–136. DOI: https://doi.org/10.1590/s0103-37862012000200005

[19] Vasco Furtado, Leonardo Ayres, Marcos de Oliveira, Eurico Vasconcelos, Carlos Caminha, Johnatas D'Orleans, and Mairon Belchior. 2010. Collective intelligence in law enforcement—The WikiCrimes system. *Information Sciences* 180, 1 (2010), 4–17. DOI: https://doi.org/10.1016/j.ins.2009.08.004

[20] Jing Gao, Qi Li, Bo Zhao, Wei Fan, and Jiawei Han. 2015. Truth discovery and crowdsourcing aggregation: A unified perspective. *PVLDB* 8, 12 (2015), 2048–2049. DOI: https://doi.org/10.14778/2824032.2824136

[21] Sandro Georgi and Reinhard Jung. 2012. Collective intelligence model: How to describe collective intelligence. In *Advances in Intelligent and Soft Computing*. Vol. 113. Springer, 53–64. DOI: https://doi.org/10.1007/978-3-642-25321-8_5

[22] Jim Giles. 2005. Internet encyclopaedias go head to head. *Nature* 438, 7070 (Dec. 2005), 900–901. DOI: https://doi.org/10.1038/438900a

[23] Jerome C. Glenn. 2013. Collective intelligence and an application by the millennium project. *World Futures Review* 5, 3 (2013), 235–243. DOI: https://doi.org/10.1177/1946756713497331

[24] Vijay Govindarajan and Chris Trimble. 2011. The CEO's role in business model reinvention. *Harvard Business Review* 89, 1–2 (2011), 108–14.

[25] Antonietta Grasso and Gregorio Convertino. 2012. Collective intelligence in organizations: Tools and studies. *Computer Supported Cooperative Work (CSCW)* 21, 4 (01 Oct 2012), 357–369. DOI: https://doi.org/10.1007/s10606-012-9165-3

[26] Dawn Gregg. 2009. Developing a collective intelligence application for special education. *Decision Support Systems* 47, 4 (2009), 455–465. DOI: https://doi.org/10.1016/j.dss.2009.04.012

[27] Dawn G. Gregg. 2010. Designing for collective intelligence. *Communications of the ACM* 53, 4 (April 2010), 134–138. DOI: https://doi.org/10.1145/1721654.1721691

[28] Tom Gruber. 2008. Collective knowledge systems: Where the social web meets the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web* 6, 1 (2008), 4–13. DOI: https://doi.org/10.1016/j.websem.2007.11.011

[29] Friedrich August Hayek. 1945. The use of knowledge in society. *The American Economic Review* 35, 4 (1945), 519–530.

[30] Sylvia Ann Hewlett, Melinda Marshall, and Laura Sherbin. 2013. How diversity can drive innovation. *Harvard Business Review* 91, 12 (2013), 30–30.

[31] Lu Hong and Scott E. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101, 46 (2004), 16385–16389. DOI: https://doi.org/10.1073/pnas.0403723101

[32] Luca Iandoli. 2009. Leveraging the power of collective intelligence through IT-enabled global collaboration. *Journal of Global Information Technology Management* 12, 3 (2009), 1–6. DOI: https://doi.org/10.1080/1097198X.2009.10856494

[33] Luca Iandoli, Mark Klein, and Giuseppe Zollo. 2009. Enabling on-line deliberation and collective decision-making through large-scale argumentation. *International Journal of Decision Support System Technology* 1, 1 (Jan. 2009), 69–92. DOI: https://doi.org/10.4018/jdsst.2009010105

[34] Joshua Introne, Robert Laubacher, Gary Olson, and Thomas Malone. 2011. The climate CoLab: Large scale model-based collaborative planning. In *Proceedings of the 2011 International Conference on Collaboration Technologies and Systems (CTS'11)*. 40–47. DOI: https://doi.org/10.1109/CTS.2011.5928663

[35] Caroline Just. 2011. A review of literature on the general factor of personality. *Personality and Individual Differences* 50, 6 (2011), 765–771. DOI: https://doi.org/10.1016/j.paid.2011.01.008

[36] Stuart A. Kauffman. 1993. *The Origins of Order: Self-organization and Selection in Evolution*. Oxford University Press.

[37] Barbara A. Kitchenham and S. Charters. 2007. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. Technical Report EBSE-2007-01. Keele University.

[38] Gary A Klein. 2004. *The Power of Intuition: How to Use Your Gut Feelings to Make Better Decisions at Work*. Crown Business.

[39] A. Kornrumpf and U. Baumöl. 2014. A design science approach to collective intelligence systems. In *2014 47th Hawaii International Conference on System Sciences*. 361–370. DOI: https://doi.org/10.1109/HICSS.2014.53

[40] Evangelos Kosmidis et al. 2018. hackAIR: Towards raising awareness about air quality in Europe by developing a collective online platform. *ISPRS International Journal of Geo-Information* 7, 5 (2018). DOI: https://doi.org/10.3390/ijgi7050187

[41] Hélène Landemore and Jon Elster. 2012. *Collective Wisdom: Principles and Mechanisms*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511846427

[42] Vita Lanfranchi and Neil Ireson. 2009. User requirements for a collective intelligence emergency response system. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology (BCS-HCI'09)*. British Computer Society, Swinton, UK, 198–203. http://dl.acm.org/citation.cfm?id=1671011.1671035.

[43] Pierre Levy. 1997. *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Perseus Books, Cambridge, MA.

[44] Pierre Lévy. 2010. From social computing to reflexive collective intelligence: The IEML research program. *Information Sciences* 180, 1 (2010), 71–94. DOI: https://doi.org/10.1016/j.ins.2009.08.001

[45] Randall M. Livingstone. 2016. Models for understanding collective intelligence on wikipedia. *Social Science Computer Review* 34, 4 (Aug. 2016), 497–508. DOI: https://doi.org/10.1177/0894439315591136

[46] Jan Lorenz, Heiko Rauhut, Frank Schweitzer, and Dirk Helbing. 2011. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* 108, 22 (2011), 9020–9025. DOI: https://doi.org/10.1073/pnas.1008636108

[47] Shuangling Luo, Haoxiang Xia, Taketoshi Yoshida, and Zhongtuo Wang. 2009. Toward collective intelligence of online communities: A primitive conceptual model. *Journal of Systems Science and Systems Engineering* 18, 2 (01 June 2009), 203–221. DOI: https://doi.org/10.1007/s11518-009-5095-0

[48] Ioanna Lykourentzou, Dimitrios J. Vergados, Epaminondas Kapetanios, and Vassili Loumos. 2011. Collective intelligence systems: Classification and modeling. *Journal of Emerging Technologies in Web Intelligence* 3, 3 (Aug. 2011). DOI: https://doi.org/10.4304/jetwi.3.3.217-226

[49] Ioanna Lykourentzou, Dimitrios J. Vergados, and Vassili Loumos. 2009. Collective intelligence system engineering. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems (MEDES'09)*. ACM, New York, NY, Article 20, 7 pages. DOI: https://doi.org/10.1145/1643823.1643848

[50] T. W. Malone and M. S. Bernstein. 2015. *Handbook of Collective Intelligence*. MIT Press. 2015029234 https://books.google.ee/books?id=Px3iCgAAQBAJ.
[51] Thomas W. Malone and Mark Klein. 2007. Harnessing collective intelligence to address global climate change. *Innovations: Technology, Governance, Globalization* 2, 3 (2007), 15–26. DOI: https://doi.org/10.1162/itgg.2007.2.3.15
[52] Thomas W. Malone, Robert Laubacher, and Chrysanthos N. Dellarocas. 2009. Harnessing crowds: Mapping the genome of collective intelligence. *SSRN Electronic Journal* (2009). DOI: https://doi.org/10.2139/ssrn.1381502
[53] Kurt Matzler, Andreas Strobl, and Franz Bailom. 2016. Leadership and the wisdom of crowds: How to tap into the collective intelligence of an organization. *Strategy & Leadership* 44, 1 (2016), 30–35. DOI: https://doi.org/10.1108/SL-06-2015-0049 arXiv:https://doi.org/10.1108/SL-06-2015-0049
[54] Gavin McCrory, Carina Veeckman, and Laurence Claeys. 2017. Citizen science is in the air—Engagement mechanisms from technology-mediated citizen science projects addressing air pollution. In *Internet Science*. Springer International Publishing, Cham, 28–38. DOI: https://doi.org/10.1007/978-3-319-70284-1_3
[55] Lev Muchnik, Sinan Aral, and Sean J. Taylor. 2013. Social influence bias: A randomized experiment. *Science* 341, 6146 (2013), 647–651. DOI: https://doi.org/10.1126/science.1240466
[56] J. Musil, A. Musil, D. Weyns, and S. Biffl. 2015. An architecture framework for collective intelligence systems. In *2015 12th Working IEEE/IFIP Conference on Software Architecture (WICSA'15)*. 21–30. DOI: https://doi.org/10.1109/WICSA.2015.30
[57] Kimberly Nehls. 2015. Crowdsourcing. *Learning, Media and Technology* 40, 1 (2015), 123–126. DOI: https://doi.org/10.1080/17439884.2014.929144
[58] Van Du Nguyen and Ngoc Thanh Nguyen. 2018. Intelligent collectives: Theory, applications, and research challenges. *Cybernetics and Systems* 49, 5-6 (2018), 261–279. DOI: https://doi.org/10.1080/01969722.2017.1418254
[59] Ikujiro Nonaka and Takeuchi Hirotaka. 1995. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press. 94040408 https://books.google.ee/books?id=B-qxrPaU1-MC.
[60] Julio M. Ottino. 2004. Engineering complex systems. *Nature* 427, 6973 (2004), 399. DOI: https://doi.org/10.1038/427399a
[61] S. E. Page and Princeton University Press. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press. 2006044678 https://books.google.ee/books?id=FAFVHnJ7uK0C.
[62] Chloe Parker. 2015. The IWA WaterWiki: An open-access publishing platform providing free content on all aspects of water, wastewater and the environment. *Environment and Urbanization* 27, 1 (2015), 137–138. DOI: https://doi.org/10.1177/0956247814547172
[63] Michael A. Peters and Richard Heraud. 2015. Toward a political theory of social innovation: Collective intelligence and the co-creation of social goods. 3, 3 (2015), 7–23. https://researchcommons.waikato.ac.nz/handle/10289/9569.
[64] Rolf Pfeifer, Jan Henrik Sieg, Thierry Bücheler, and Rudolf Marcel Füchslin. 2010. Crowdsourcing, open innovation and collective intelligence in the scientific method: A research agenda and operational framework. (2010). DOI: https://doi.org/10.21256/zhaw-4094
[65] P. Clint Rogers, Stephen W. Liddle, Peter Chan, Aaron Doxey, and Brady Isom. 2007. WEB 2.0 learning platform: Harnessing collective intelligence. *Turkish Online Journal of Distance Education* 8, 3 (2007), 16–33. http://tojde.anadolu.edu.tr/makale_goster.php?id=348.
[66] Frantz Rowe. 2014. What literature review is not: Diversity, boundaries and recommendations. *European Journal of Information Systems* 23, 3 (01 May 2014), 241–255. DOI: https://doi.org/10.1057/ejis.2014.7
[67] Juho Salminen. 2015. *The Role of Collective Intelligence in Crowdsourcing Innovation*. PhD dissertation. Lappeenranta University of Technology.
[68] H. Santos, L. Ayres, C. Caminha, and V. Furtado. 2012. Open government and citizen participation in law enforcement via crowd mapping. *IEEE Intelligent Systems* 27 (2012), 63–69. DOI: https://doi.org/10.1109/MIS.2012.80
[69] Detlef Schoder, Peter A. Gloor, and Panagiotis Takis Metaxas. 2013. Social media and collective intelligence—Ongoing and future research streams. *KI - Künstliche Intelligenz* 27, 1 (1 Feb. 2013), 9–15. DOI: https://doi.org/10.1007/s13218-012-0228-x
[70] Martijn C. Schut. 2010. On model design for simulation of collective intelligence. *Information Sciences* 180, 1 (2010), 132–155. DOI: https://doi.org/10.1016/j.ins.2009.08.006 Special Issue on Collective Intelligence.
[71] Toby Segaran. 2007. *Programming Collective Intelligence : Building Smart Web 2.0 Applications*. O'Reilly, Beijing/Sebastapol, CA.
[72] V. Singh, G. Singh, and S. Pande. 2013. Emergence, self-organization and collective intelligence—Modeling the dynamics of complex collectives in social and organizational settings. In *2013 UKSim 15th International Conference on Computer Modelling and Simulation*. 182–189. DOI: https://doi.org/10.1109/UKSim.2013.77
[73] Aelita Skarzauskiene and Monika Maciuliene. 2015. Modelling the index of collective intelligence in online community projects. In *International Conference on Cyber Warfare and Security*. Academic Conferences International Limited, 313.

[74] Daniel Stieger, Kurt Matzler, Sayan Chatterjee, and Florian Ladstaetter-Fussenegger. 2012. Democratizing strategy: How crowdsourcing can be used for strategy dialogues. *California Management Review* 54, 4 (2012), 44–68. DOI: https://doi.org/10.1525/cmr.2012.54.4.44

[75] Tim Sullivan. 2014. Embracing Complexity. Retrieved on March 8, 2019 from https://hbr.org/2011/09/embracing-complexity.

[76] Shweta Suran, V. Pattanaik, Sadok B. Yahia, and Dirk Draheim. 2019. Exploratory analysis of collective intelligence projects developed within the EU-Horizon 2020 framework. In *Computational Collective Intelligence*. Springer International Publishing, 285–296.

[77] James Surowiecki. 2005. *The Wisdom of Crowds*. Anchor.

[78] Guy Theraulaz and Eric Bonbeau. 1999. A brief history of stigmergy. *Artificial Life* 5, 2 (April 1999), 97–116. DOI: https://doi.org/10.1162/106454699568700

[79] O'Reilly Tim. 2007. *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. MPRA Paper. University Library of Munich, Germany. https://EconPapers.repec.org/RePEc:pra:mprapa:4578.

[80] John P. van Gigch. 1991. *System Design Modeling and Metamodeling*. Springer US. DOI: https://doi.org/10.1007/978-1-4899-0676-2

[81] Dimitrios J. Vergados, Ioanna Lykourentzou, and Epaminondas Kapetanios. 2010. A resource allocation framework for collective intelligence system engineering. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems (MEDES'10)*. ACM, New York, NY, 182–188. DOI: https://doi.org/10.1145/1936254.1936285

[82] Jennifer H. Watkins. 2007. Prediction markets as an aggregation mechanism for collective intelligence. Retrieved from https://escholarship.org/uc/item/8mg0p0zc.

[83] Sean Wise, Robert A. Paton, and Thomas Gegenhuber. 2012. Value co-creation through collective intelligence in the public sector: A review of US and European initiatives. *VINE* 42, 2 (2012), 251–276. DOI: https://doi.org/10.1108/03055721211227273

[84] Michael A. Woodley and Edward Bell. 2011. Is collective intelligence (mostly) the General Factor of Personality? A comment on Woolley, Chabris, Pentland, Hashmi and Malone (2010). *Intelligence* 39, 2 (2011), 79–81. DOI: https://doi.org/10.1016/j.intell.2011.01.004

[85] Anita Williams Woolley, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science* 330, 6004 (2010), 686–688. DOI: https://doi.org/10.1126/science.1193147

[86] Koji Zettsu and Yasushi Kiyoki. 2006. Towards knowledge management based on harnessing collective intelligence on the web. In *Managing Knowledge in a World of Networks*. Springer, Berlin, 350–357. DOI: https://doi.org/10.1007/11891451_31

Received March 2019; revised October 2019; accepted October 2019

Supplementary Material for: Frameworks for Collective Intelligence: A Systematic Literature Review

SHWETA SURAN, VISHWAJEET PATTANAIK, and DIRK DRAHEIM, Tallinn University of Technology, Estonia

A SUPPLEMENTARY MATERIAL

A.1 Details of Selected CI Platforms

Table 1 presents the short name, full name, URL, types of related publications, and organizations of the examined CI platforms.

Table 1. Details of Examined CI Platforms

Platform	Full name	URL	Resources*	Organization
CAPSELLA	Collective Awareness Platform for Environmentally Sound Land Management based on Data Technologies and Agrobiodiversity	capsella.eu	TR, CP	CORDIS (H2020)
hackAIR	Collective awareness platform for outdoor air pollution	hackair.eu	TR, CP	CORDIS (H2020)
openIDEO	Collaborative platform for the design process	openideo.com	CP	IDEO
Climate CoLab	Climate CoLab	climatecolab.org	CP	MIT CCI
WikiCrimes	WikiCrimes	wikicrimes.org	CP	CNPq
Threadless	Threadless	threadless.com	CP	SkinnyCorp LLC

* TR: *Technical Reports.* CP: *Conference Papers.*

A.2 List of Primary Studies

Table 2 presents the list of 219 primary studies selected after the study section stage (see Section 2.2.1). Each study is denoted by a unique ID, followed by the title of the publication, author(s) name, publication type, year, and publisher.

0360-0300/2020/02-ART14 $15.00
https://doi.org/10.1145/3368986

Table 2. List of Primary Studies and Publication Details

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS1	Intelligent collectives: Theory, applications, and research challenges	Van Du Nguyen and Ngoc Thanh Nguyen	Journal Article	2018	Taylor & Francis (Google Scholar)
PS2	Literature review on collective intelligence: a crowd science perspective	Chao Yu et al.	Journal Article	2018	Emerald Insight (Google Scholar)
PS3	Geospatial Collective Intelligence for Health Planning: A Case Study for Screening Tests in the City of Esmeraldas, Ecuador	A. F. Jimenez Velez et al.	Conference Paper	2017	IEEE Xplore
PS4	Collaborative Optimization for Collective Decision-Making in Continuous Spaces	Nikhil Garg et al.	Conference Paper	2017	ACM Digital Library
PS5	Collective intelligence: From the enlightenment to the crowd science	Kai Wang et al.	Journal Article	2017	ACM Digital Library
PS6	Awareness Supporting Technologies Used in Collaborative Systems: A Systematic Literature Review	Gustavo Lopez and Luis A. Guerrero	Conference Paper	2017	ACM Digital Library
PS7	Deep Structures of Collaboration: Physiological Correlates of Collective Intelligence and Group Satisfaction	Prerna Chikersal et al.	Conference Paper	2017	ACM Digital Library
PS8	From Crowd to Community: A Survey of Online Community Features in Citizen Science Projects	Neal Reeves et al.	Conference Paper	2017	ACM Digital Library
PS9	Group-in-the-loop : Architecture for Harnessing and Creating Collective Intelligence	Sayed Nafiz Haider et al.	Conference Paper	2017	ICOCI (Google Scholar)
PS10	Improving Health Empowerment and Evidence-Based Decision-Making Using Collective Intelligence and Self-Management Health System: Em-Phasys Project	Pedro Perez-Alcantara and Manuel Herrera-Usagre	Conference Paper	2017	ACM Digital Library
PS11	Prediction Markets as a Vital Part of Collective Intelligence	R. Palak and N. T. Nguyen	Conference Paper	2017	IEEE Xplore
PS12	Research and Trends in the Studies of Collective Intelligence from 2012 to 2016	Francisca Grimon et al.	Conference Paper	2017	SpringerLink
PS13	Computational collective intelligence with big data: Challenges and opportunities	Jason J. Jung	Journal Article	2017	ScienceDirect
PS14	Holism, collective intelligence, climate change and sustainable cities	Monika M. L. dos Santos	Journal Article	2017	ScienceDirect
PS15	Leveraging collective intelligence: How to design and manage crowd-based business models	Karl Tauscher	Journal Article	2017	ScienceDirect

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS16	Optimal incentives for collective intelligence	Richard P. Mann and Dirk Helbing	Journal Article	2017	National Academy of Sciences (IN) (Google Scholar)
PS17	The potential of collective intelligence in emergency medicine: Pooling medical students' independent decisions improves diagnostic performance	Juliane E. Kammer et al.	Journal Article	2017	SAGE (Google Scholar)
PS18	Governing with Collective Intelligence	Tom Saunders and Geoff Mulgan	Book	2017	SpringerLink
PS19	Leadership and the wisdom of crowds: How to tap into the collective intelligence of an organization	Kurt Matzler et al.	Journal Article	2016	Emerald Insight (Google Scholar)
PS20	Do It Yourself Diagnosis: A Study on Acquiring Health-Related Information Online	Duwaraka Murugadas and Sergej Sizov	Conference Paper	2016	ACM Digital Library
PS21	A distributed, collective intelligence framework for collision-free navigation through busy intersections	R. Kalantari et al.	Conference Paper	2016	IEEE Xplore
PS22	A Method Using Collective Intelligence for Communication Activation Among Elderly People Living Alone	Hiroshi Yajima et al.	Conference Paper	2016	SpringerLink
PS23	A research framework of smart education	Zhi-Ting Zhu et al.	Journal Article	2016	SpringerLink
PS24	Internet of things: Applications and challenges in smart cities: A case study of IBM smart city projects	Veronica Scuotto et al.	Journal Article	2016	Emerald Insight (Google Scholar)
PS25	Collective Intelligence and Collaboration: A Case Study in Airline Industry	Sonia A. C. Teixeira et al.	Conference Paper	2016	SpringerLink
PS26	Collective Intelligence Development in Business	Patricia Bouvard and Herv Suzanne	Book	2016	Wiley (Google Scholar)
PS27	Critical review of the millennium project in Nepal	Ashma Vaidya and Audrey L. Mayer	Journal Article	2016	MDPI (Google Scholar)
PS28	Cross-Fertilization of Ideas in Collective Intelligence Model (CIM)	S. S. Gunasekaran et al.	Conference Paper	2016	IEEE Xplore
PS29	Development of Failure Detection System for Network Control Using Collective Intelligence of Social Networking Service in Large-Scale Disasters	Chihiro Maru et al.	Conference Paper	2016	ACM Digital Library
PS30	Fostering Collective Intelligence Education	Jaime Meza et al.	Conference Paper	2016	SpringerLink
PS31	Geospatial Model e-Health Planning Collective Intelligence	A. F. J. Velez et al.	Conference Paper	2016	IEEE Xplore

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS32	Harnessing Crowds to Avert or Mitigate Acts Terrorism: A Collective Intelligence Call for Action	H. Li and A. Hedman	Conference Paper	2016	IEEE Xplore
PS33	Impact of ICT-mediated collective awareness on urban mobility	Kashif Zia et al.	Journal Article	2016	SpringerLink
PS34	Major Variants of the SIS Architecture Pattern for Collective Intelligence Systems	Angelika Musil et al.	Conference Paper	2016	ACM Digital Library
PS35	Managing intellectual capital through a collective intelligence approach: An integrated framework for universities	Giustina Secundo et al.	Journal Article	2016	Emerald Insight (Google Scholar)
PS36	micROS: A morphable, intelligent and collective robot operating system	Xuejun Yang et al.	Journal Article	2016	SpringerLink
PS37	New ICTs for knowledge management in organizations	Pedro Soto-Acosta and Juan-Gabriel Cegarra-Navarro	Journal Article	2016	Emerald Insight (Google Scholar)
PS38	Next Generation Crowdsourcing for Collective Intelligence	John Prpic	Conference Paper	2016	SSRN (Google Scholar)
PS39	The Collective Intelligence Concept: A Literature Review from the Behavioral and Cognitive Perspective	S. S. Gunasekaran et al.	Conference Paper	2016	IEEE Xplore
PS40	Towards Air Quality Estimation Using Collected Multimodal Environmental Data	Anastasia Moumtzidou et al.	Conference Paper	2016	SpringerLink
PS41	Collective Intelligence Systems	Jerome C. Glenn	Book Section	2016	SpringerLink
PS42	Collective decision making, leadership, and collective intelligence: Tests with agent-based simulations and a field study	Kristie A. McHugh et al.	Journal Article	2016	ScienceDirect
PS43	Collective intelligence in project groups: Reflections from the field	Morten Juel Hansen and Hajnalka Vaagen	Journal Article	2016	ScienceDirect
PS44	Collective intelligence in medical diagnosis systems: A case study	Gandhi S. Hernández-Chan et al.	Journal Article	2016	ScienceDirect
PS45	Impact of a collective intelligence tailored messaging system on smoking cessation: The perspect randomized experiment	Rajani Shankar Sadasivam et al.	Journal Article	2016	JMIR (Google Scholar)
PS46	Models for understanding collective intelligence on Wikipedia	Randall M. Livingstone	Journal Article	2016	SAGE (Google Scholar)
PS47	The Horizon 2020 CAPSELLA Project: Collective Awareness PlatformS for Environmentally-Sound Land Management based on Data TechnoLogies and Agrobiodiversity	Mariateresa Lazzaro et al.	Article	2016	CAPSELLA (Google Scholar)
PS48	Modelling the Index of Collective Intelligence in Online Community Projects	Aelita Skarzauskiene and Monika Maciuliene	Conference Paper	2015	ICCWS (Google Scholar)

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS49	A Case Study on Collective Intelligence based on Energy Flow	K. Hassani et al.	Conference Paper	2015	IEEE Xplore
PS50	All Together Now: Collective Intelligence for Computer-Supported Collective Action	G. Valetto et al.	Conference Paper	2015	IEEE Xplore
PS51	An Architecture Framework for Collective Intelligence Systems	J. Musil et al.	Conference Paper	2015	IEEE Xplore
PS52	An Intelligent Collaborative Environment for Sharing Information in a Blood Supply Network	Oscar Salviano, Silva Filho, and Frederic Andres	Conference Paper	2015	ACM Digital Library
PS53	Collective intelligence: Analysis and modelling	Erika Suarez Valencia et al.	Journal Article	2015	Emerald Insight (Google Scholar)
PS54	Collective Intelligence Approach for Free Software Adoption by Municipalities	Jarbas Lopes Cardoso et al.	Conference Paper	2015	ACM Digital Library
PS55	Collective Intelligence in Computer-Mediated Collaboration Emerges in Different Contexts and Cultures	David Engel et al.	Conference Paper	2015	ACM Digital Library
PS56	Collective Intelligence or Group Think?: Engaging Participation Patterns in World Without Oil	Nassim JafariNaimi and Eric M. Meyers	Conference Paper	2015	ACM Digital Library
PS57	Collective Intelligence Support Protocol	Alexandru Senciuc et al.	Conference Paper	2015	SpringerLink
PS58	Collective intelligence systems and an application by The Millennium Project for the Egyptian Academy of Scientific Research and Technology	Jerome C. Glenn	Journal Article	2015	ScienceDirect
PS59	Coordinating Measurements for Air Pollution Monitoring in Participatory Sensing Settings	Alexandros Zenonos et al.	Conference Paper	2015	International Foundation for Autonomous Agents and Multiagent Systems (Google Scholar)
PS60	Encouraging Collective Intelligence for the Common Good: How Do We Integrate the Disparate Pieces?	Douglas Schuler et al.	Conference Paper	2015	ACM Digital Library
PS61	Health care crowds: Collective intelligence in public health	John Prpic	Journal Article	2015	SSRN (Google Scholar)
PS62	Path Planning in GPS-Denied Environments: A Collective Intelligence Approach	P. Chattopadhyay et al.	Conference Paper	2015	IEEE Xplore
PS63	SIS: An Architecture Pattern for Collective Intelligence Systems	Juergen Musil et al.	Conference Paper	2015	ACM Digital Library
PS64	The Interface Design for Serendipity with Collective Intelligence	Makiko Harada and Hidenori Watanave	Conference Paper	2015	ACM Digital Library

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS65	Using Collective Intelligence to Generate Trend-Based Travel Recommendations	S. Schlick et al.	Conference Paper	2015	IEEE Xplore
PS66	Using Collective Intelligence to Support Multi-Objective Decisions: Collaborative and Online Preferences	D. Cinalli et al.	Conference Paper	2015	IEEE Xplore
PS67	Collective intelligence to solve creative problems in conceptual design phase	Rene Lopez Flores et al.	Journal Article	2015	ScienceDirect
PS68	Using the collective intelligence for inventive problem solving: A contribution for open computer aided innovation	Rene Lopez et al.	Journal Article	2015	ScienceDirect
PS69	Collective intelligence applied to legal e-discovery: A ten-year case study of Australia franchise and trademark litigation	Charles V. Trappey and Amy J. C. Trappey	Journal Article	2015	ScienceDirect
PS70	Collective intelligence meets medical decision-making: The collective outperforms the best radiologist	Max Wolf et al.	Journal Article	2015	Public Library of Science (Google Scholar)
PS71	Integration computing and collective intelligence	Marcin Maleszka and Ngoc Thanh Nguyen	Journal Article	2015	ScienceDirect
PS72	Collective intelligence and group performance	Anita Williams Woolley et al.	Journal Article	2015	SAGE (Google Scholar)
PS73	The Role of Collective Intelligence in Crowdsourcing Innovation	Juho Salminen	Thesis	2015	Doria: Lappeenranta University of Technology (IS) (Google Scholar)
PS74	Emergent Collective Decision-Making: Control, Model and Behavior	T. Shen	Thesis	2015	Princeton University Press (Google Scholar)
PS75	Collective Intelligence in Patient Organisations	Lydia Nicholas and Stefana Broadbent	Technical Report	2015	Nesta (Google Scholar)
PS76	Collective Intelligence in Organisations: Uses and Challenges	Silverman	Technical Report	2015	Silverman (Google Scholar)
PS77	A Design Science Approach to Collective Intelligence Systems	A. Kornrumpf and U. Baumol	Conference Paper	2014	IEEE Xplore
PS78	D-CENT. Decentralised Citizens ENgagement Technologies. D2. 1 Collective Intelligence Framework	Javier Toret and Antonio Calleja	Technical Report	2014	D-CENT (Google Scholar)
PS79	How Collective Intelligence Emerges: Knowledge Creation Process in Wikipedia from Microscopic Viewpoint	Kyungho Lee	Conference Paper	2014	ACM Digital Library

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS80	Analysis of collaborative design networks: A case study of OpenIDEO	Mark Fuge et al.	Journal Article	2014	ASME (Google Scholar)
PS81	Power through things: Following traces of collective intelligence in internet of things	Monika Maciulien	Journal Article	2014	MRUNI (Google Scholar)
PS82	Smart Asia: A New Platform for Collective Intelligence	Christopher Grant Kirwan	Conference Paper	2014	SpringerLink
PS83	Spatial collective intelligence? Credibility, accuracy, and volunteered geographic information	Seth E. Spielman	Journal Article	2014	Taylor & Francis (Google Scholar)
PS84	Toward an Ethical Framework for Web-Based Collective Intelligence	Khaled Saleh Al Omoush	Conference Paper	2014	SpringerLink
PS85	Using the Internet as a Collective Intelligence Platform in Harnessing Issues on Climate Change	S. S. Gunasekaran et al.	Conference Paper	2014	IEEE Xplore
PS86	The surprising power of collective intelligence	Michael Brooks	Journal Article	2014	ScienceDirect
PS87	Social Collective Intelligence	D. Miorandi	Book	2014	SpringerLink
PS88	A systematic review of software robustness	Ali Shahrokni and Robert Feldt	Journal Article	2013	ScienceDirect
PS89	Anatomy of a Collective Intelligence Blood Supply Chain	Frederic Andres et al.	Conference Paper	2013	ACM Digital Library
PS90	Challenges and Perspectives of Innovative Digital Ecosystems Designed to Monitor and Warn Natural Disasters in Brazil	Luciana S. Soler et al.	Conference Paper	2013	ACM Digital Library
PS91	Collective Intelligence Based Place Recommendation System	Jehwan Oh et al.	Conference Paper	2013	SpringerLink
PS92	Collective Intelligence for Suicide Surveillance in Web Forums	Tim M. H. Li et al.	Conference Paper	2013	SpringerLink
PS93	Delivering Patients to Sacre Coeur: Collective Intelligence in Digital Volunteer Communities	Kate Starbird	Conference Paper	2013	ACM Digital Library
PS94	Collective Intelligence in Toursplan: An Online Tourism Social Network with Planning and Recommendation Services	Nuno Luz et al.	Conference Paper	2013	ACM Digital Library
PS95	Collective Transport of Complex Objects by Simple Robots: Theory and Experiments	Michael Rubenstein et al.	Conference Paper	2013	International Foundation for Autonomous Agents and Multiagent Systems (Google Scholar)

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS96	Designing a Collective-Intelligence System for Evaluating Complex, Crowd-Generated Intellectual Artifacts	Yiftach Nagar	Conference Paper	2013	ACM Digital Library
PS97	Elements of Software Ecosystem Early-Stage Design for Collective Intelligence Systems	Juergen Musil et al.	Conference Paper	2013	ACM Digital Library
PS98	Evolution of the software-as-a-service innovation system through collective intelligence	Kibae Kim and Jorn Altman	Journal Article	2013	World Scientific (Google Scholar)
PS99	From collective intelligence to collective intelligence systems: Definitions and a semi-structured model	Alexander Kornrumpf and Ulrike Baumol	Journal Article	2013	World Scientific (Google Scholar)
PS100	Mining Collective Intelligence in Diverse Groups	Guo-Jun Qi et al.	Conference Paper	2013	ACM Digital Library
PS101	Social media and collective intelligence-ongoing and future research streams	Detlef Schoder et al.	Journal Article	2013	SpringerLink
PS102	The emergence of collective intelligence	S. S. Gunasekaran et al.	Conference Paper	2013	IEEE Xplore
PS103	Topology of Social Networks and Efficiency of Collective Intelligence Methods	Alan Godoy and Fernando J. Von Zuben	Conference Paper	2013	ACM Digital Library
PS104	Toward Collective Intelligence for Fighting Obesity	I. D. Addo et al.	Conference Paper	2013	IEEE Xplore
PS105	Water: How collective intelligence initiatives can address this challenge	Tony Diggle	Journal Article	2013	Emerald Insight (Google Scholar)
PS106	Collective intelligence as mechanism of medical diagnosis: The iPixel approach	Yuliana PeRez-Gallardo et al.	Journal Article	2013	ScienceDirect
PS107	Value co-creation through collective intelligence in the public sector: A review of US and European initiatives	Sean Wise et al.	Journal Article	2012	Emerald Insight (Google Scholar)
PS108	Collective Intelligence Model: How to Describe Collective Intelligence	Sandro Georgi and Reinhard Jung	Conference Paper	2012	SpringerLink
PS109	Emerging Collective Intelligence Business Models	Miguel de Castro Neto and Ana Espirtio Santo	Conference Paper	2012	MCIS (Google Scholar)
PS110	A Collective Intelligence Based System for Visualizing Problems in Public Roads	M. K. Nishi et al.	Conference Paper	2012	IEEE Xplore
PS111	Advances in Collective Intelligence 2012	Jorn Altmann et al.	Book	2011	SpringerLink
PS112	Collective Intelligence and Social Computing: A Literature Review	Luca Cremona and Aurelio Ravarini	Conference Paper	2012	SpringerLink
PS113	Collective Intelligence Applications in IT Services Business	M. Vukovic and O. Stewart	Conference Paper	2012	IEEE Xplore

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS114	Collective Intelligence As a Source for Machine Learning Self-Supervision	Saulo D. S. Pedro et al.	Conference Paper	2012	ACM Digital Library
PS115	Collective intelligence in organizations: Tools and studies	Antonietta Grasso and Gregorio Convertino	Journal Article	2012	SpringerLink
PS116	Collective Intelligence in the Online Social Network of Yahoo!Answers and Its Implications	Ze Li et al.	Conference Paper	2012	ACM Digital Library
PS117	Computational intelligence techniques for communities network formation	Iulia Maries and Emil Scarlat	Journal Article	2012	Emerald Insight (Google Scholar)
PS118	Construction of a Personally Adapted e-Learning System using Collective Intelligence	J. She et al.	Conference Paper	2012	IEEE Xplore
PS119	Crowd Computation: Organizing Information During Mass Disruption Events	Kate Starbird	Conference Paper	2012	ACM Digital Library
PS120	CrowdCamp: Rapidly Iterating Ideas Related to Collective Intelligence and Crowdsourcing	Paul Andre et al.	Conference Paper	2012	ACM Digital Library
PS121	Designing a Facilitator's Cockpit for an Idea Management System	Marcos Baez and Gregorio Convertino	Conference Paper	2012	ACM Digital Library
PS122	Estimating the Number of Lanes on Rapid Road Map Survey System Using GPS Trajectories as Collective Intelligence	N. Sato et al.	Conference Paper	2012	IEEE Xplore
PS123	From Turing Machine Intelligence to Collective Intelligence	L. Huang et al.	Conference Paper	2012	IEEE Xplore
PS124	Harnessing collective intelligence of Web 2.0: Group adoption and use of Internet-based collaboration technologies	Xiao-Liang Shen et al.	Journal Article	2012	Taylor & Francis (Google Scholar)
PS125	Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many	Helene Landemore	Book Section	2012	Princeton University Press (Google Scholar)
PS126	Interaction and Collective Intelligence in Internet Computing	Deyi Li	Conference Paper	2012	ACM Digital Library
PS127	Onto Collective Intelligence in Social Media	Martin Atzmueller	Conference Paper	2012	ACM Digital Library
PS128	Open Geometry Textbook: A Case Study of Knowledge Acquisition via Collective Intelligence	Xiaoyu Chen et al.	Conference Paper	2012	SpringerLink
PS129	Open government and ctizen participation in law enforcement via crowd mapping	V. Furtado et al.	Journal Article	2012	IEEE Xplore

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS130	QLIM—A Tool to Support Collective Intelligence	Y. Veilleroy et al.	Conference Paper	2012	IEEE Xplore
PS131	KikuNavi: Real-Time Pedestrian Navigation Based on Social Networking Service and Collective Intelligence	H. Nagasaka et al.	Conference Paper	2012	IEEE Xplore
PS132	The Human-Robot Cloud: Situated Collective Intelligence on Demand	N. Mavridis et al.	Conference Paper	2012	IEEE Xplore
PS133	The Mustawa Framework: When Collective Intelligence Meets Intercultural Communication	Salwa Al-Mannai and Andreas Karatsolis	Conference Paper	2012	ACM Digital Library
PS134	What Do You Think?: The Structuring of an Online Community As a Collective-Sensemaking Process	Yiftach Nagar	Conference Paper	2012	ACM Digital Library
PS135	Collective intelligence with web-based information aggregation markets: The role of market facilitation in idea management	Efthimios Bothos et al.	Journal Article	2012	ScienceDirect
PS136	Crowd-Powered Systems	Michael Scott Bernstein	Thesis	2012	Massachusetts Institute of Technology (Google Scholar)
PS137	Collective Intelligence Capitalizing on the crowd	Eric Lesser	Technical Report	2012	IBM (Google Scholar)
PS138	Collective intelligence systems: Classification and modeling	Ioanna Lykourentzou et al.	Journal Article	2011	Journal of Emerging Technologies in Web Intelligence (Google Scholar)
PS139	A Collective Intelligence Based Business-Matching and Recommending System for Next Generation e-Marketplaces	Y. Iguider and H. Morita	Conference Paper	2011	IEEE Xplore
PS140	A Collective Intelligence Resource Management Dynamic Approach for Disaster Management: A Density Survey of Disasters Occurrence	E. Asimakopoulou et al.	Conference Paper	2011	IEEE Xplore
PS141	Collaborative Cyberporn Filtering with Collective Intelligence	Lung-Hao Lee and Hsin-Hsi Chen	Conference Paper	2011	ACM Digital Library
PS142	Collective Intelligence and Online Learning Communities	M. Gea et al.	Conference Paper	2011	IEEE Xplore
PS143	Collective Intelligence Applications—Algorithms and Visualization	A. Nagalakshmi and S. Joglekar	Conference Paper	2011	IEEE Xplore
PS144	Design for Public Service Application Based on Collective Intelligence in China	Zhiyong Fu	Conference Paper	2011	ACM Digital Library
PS145	Human Computation: A Survey and Taxonomy of a Growing Field	Alexander J. Quinn and Benjamin B. Bederson	Conference Paper	2011	ACM Digital Library

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS146	Modeling and simulating crowdsourcing as a complex biological system: Human crowds manifesting collective intelligence on the internet	Rolf Pfeifer et al.	Journal Article	2011	MIT Press (Google Scholar)
PS147	Research on the Application of Recommendation Engine Based on Collective Intelligence in Intelligent Web	Xiaolin Xu and Guanglin Xu	Conference Paper	2011	IEEE Xplore
PS148	Scaling Up: From Individual Design to Collaborative Design to Collective Design	Mary Lou Maher et al.	Conference Paper	2011	SpringerLink
PS149	Survey on Governance of User-Generated Content in Web Communities	Felix Schwagereit et al.	Conference Paper	2011	ACM Digital Library
PS150	The Climate CoLab: Large Scale Model-Based Collaborative Pplanning	J. Introne et al.	Conference Paper	2011	IEEE Xplore
PS151	The Tree of Knowledge: A Localized Collective Intelligence Tool	Gyu Hyun Kwon et al.	Conference Paper	2011	ACM Digital Library
PS152	Hospital-based nurses' perceptions of the adoption of Web 2.0 tools for knowledge sharing, learning, social interaction and the production of collective intelligence	Adela S. M. Lau	Journal Article	2011	JMIR (Google Scholar)
PS153	The Global Brain Institute Vision: Past, Present and Future Context of Global Brain Research	Francis Heylighen	Working paper	2011	The Global Brain Institute (Google Scholar)
PS154	Designing for collective intelligence	Dawn G. Gregg	Journal Article	2010	ACM Digital Library
PS155	On model design for simulation of collective intelligence	Martijn C. Schut	Journal Article	2010	ScienceDirect
PS156	A Resource Allocation Framework for Collective Intelligence System Engineering	Dimitrios J. Vergados et al.	Conference Paper	2010	ACM Digital Library
PS157	A Novel Recommendation System with Collective Intelligence	J. Zhou et al.	Conference Paper	2010	IEEE Xplore
PS158	Collective Decision-Making in Multi-Agent Systems by Implicit Leadership	Chih-Han Yu et al.	Conference Paper	2010	ACM Digital Library
PS159	Collective Intelligence for the Design of Emergency Response	A. S. Vivacqua and M. R. S. Borges	Conference Paper	2010	IEEE Xplore
PS160	Collective intelligence in law enforcement—The WikiCrimes system	Vasco Furtado et al.	Journal Article	2010	ScienceDirect
PS161	Collective Intelligence in Mobile Consumer Social Applications	S. Diplaris et al.	Conference Paper	2010	IEEE Xplore

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS162	IDE 2.0: Collective Intelligence in Software Development	Marcel Bruch et al.	Conference Paper	2010	ACM Digital Library
PS163	Collective Intelligence in the Social Web	John Riedl	Conference Paper	2010	SpringerLink
PS164	Crowdsourcing, Open Innovation and Collective Intelligence in the Scientific Method: A Research Agenda and Operational Framework	T. Bucheler et al.	Conference Paper	2010	MIT Press (Google Scholar)
PS165	Emergence of Collective Intelligence in Distance Education System	Hui Li and Zhen Li	Conference Paper	2010	IEEE Xplore
PS166	From Collective Knowledge to Intelligence: Pre-Requirements Analysis of Large and Complex Systems	Peng Liang et al.	Conference Paper	2010	ACM Digital Library
PS167	Personalized Reading Support for Second-Language Web Documents by Collective Intelligence	Yo Ehara et al.	Conference Paper	2010	ACM Digital Library
PS168	Towards a framework for Collective Intelligence	J. Berri	Conference Paper	2010	IEEE Xplore
PS169	Workings of Collective Intelligence within Open Source Communities	Everett Stiles and Xiaohui Cui	Conference Paper	2010	SpringerLink
PS170	Using collective intelligence to fine-tune public health policy	Marsh Andy et al.	Journal Article	2010	IOS Press (Google Scholar)
PS171	Collective intelligence	Jan Marco Leimeister	Journal Article	2010	SpringerLink
PS172	The collective intelligence genome	T. W. Malone et al.	Journal Article	2010	IEEE Xplore
PS173	Harnessing crowds: Mapping the genome of collective intelligence	Thomas W. Malone et al.	Journal Article	2009	SSRN (Google Scholar)
PS174	Leveraging the power of collective intelligence through IT-enabled global collaboration	Luca Iandoli	Journal Article	2009	Taylor and Francis (Google Scholar)
PS175	Applying Collective Intelligence for Search Improvement on Thai Herbal Information	V. Lertnattee et al.	Conference Paper	2009	IEEE Xplore
PS176	Building Term Suggestion Relational Graphs from Collective Intelligence	Jyh-Ren Shieh et al.	Conference Paper	2009	ACM Digital Library
PS177	The Climate Collaboratorium: Project overview	Thomas Malone et al.	Journal Article	2009	Massachusetts Institute of Technology (Google Scholar)
PS178	Collaborative Platform for Multicultural Herbal Information Creation	Verayuth Lertnattee et al.	Conference Paper	2009	ACM Digital Library
PS179	Collective Intelligence Approach for Formulating a BOK of Social Informatics, an Interdisciplinary Field of Study	Yoshifumi Masunaga et al.	Conference Paper	2009	ACM Digital Library

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS180	Collective Intelligence System Engineering	Ioanna Lykourentzou et al.	Conference Paper	2009	ACM Digital Library
PS181	Developing a collective intelligence application for special education	Dawn Gregg	Journal Article	2009	ScienceDirect
PS182	Decisions 2.0: The power of collective intelligence	Eric Bonabeau	Journal Article	2009	MIT Press (Google Scholar)
PS183	Generic Framework for Recommendation System Using Collective Intelligence	A. Patel and A. Balakrishnan	Conference Paper	2009	IEEE Xplore
PS184	Recommendation system using location-based ontology on wireless internet: An example of collective intelligence by using 'mashup' applications	YoungHoon Yu et al.	Journal Article	2009	ScienceDirect
PS185	Collective Intelligence: Toward Classifying Systems of Systems	Alan J. Ramsbotham Jr.	Conference Paper	2009	ACM Digital Library
PS186	Toward collective intelligence of online communities: A primitive conceptual model	Shuangling Luo et al.	Journal Article	2009	SpringerLink
PS187	User Requirements for a Collective Intelligence Emergency Response System	Vita Lanfranchi and Neil Ireson	Conference Paper	2009	ACM Digital Library
PS188	Enabling in-line deliberation and collective decision-making through large-scale argumentation: A new approach to the design of an internet-based mass collaboration platform	Luca Iandoli et al.	Journal Article	2009	IGI Global (Google Scholar)
PS189	Designing Human-Centered Collective Intelligence	Ivor Addo	Thesis	2009	Marquette University (IS) (Google Scholar)
PS190	A Survey of Web-Based Collective Decision Making Systems	Jennifer H. Watkins and Marko A. Rodriguez	Book Section	2008	SpringerLink
PS191	Collective Intelligence and Bush Fire Spotting	David Howden and Tim Hendtlass	Conference Paper	2008	ACM Digital Library
PS192	Collective Intelligence for Demand-Responsive Transportation Systems: A Self Organization Model	Besma Zeddini	Conference Paper	2008	ACM Digital Library
PS193	Collective Intelligence in Knowledge Management	Wenyan Yuan et al.	Conference Paper	2008	SpringerLink
PS194	Crowdsourcing as a Model for Problem Solving: An Introduction and Cases	Daren C. Brabham	Journal Article	2008	SAGE (Google Scholar)
PS195	Collective Intelligence in Action	Satnam Alag	Book	2008	Manning Publications (Google Scholar)

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS196	Quo Vadis computer science: From Turing to personal computer, personal content and collective intelligence	Epaminondas Kapetanios	Journal Article	2008	ScienceDirect
PS197	Service-Oriented Collective Intelligence for Intercultural Collaboration	T. Ishida	Conference Paper	2008	IEEE Xplore
PS198	Web 2.0 and the Collective Intelligence	Kari A. Hintikka	Conference Paper	2008	ACM Digital Library
PS199	Collective intelligence approaches to malware recognition	Iñaki Urzay	Journal Article	2008	ScienceDirect
PS200	Can we exploit collective intelligence for collaborative deliberation? The case of the climate change collaboratorium	Luca Iandoli et al.	Journal Article	2008	SSRN (Google Scholar)
PS201	What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software	Tim O'Reilly	MRPA Paper	2007	University Library of Munich (Google Scholar)
PS202	Collective Intelligence for Decision Support in Very Large Stakeholder Networks: The Future US Energy System	S. Rasmussen et al.	Conference Paper	2007	IEEE Xplore
PS203	Programming Collective Intelligence	Toby Segaran	Book	2007	O'Reilly (Google Scholar)
PS204	Collective intelligence: A keystone in knowledge management	Andre Boder	Journal Article	2006	Emerald Insight (Google Scholar)
PS205	Collective intelligence	A. Pentland	Journal Article	2006	IEEE Xplore
PS206	Collective intelligence: It's all in the numbers	K. S. Jones	Journal Article	2006	IEEE Xplore
PS207	The power of collective intelligence	Aaron Weiss	Journal Article	2005	ACM Digital Library
PS208	Collective Intelligence as a Framework for Supply Chain Management	L. Sheremetov and L. Rocha-Mier	Conference Paper	2004	IEEE Xplore
PS209	Theory of Collective Intelligence	David Wolpert	Book Section	2004	SpringerLink
PS210	The SMSMS Project: Collective intelligence machines in the digital city	Maurizio Bolognini	Journal Article	2004	MIT Press (Google Scholar)
PS211	Collective intelligence of the artificial life community on its own successes, failures, and future	Steen Rasmussen et al.	Journal Article	2003	MIT Press (Google Scholar)
PS212	Universal Formal Model of Collective Intelligence and Its IQ Measure	Tadeusz Szuba	Conference Paper	2002	SpringerLink

(Continued)

Table 2. Continued

ID	Study Title	Author (s)	Publication Type	Year	Publisher (Source)
PS213	Robustness in Complex Systems	Steven D. Gribble	Conference Paper	2001	IEEE Xplore
PS214	A formal definition of the phenomenon of collective intelligence and its IQ measure	Tadeusz Szuba	Journal Article	2001	ScienceDirect
PS215	Collective intelligence as a field of multi-disciplinary study and practice	Tom Atlee and George Por	Journal Article	2000	En ligne (Google Scholar)
PS216	Toward a Computational Model of Collective Intelligence and Its IQ Measure	Mohammed Almulla and Tadeusz Szuba	Conference Paper	1999	ACM Digital Library (Google Scholar)
PS217	Using Collective Intelligence to Route Internet Traffic	D. H. Wolpert et al.	Journal Article	1999	NIPS (Google Scholar)
PS218	Collective Intelligence: Mankind's Emerging World in Cyberspace	Pierre Lévy	Book	1997	Perseus Books (Google Scholar)
PS219	Fundamental concepts of collective intelligence	William Sulis	Journal Article	1997	SpringerLink (Google Scholar)

Appendix 4

[IV]

S. Suran, V. Pattanaik, and D. Draheim. CommunityCare: Tackling mental health issues with the help of community. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2020 – the 22nd International Conference on Information Integration and Web-based Applications and Services*, pages 377–382. ACM, 2020

CommunityCare: Tackling Mental Health Issues With The Help Of Community

Shweta Suran
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
shweta@taltech.ee

Vishwajeet Pattanaik
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
vishwajeet.pattanaik@taltech.ee

Dirk Draheim
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
dirk.draheim@taltech.ee

ABSTRACT

Building on the rising interest in online crowdsourcing platforms, and the ever-rising concerns over mental health issues worldwide; in this paper, we propose a novel citizen-oriented web-based Collective Intelligence (CI) platform called 'CommunityCare'. The platform is meant to be focused on end-users' communities and aims to empower its users by allowing them to work collectively when helping those suffering from mental health issues such as depression, anxiety, and stress. Through our work, we attempt to make two distinct contributions: first, we elucidate an abstract yet novel CI platform for mental health, that could enable citizen volunteers and medical practitioners to work together to help those suffering from psychological/behavioural issues; and second, we evaluate our previously proposed 'generic' CI framework by utilizing the said platform as a use case for the same. We describe the 'CommunityCare' platform through the four primary components of CI systems, namely: staffing, processes, goals and motivation.

CCS CONCEPTS

• **Information systems** → **Crowdsourcing**; • **Applied computing** → **Health care information systems**; • **Software and its engineering** → *Development frameworks and environments*; *Use cases*; Abstraction, modeling and modularity; Designing software.

KEYWORDS

collective intelligence, crowdsourcing, e-health, framework, information system, mental health, use case

ACM Reference Format:
Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2020. CommunityCare: Tackling Mental Health Issues With The Help Of Community. In *The 22nd International Conference on Information Integration and Web-based Applications & Services (iiWAS '20), November 30-December 2, 2020, Chiang Mai, Thailand*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3428757.3429114

iiWAS '20, November 30-December 2, 2020, Chiang Mai, Thailand

ACM ISBN 978-1-4503-8922-8/20/11…$15.00
https://doi.org/10.1145/3428757.3429114

1 INTRODUCTION

The rise in mental health issues among citizens, especially over the last few decades has been a critical challenge worldwide [5, 9]. Such issues continue to grow as a burden on our societies, and have significant impacts on not only the social aspects of an individual's life; but also have socioeconomic consequences world wide [1]. Some reports even go on to suggest that 1 in 4 people worldwide are expected to suffer from some kind of mental illness at some point in their lives [1]. Beyond the psychological impact of mental health issues, several studies suggest that these issues can also lead to other medical conditions such as cardiovascular disorders [21], strokes, diabetes [19], spinal problems [23] and even complete mental collapse. In worst cases, it can also force a person to do self-harm, or even commit suicide [13]. The situation is so daunting that reports claim that globally approximately 264 million people are affected by depression (and other similar mental health ailments).

Now, given the more recent COVID-19 outbreak and the following global lockdown, the situation seems to have become direr. So far so, that studies indicate that the outbreak is "leading to additional health problems such as stress, anxiety, depressive symptoms, insomnia, denial, anger and fear globally" [26]. According to the Organisation for Economic Cooperation and Development's (OECD) report, in Europe alone 10 million people suffer from mental disorders every year [16]. Also, mental health problems such as depression and stress, excessive drinking, suicides and social exclusion are gradually becoming more severe, in quite a few countries in the EU [15].

That said, typically as citizens, we would rely on the medical infrastructures of our nations to help us tackle these challenges. However, there is another less explored avenue, which we would like to draw attention towards in this paper; that is, the World Wide Web, and more specifically, Crowdsourcing / Collective Intelligence (CI) platforms on the Web. Although, in essence, the concept of CI viz. 'wisdom of the crowd' has been around for centuries; CI in Information and Communications Technology (ICT) has only been around for the previous two decades [6, 11, 12]. And, the field has only gained interest among researchers more recently, thanks to the rising influence of the Social Web [24].

CI platforms, in particular, have enabled mass collaborations [20], where a large number of individuals can come together to exchange new information, solve problems and generate new innovative solutions that benefit both the individuals as well as the society as a whole. There are in-fact various types of CI platforms available on

[1] World Health Organization's Report on Mental Disorders (2019)

the web, however, the underlying objective of said platforms is always the same i.e., to connect individuals and to harness their collective intelligence for various purposes. On one side of the spectrum, CI platforms like (i) ClimateCoLab utilize the collective knowledge of individuals to solve complex effects of climate change [7], (ii) Tippanee uses collaborative knowledge sharing to improve the quality of textual content on the web [18] and (iii) Wikipedia exploits the wisdom of the crowd to build an online collaborative encyclopedia [12]; while on the other side of the spectrum organizations like InnoCentive, OpenIDEO, NESTA [24] and AccelerateEstonia (aE!) harness the collective innovation and knowledge of the crowd to tackle wicked societal issues with the help of experts. A more recent example of this is, InnoCentive's initiative to tackle the COVID-19 pandemic. Under the initiative citizens (as solvers and independent group thinkers) were asked to submit creative ideas and propose novel solutions/services to prevent the transmission of COVID-19, and to present new designs for ventilator systems, anti-COVID-19 infection protective films/masks, and more. Other similar examples of utilisation of CI can be seen in Nonaka's SECI (Socialization, Externalization, Combination and Internalization) model [14] that is (more recently) being utilized by organizations [10] and communities [3, 17] to generate new knowledge; thereby promoting sustainable innovation and enhancing knowledge creation within both organizations and communities.

Drawing from the ever-rising risks of mental health issues worldwide, in this paper we set out to detail a community-oriented, mental health online CI platform called 'CommunityCare'. The platform aims to provide citizens with a confidential, community-based online platform where citizens suffering from mental health issues could be offered emotional and mental health support by not only medical practitioners but also by citizen volunteers. The platform aims to bring together citizens (with mental health issues), volunteers, medical practitioners, law enforcement agencies and policymakers on a common platform; while providing easy-to-access, secure (preventing social stigma) cognitive-behavioural therapy (CBT) psychological support to those in need. The proposed platform aims to do so by,

(1) allowing citizens to share their feelings with other citizens anonymously,
(2) by allowing citizen volunteers to help their fellow-citizen by providing emotional care,
(3) by enabling medical organizations to provide training to volunteers and extensive support to patients based on severity,
(4) by assisting law enforcement agencies in gathering information about citizens at higher risk of doing self-harm,
(5) by empowering governing bodies in protecting the rights of citizens and supporting other related activities.

Apart from the proposed novel CI platform, we make an additional contribution in this paper i.e., we build on our previous work in CI [24, 25] and examine the proposed CommunityCare platform with the help of our 'generic' CI framework [24]. To do so, we describe the various components of the proposed platform through the building blocks of the 'generic' CI model and therefore evaluate the framework in the process. This is vital, as a 'generic' framework for CI should inherently be able to describe any CI platform, irrespective of the platforms' domain or application [24]. Finally, we provide additional insights into user behaviour and limitations of a CI-based mental health platform.

2 RELATED WORK

As mentioned in the previous section, web-based CI platforms have gained tremendous interest in the past few years; so much so, that quite a few online platforms oriented towards mental health issues have been designed and deployed around the world. In this section, we briefly discuss some such platforms. We enumerate the list of features these platforms offer and establish the key drawbacks of each of these systems. This is vital, as we would like to incorporate said features into the proposed CommunityCare platform, thereby improving its overall usefulness and novelty.

Kooth [2] is a free and completely confidential online service based in the UK. The platform offers emotional and mental health support to children and young people; its key features include: counselling services, online self-help materials, news feeds, live moderated forums, asynchronous messaging, online chats and emotion status updates. Kooth also offers mood-tracking and goal setting tools to assist the patients to deal with exam stress, depression, self-harm, relationship problems and any other kind of mental stresses. Unfortunately, the platform has two major drawbacks; first, it only offers its services to young people aged between 11 and 25 years; and second, it does not provide 24/7 online support.

iFightDepression [3] is a free, web-based, multilingual, self-management tool meant for young people and adults who are suffering from mild depression [8]. The platform offers its services in 9 different languages including English, German, Spanish, Catalonian, Dutch, Hungarian, Estonian, Italian, and Bulgarian [2]. iFightDepression is intended to promote self-monitoring, cognitive restructuring and sleep regulation for individuals who are suffering from moderate depression. A key unique feature of the platform is that it provides self-test checking for symptoms of depression. Similar to Kooth, iFightDepression also has its own set of limitations. In particular, the platform is not meant for kids and elderly citizens, and can only be used after the referral/recommendation of a general physician (GP) or healthcare professional/counsellors.

The InnoWell platform [4] is a digital clinical tool for treatment and self-management for young people in Australia. Shared decision making, early intervention, treatment for children with mental health issues, and person-centred care are some key features of the platform. The platform achieves its goal by gathering and reporting mental health information from and to patients and their carers/counsellors; and hence, promotes patient satisfaction and collaborative care. One limitation of the platform is that it does not provide any advice on the medical conditions of an individual and like aforementioned platforms does not extend its services to elderly citizens.

MoodPath [5] (recently renamed to MindDoc) is a mental health tracking app for detection and treatment of depression. The key features of the app are that it tracks and monitors depressive symptoms, and conveys the depressive, anxiety and stress symptoms

[2]Kooth homepage: https://www.kooth.com/
[3]iFightDepression homepage: https://ifightdepression.com/en/
[4]InnoWell homepage: https://www.innowell.org/
[5]MoodPath homepage: https://mymoodpath.com/en/

of the patients to their care providers. For tracking and monitoring within the app, patients need to follow a two-week interactive program and their responses are utilized to provide users with personalized reports. Additionally, the app provides patients with a synopsis of their emotional states for each day, which helps patients track their mood changes throughout the day. Unfortunately, the app can only be used by patients who are above the age of 12; and does not provide any discussion forums, counselling or advice from healthcare professionals/counsellors.

Finally, Talkspace [6] is an online platform that offers psychotherapy services from licensed therapists. The platform allows its users to exchange unlimited messages with their therapists; users can also get couples therapies and even take part in anxiety and depression tests. Based on the provided information, patients are then linked with respective psychotherapists, who then provide personalized treatments including medications and follow-ups. Unfortunately, the Talkspace services are not available 24/7 and to access online therapy patients are required to pay for different plans/subscriptions; for instance, the unlimited messaging therapy costs $65/week and the couples therapy plan costs $99/week.

Although the above-mentioned platforms offer a wide variety of features, each of them has its own drawbacks and pitfalls. A key drawback of all of these platforms is that they are owned by private organizations, and therefore said organizations have total control over the users' personal data. Also, most of the platforms have restricted access based age, gender, region and medical conditions. Additionally, some of these platforms only offer their services at specific periods, are often too expensive for the general public, and are only focused towards providing tailored/personalized services; rather than providing general help and guidance.

Apart for the rather minor drawbacks mentioned above, there is one other key pitfall that is an inevitable part of most mental health related services/web-based platforms. And that is, that even with the numerous publicly and privately owned e-health services, platforms and apps; the number of citizens suffering from such disorders are ever-rising. A common pattern that can be observed in the previously mentioned OCED (and other similar) reports is that the human resources in the medical field are unable to keep up these rising number of citizens in need. And this is also true for the platforms we just discussed. An unexplored avenue, that could have potential in assisting us solve this deteriorating situation is the use of 'the crowd'. We are of the opinion, that given the rising societal and scientific interest in CI in the past few years, CI could indeed help us solve this wicked issue while also promoting sustainability.

3 COMMUNITYCARE: A USE CASE FOR THE 'GENERIC' CI FRAMEWORK

Building on the above-mentioned gaps and motivations, in this section we describe the proposed CommunityCare platform. The platform is detailed using the vocabulary of the 'generic' CI framework [24], which describes CI platforms through twenty-four unique attributes. The attributes are segregated into three parts: *the CI model*, *additional requisites*, and *properties that make CI systems complex-adaptive*. The CI model is further divided into four primary components: *staffing*, *processes*, *goals*, and *motivation*; and these are again segregated into *types*, *properties*, and *interactions* [24].

3.1 Staffing

As proposed by Suran et al. [24], the staffing i.e., users of a CI platform can be categorized as *beneficiaries* and *contributors*. Beneficiaries are platform users who aim to gain from the knowledge or artefacts produced by the platform, but do not contribute to the co-creation process; whereas contributors are users who actively participate in create-decide activities. The contributors on the CI platform can be further classified as *crowd* and *hierarchy*. While the members of the crowd actively contribute by creating solutions and artefacts on the platform, the members of the hierarchy are primarily responsible for decision making and work allocation.

Within the CommunityCare platform, the patients (i.e., citizens requiring mental health assistance), their family members and the state (i.e., governing body) are classified as beneficiaries. Both the citizens and the state, would not participate in the co-creation/decide activities, however, they would still seek to gain from the platform's innovations. Since the platform is oriented towards citizens suffering from mental health issues and their families, such citizens would be the primary beneficiaries of the system. These beneficiaries would be citizens who would like to receive psychological support and therapies from trained volunteers and medical practitioners. Whereas, the family members/carers of said citizens would aim to learn self-help techniques to assist them in their day-to-day lives. Finally, the state would aim to maintain a stable and secure mental health platform for societal welfare. For the state, the platform would be a win-win situation anyhow; as the platform would support the state by not just allowing citizens to live a better lifestyle, but also help the state generate revenue and may-be even provide a means for sustainable income for the state's future endeavours.

Citizen volunteers within the platform would assist psychiatrists and medical practitioners by talking or chatting to citizens with initial psychological symptoms. Since the volunteers would actively contribute to the platform, they can be classified as the crowd (contributors). Said volunteers would be primarily responsible for assisting people suffering from early symptoms of mental disorders, and would, therefore, help in reducing the workload of counsellors and medical practitioners. This is key, as in many cases people suffering from mental disorders often show early symptoms, that if tackled instantly can prevent their condition from deteriorating.

Lastly, organizations and institutions that work in the health sector such as Medical Health Centres, Medical Associations, Non-Profits, and Academic Institutes would constitute the hierarchy (contributors) on the platform. These organizations would be responsible for disseminating knowledge, training citizen volunteers, providing psychiatric services to those in need, and creating/deciding new policies and strategies.

Based on the 'generic' CI framework, the crowd on the platform would be required to have following key properties: the user base should be diverse i.e., the users should belong to different economical/sociotechnical backgrounds; the users should be allowed to act as independent entities; and the crowd should have a critical mass

[6] Talkspace homepage: https://www.talkspace.com/

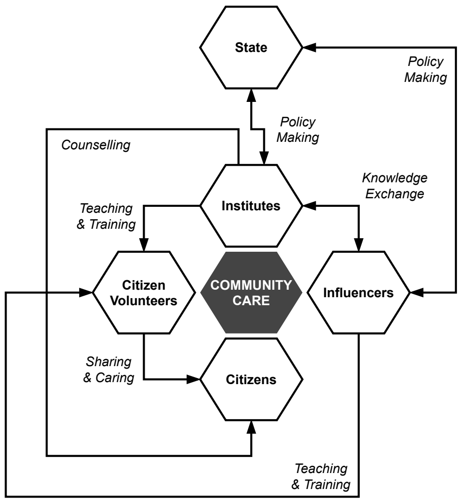


Figure 1: Illustration of relationship between 'staffing' in CommunityCare.

i.e., there should be enough users on the platform who could provide contradictory ideas, thus promoting debates and discussions. Additionally, to promote co-creation said members of the crowd must treat each other with trust and respect; and should socialize, externalize, combine and internalize both old and new ideas and experiences, as doing so could help foster new and innovative ideas.

3.2 Processes

The processes in the 'generic' CI framework [24], are categorized into two unique types and two subsequent interactions. Combining these types and interactions, the framework classifies processes in a CI platform as *collection*, *collaboration*, *individual decisions*, and *group decisions*. Activities where platform users participate in individual creation tasks are referred to as collection, while group creation tasks are referred to as collaboration. When an individual makes a decision on their own, its referred to as an individual decision, and when multiple users decide something after a discussion or debate it is referred to as a group decision.

Based on these categorizations, the processes in the CommunityCare platform can be classified under the said categories. Using the same analogy, the processes (i) by which citizen volunteers and medical practitioner would provide psychological support to those in need, and (ii) individual researchers/medical practitioners coming up with new ideas/policies/treatment methodologies, can be classified as collection. Whereas, group activities such as (i) practitioners providing training to volunteers, and (ii) all platform members creating and adopting new policies and future short-term/long-term goals, can be classified as collaboration.

For decide activities, beneficiaries providing feedback can be considered as individual decisions; while policy-making decisions achieved after discussions and deliberations can be considered group decisions.

3.3 Goals

CI platform goals are categorized into two types: *individual goals*, and *community goals* [24]. Additionally, the goals need to be *well defined*, and *objective*. Given these categorizations, the individual and community goals of the proposed CommunityCare platform can be enumerated as follows. The individual goals of the platform's staffing would be (i) to receive/provide mental health services (psychological support and Cognitive Behavioural Therapy) and (ii) to promote knowledge exchange and training on mental health.

Whereas, the community goals would be (i) to make mental health-related services more accessible to all citizens, by allowing users to have access to the services 24/7 and online; (ii) to provide a safe and stigma-free environment for citizens to get help for their mental health issues, without worrying about issues such as data monitoring (thereby encouraging more suffering citizens to come forward to avail help); (iii) to bring communities closer to each other by promoting empathy among citizens, thereby motivating citizen volunteers to come forward to assist their fellow-citizen; (iv) to aggregate data and knowledge on mental health issues (keeping in mind the ethical implications), and (v) to create innovative solutions and methodologies meant to tackle mental health issues.

3.4 Motivation

The 'generic' CI framework [24], categorizes user motivations into two primary types *intrinsic* and *extrinsic*. User motivations such as *interest*, *passion*, *keenness to learn*, and *empathy towards a social cause* are classified as intrinsic motivators; whereas, *money (tangible)* and *glory (intangible)* are classified as extrinsic motivators.

In the proposed CommunityCare platform, citizens (and their family members) who aim to seeks assistance would be motivated by interest and knowledge; whereas citizen volunteers who would like to assist those in need, would be motivated by passion, knowledge, and social cause. The state and other members of the hierarchical staffing would have both intrinsic and extrinsic motivations.

The intrinsic motivation for these hierarchical staffing to participate in the platform would be twofold, (i) the influx of volunteers would help reduce the stress and workload on medical practitioners, and their participation in the platform might motivate more people to study, learn, and practice in health services (i.e., more students, and more future nurses and doctors), and (ii) the institutions would aim to gain more visibility through the platform's communication channels as this could potentially open up new avenues for research fundings (which institutions are typically always in need of). While their extrinsic motivations would be money for sustainability and glory (i.e., reputation) for future research endeavours. Additionally, the influencers (i.e., associations and non-profits) would be motivated by their own self-interests as they would be responsible for making sure that the rights of the practitioners participating within the platform are taken into consideration when creating new policies.

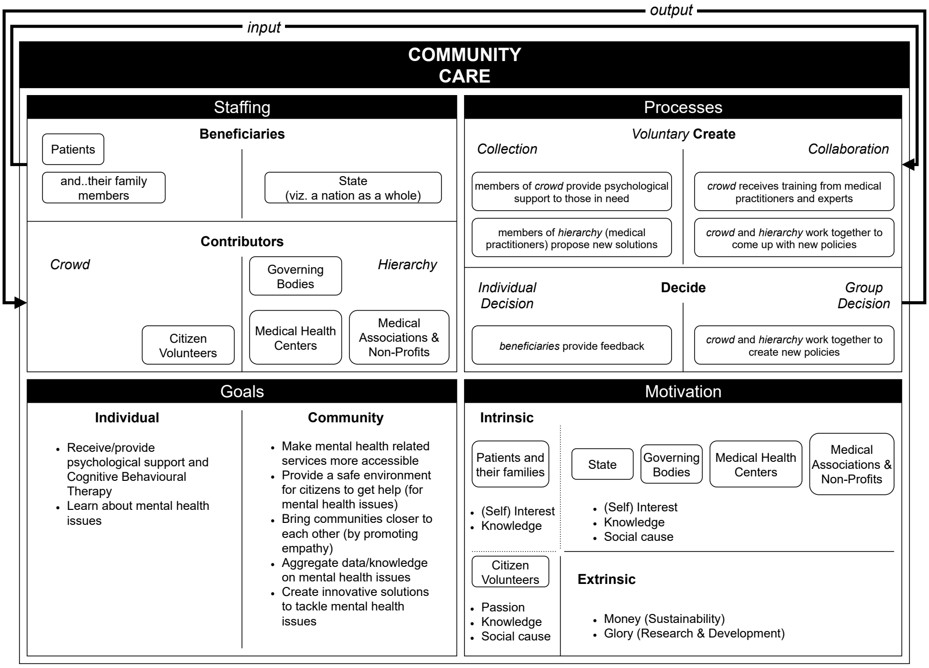


Figure 2: Components of CommunityCare illustrated via. the 'generic' CI model.

3.5 Additional Attributes

Since the effectiveness of the CommunityCare platform depends mainly on user-generated data/information, therefore the platform would be required to provide mechanisms like social tagging [22, 24] and aggregating functions [4] to aggregate the data/information gathered from users. The platform's services would have to be accessible to all citizens and on multiple devices (including PCs and smartphones). The user interface of the platform would have to be easy to use and visually rich, so much so that even a layperson without any technical background could use the platform. Finally, the platform would have to provide its users with access to the aggregated knowledge, based on the users' role; and would require well-ordered task-and-workload allocations that would enable coordination and collaboration activities in an effective manner.

4 DISCUSSION

Apart from the few key differences illustrated by the comparison in Table 1, the CommunityCare platform aims to provide several key features that add to the platform's novelty as an online mental e-health platform. First, the platform's feature sets and architecture revolve around the concept of 'community', and the platform's primary goal is to use the literal 'power of the crowd' to assist those who require psychological support. This is vital as even with the plethora of social apps on the Web, human-centred platforms where users can share/discuss their inner thoughts and feelings (in a safe environment) are actually declining. Building the platform around the idea of a community should prompt users from multiple facets of life to come together, and could therefore inculcate a sense of oneness among the platform users. Second, by allowing general users to participate and contribute as volunteers, the platform aims to instil a sense of empathy among the platform's users. Also, with its unique volunteer training feature the platform could potentially bring about adaptivity and self-organization, as trained volunteers could go on to study human psychology and mental health, and later come back to the platform as medical practitioners. The platform's adaptivity could be further escalated by its hierarchical contributors, as novel solutions proposed by experts (from outside the platform) could be easily adopted and utilized within the system (based on community feedback and suggestions). This adaptive and self-organizing behaviour of the platform could potentially lead to the emergence in the platform's future versions.

5 CONCLUSION

Tackling mental health issues has been a global issue for quite a few decades now. Given the rising interest in web services and Collective Intelligence platforms, in this paper, we set out to devise an online, crowd-focused, Collective Intelligence platform that aims to overcome the pitfalls of the state-of-the-art in mental e-health systems. As an added contribution we evaluate our previously proposed 'generic' CI framework, by describing the proposed platform

Table 1: Comparison of CommunityCare vs. other mental health platforms (discussed in Section 2)

Platform	User Age Group	24/7 Service	Free/ Paid	Access Type	Platform Ownership	MA/C[f]	Citizen Volunteers
Kooth	11-24 years		Free	Open for all	Private	✓	
iFightDepression	15-24 years	✓	Free	Only after recommendation from a general physician or healthcare professional	Private	✓	
InnoWell	16-25 years	✓	Free	Invitation required from healthcare professional	Private		
MoodPath	all-ages	✓	Free	Open for all	Private		
Talkspace	13+ years		Paid	Open for all**	Private	✓	
CommunityCare	all-ages	✓	Free	Open for all**	Public	✓	✓

[f] Medical Advice / Counselling

**To access services, children are required to get parental consent.

using the vocabulary of the same. We delve into the four primary components of the CI platform (namely *staffing*, *process*, *goal* and *motivation*), and then explain the additional requisites such a e-health system should fulfil in order to be accepted as a CI platform. Finally, we provide insights into how, the proposed platform could potentially become adaptive and self-organizing over time, and even lead to emergence given enough time.

REFERENCES

[1] Jordi Alonso. 2012. Burden of Mental Disorders based on the World Mental Health Surveys. *Revista Brasileira de Psiquiatria* 34, 1 (Mar 2012), 7–11. https://doi.org/10.1590/s1516-44462012000100003

[2] Ella Arensman, Nicole Koburger, Celine Larkin, Gillian Karwig, Claire Coffey, Margaret Maxwell, Fiona Harris, Christine Rummel-Kluge, Chantal vanAudenhove, Merike Sisask, Anna Alexandrova-Karamanova, Victor Perez, György Purebl, Annabel Cebria, Diego Palao, Susana Costa, Lauraliisa Mark, Mónika Ditta Tóth, Marieta Gecheva, Angela Ibelshäuser, Ricardo Gusmão, and Ulrich Hegerl. 2015. Depression Awareness and Self-Management Through the Internet: Protocol for an Internationally Standardized Approach. *JMIR Res Protoc* 4, 3 (06 Aug 2015), 99. https://doi.org/10.2196/resprot.4358

[3] Andre Boder. 2006. Collective intelligence: a keystone in knowledge management. *Journal of Knowledge Management* 10, 1 (2006), 81–93. https://doi.org/10.1108/13673270610650120

[4] Simin Cai, Barbara Gallina, Dag Nyström, and Cristina Seceleanu. 2019. Data aggregation processes: a survey, a taxonomy, and design guidelines. *Computing* 101, 10 (2019), 1397–1429. https://doi.org/10.1007/s00607-018-0679-5

[5] Abraham Francis. 2014. *Social Work in Mental Health: Contexts and Theories for Practice.* SAGE Publications, New Delhi. https://doi.org/10.4135/9789351507864

[6] Jerome C. Glenn. 2013. Collective Intelligence and an Application by The Millennium Project. *World Futures Review* 5, 3 (2013), 235–243. https://doi.org/10.1177/1946756713497331

[7] Joshua Introne, Robert Laubacher, Gary Olson, and Thomas Malone. 2011. The Climate CoLab: Large scale model-based collaborative planning. In *2011 International Conference on Collaboration Technologies and Systems (CTS).* IEEE, 40–47. https://doi.org/10.1109/cts.2011.5928663

[8] Azucena Justicia, Matilde Elices, Ana Isabel Cebria, Diego J Palao, Jesús Gorosabel, Dolors Puigdemont, Javier de Diego-Adeliño, Andrea Gabilondo, Alvaro Iruin, Ulrich Hegerl, et al. 2017. Rationale and methods of the iFightDepression study: A double-blind, randomized controlled trial evaluating the efficacy of an internet-based self-management tool for moderate to mild depression. *BMC psychiatry* 17, 1 (2017), 143. https://doi.org/10.1186/s12888-017-1306-2

[9] Ronald C. Kessler, Sergio Aguilar-Gaxiola, Jordi Alonso, Somnath Chatterji, Sing Lee, Johan Ormel, T. Bedirhan Üstün, and Philip S. Wang. 2009. The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys. *Epidemiologia e Psichiatria Sociale* 18, 1 (2009), 23–33. https://doi.org/10.1017/S1121189X00001421

[10] R. Korsakiene and D. Diskiene. 2015. Do competencies of entrepreneurs and managers influence internationalization processes? Investigation of Lithuanian SMEs. *Proceedings of the European Conference on Innovation and Entrepreneurship, ECIE* 2015-January (2015), 384–391.

[11] Pierre Levy. 1997. *Collective Intelligence: Mankind's Emerging World in Cyberspace.* Perseus Books, Cambridge, MA, USA.

[12] T.W. Malone and M.S. Bernstein. 2015. *Handbook of Collective Intelligence.* MIT Press.

[13] Raheel Mushtaq. 2014. Relationship Between Loneliness, Psychiatric Disorders and Physical Health? A Review on the Psychological Aspects of Loneliness. *Journal of Clinical and Diagnostic Research* 8, 9 (2014), WE01 – WE04. https://doi.org/10.7860/jcdr/2014/10077.4828

[14] Ikujiro Nonaka and Takeuchi Hirotaka. 1995. *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation.* Oxford University Press.

[15] OECD. 2012. *Health at a Glance: Europe 2012.* OECD Publishing. 154 pages. https://doi.org/10.1787/9789264183896-en

[16] OECD and European Union. 2018. *Health at a Glance: Europe 2018.* OECD Publishing. 212 pages. https://doi.org/10.1787/health_glance_eur-2018-en

[17] Vishwajeet Pattanaik, Alex Norta, Michael Felderer, and Dirk Draheim. 2018. Systematic Support for Full Knowledge Management Lifecycle by Advanced Semantic Annotation Across Information System Boundaries. In *Information Systems in the Big Data Era*, Jan Mendling and Haralambos Mouratidis (Eds.). Springer International Publishing, Cham, 66–73.

[18] Vishwajeet Pattanaik, Shweta Suran, and Dirk Draheim. 2019. Enabling Social Information Exchange via Dynamically Robust Annotations. In *Proceedings of the 21st International Conference on Information Integration and Web-Based Applications & Services* (Munich, Germany) *(iiWAS2019).* Association for Computing Machinery, New York, NY, USA, 176–184. https://doi.org/10.1145/3366030.3366060

[19] Emilio Sacchetti, Alessandro Galluzzo, Adelaide Panariello, Giovanni Parrinello, and Stefano Francesco Cappa. 2008. Self-ordered pointing and visual conditional associative learning tasks in drug-free schizophrenia spectrum disorder patients. *BMC Psychiatry* 8, 1 (Jan. 2008), 8 p. https://doi.org/10.1186/1471-244x-8-6

[20] K. Schrier. 2016. *Knowledge Games: How Playing Games Can Solve Problems, Create Insight, and Make Change.* Johns Hopkins University Press.

[21] S U Shah, A White, S White, and W A Littler. 2004. Heart and mind: (1) relationship between cardiovascular and psychiatric conditions. *Postgraduate Medical Journal* 80, 950 (2004), 683–689. https://doi.org/10.1136/pgmj.2003.014662

[22] S.K. Shivakumar. 2016. *Enterprise Content and Search Management for Building Digital Platforms.* John Wiley & Sons, Ltd.

[23] Mayo Clinic Staff. 2019. Mental illness. https://www.mayoclinic.org/diseases-conditions/mental-illness/symptoms-causes/syc-20374968

[24] Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2020. Frameworks for Collective Intelligence: A Systematic Literature Review. *ACM Comput. Surv.* 53, 1, Article 14 (Feb 2020), 36 pages. https://doi.org/10.1145/3368986

[25] Shweta Suran, Vishwajeet Pattanaik, Sadok Ben Yahia, and Dirk Draheim. 2019. Exploratory Analysis of Collective Intelligence Projects Developed Within the EU-Horizon 2020 Framework. In *Computational Collective Intelligence*, Ngoc Thanh Nguyen, Richard Chbeir, Ernesto Exposito, Philippe Aniorté, and Bogdan Trawiński (Eds.). Springer International Publishing, Cham, 285–296. https://doi.org/10.1007/978-3-030-28374-2_25

[26] Julio Torales, Marcelo O'Higgins, João Mauricio Castaldelli-Maia, and Antonio Ventriglio. 2020. The outbreak of COVID-19 coronavirus and its impact on global mental health. *International Journal of Social Psychiatry* 66, 4 (2020), 317–320. https://doi.org/10.1177/0020764020915212 PMID: 32233719.

Appendix 5

[V]

S. A. Peious, S. Suran, V. Pattanaik, and D. Draheim. Enabling sensemaking and trust in communities. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2021 – the 23rd International Conference on Information Integration and Web-based Applications and Services*, pages 1–9. ACM, 2021

Enabling Sensemaking and Trust in Communities: An Organizational Perspective

Sijo Arakkal Peious
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
sijo.arakkal@taltech.ee

Shweta Suran
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
shweta@taltech.ee

Vishwajeet Pattanaik
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
vishwajeet.pattanaik@taltech.ee

Dirk Draheim
Information Systems Group,
Tallinn University of Technology
Tallinn, Estonia
dirk.draheim@taltech.ee

ABSTRACT

The large volume of information being produced in organizations today poses new challenges to the accuracy and effectiveness of any organizations' decision-making processes. These challenges, namely sensemaking and trust, can critically impact the decision-making processes, even if the organizations are relying on business intelligence (BI) strategies. Given the critical impact an organizations' BI can have on its sustainability and thus its success, in this work, we attempt to draw insights from the literature on collective intelligence and, based on these, present a novel artifact that aims to empower organizations' BI by supporting the organizations' employees in establishing trust and sense when working up with new ideas and solutions. The proposed artifact utilizes a novel reputation model, which calculates reputation based on an individual's area of expertise and reputation score, in order to assist in establishing trust among system users, and thus helps improve decision-making processes.

CCS CONCEPTS

• **Information systems → Crowdsourcing**; • **Software and its engineering →** *Development frameworks and environments*; *Use cases*; Abstraction, modeling and modularity; Designing software.

KEYWORDS

business intelligence, collective intelligence, crowdsourcing, sensemaking, trust, reputation model

ACM Reference Format:
Sijo Arakkal Peious, Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2021. Enabling Sensemaking and Trust in Communities: An Organizational Perspective. In *The 23rd International Conference on Information Integration and Web Intelligence (iiWAS2021), November 29-December 1, 2021, Linz, Austria.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3487664.3487678

iiWAS2021, November 29-December 1, 2021, Linz, Austria

ACM ISBN 978-1-4503-9556-4/21/11…$15.00
https://doi.org/10.1145/3487664.3487678

1 INTRODUCTION

Today's business organizations face endless instabilities and volatilities, which can lead to creation of massive volumes of data; being produced by organizations both internally and externally [!REF]. To harness the possibilities of this transformation, several organizations now aspire (but often even struggle) to convert these large volumes of existing data into a clear understandable chunks that could be utilized in their business processes. In order to achieve this businesses reply on Business Intelligence (BI); a strategy that enables organizations to examine their past actions and decisions, and thus consequently, predict the future. BI denotes a wide range of technologies, processes and applications that assist organizations in gathering, storing, evaluating, and granting access to data for refining business's processes and over-all decision-making [17, 38]. It aids organizations by continuously collecting and analyzing organizational information (including performance metrics) and assists by making the decision-making processes more efficient.

Although BI is a powerful tool and can be typically used in an organization's almost all decision-making processes (both long-term and short-term), however, business organizations today only use BI for day-to-day (i.e., short-term) decision-making [20] and, presently, BI abilities are not necessarily utilized for identifying the organizations' long-term progression, which could indeed help them in improving their methods when undertaking tactical decisions [8]. Another problem that can arise when using BI (which is also often discussed in literature) is sensemaking [35]; this is a key precondition to reach an informed decision and is based on the prior actions of humans [3]. This is to say, that given BI relies on both machine intelligence and human intelligence, when assisting organizations in decision-making; the humans involved in analysis tasks can often get confused by the lack of sense in an idea or an outcome.

Now given that by gaining a better 'sense' of the organization overall, managers (and other decision makers) could better understand their business's organizational environment and hence make healthier decisions [34]; BI applications and related strategies can play a critical part in sensible decision making, and even added advantages beyond conventional decision-making. It is key to note here that, the decisions that are made using BI should be both sensible and explainable and should cover various potential possibilities.

In BI, data/information is used to create reports, summarize past actions, forecast actions, and to understand current and future risks. When relying on BI, the precision of predictions (made using BI strategies) depends on the quality of the information and its sources [39]; if the information and it's source are not trustable, the entire action and its outputs can become futile (even counter-productive). Managers and decision makers who use these outputs typically understand on-going scenarios, and hence create productive decisions or implement their decisions keeping the scenarios in mind [10]; however, a key factor that can influence the decision-making process in such scenarios is 'trust'; specially since humans are involved in the process. Consider this for example, if managers from distinct departments/sections of an organization are working together on creating a solution for given scenario, individuals who have encountered similar scenarios before might be able to contribute more to the solution, however, if the managers are not aware of the past experiences of their colleagues, they might end up not considering ideas of the individual that could contribute the most. This is in line with literature, where researchers have found that, when working together in groups humans tend to make better decisions when there is trust among group-members [29].

That said, in this work we attempt to tackle the above mentioned issues of sensemaking and trust, and propose a novel platform designed as discussion forum oriented towards managers, decision-makers and other employees working in organizations. To achieve this, we draw influence from another domain (one that dates back to Aristotle), that is, Collective Intelligence (CI, defined as, "groups of individuals acting collectively in ways that seem intelligent" [21]); as the domain has recently gained traction in a wide variety of domains [29]. So much so, that it is actively being used by both governing bodies and organizations today; not only to collect citizen/end-user feedback, but also in the design processes for solving critical issues and developing new products, respectively (for example, in Crowd4Roads and CAPSELLA [30], and openIDEO and Threadless [29]). In general, through its fundamental concepts of collection and collaboration, CI has allowed organizations to make better use of the (collective) intelligence of their employees and their users, and thus helps enhance their decision-making processes, when gathering information from numerous sources and creating valuable outputs using CI methods.

With this in mind, the overall aim of this study, is to discover how BI strategies could contribute to better decision-making in presence of sensemaking and trust. The study mainly focuses on the organizations' employee's perspective and tries to identify factors that generate trust between employees and attempts to understand how this trust helps in sensible decision-making processes. In particular, we would like to answer to answer the following research questions:

Q1: *Can we solve the issues of trust and sensemaking in BI using the concepts from CI?*

Q2: *How can we design a reputation model for such a BI system while solving well-known challenges related to reputation in CI platforms?*

The remaining paper is organized as follows, in Section 2, background and related work of BI, CI, trust and reputation systems are described. Then, Section delves into the novel reputation model of trust and sensemaking, and Section discusses the proposed artifact (i.e., the CI-Forum). In Section 5 we describe the evaluation process for the developed forum and reputation model; and finally, Section provides a brief discussion on the findings of this work before concluding the paper.

2 BACKGROUND AND RELATED WORK

Business organizations' performance relies on real-time and effective organizational information. BI systems analyze this information and identify shortcomings and problems within an organization, they provide businesses with insights and suggestions in real-time and support decision-makers in coming up with better conclusions; which subsequently helps organizations sustain and improve productivity [2, 33]. By implementing innovative ideas and new technologies in the their processes, businesses can achieve competitive advantage and success in rapidly changing business conditions [13, 22].

2.1 Business Intelligence

The decision-making processes change according to the information businesses are using to make decisions within their organizations [18]. We can characterize a BI system as a framework that collects, makes modification and generates business's organizational information from different resources. This reduces the time required for analyzing important business information and helps managers to make efficient decisions that can be utilized to improve business strategies. BI is the process of combining different series of actions and business information to provide a competitive advantage to business organizations by helping decision-makers [25]. It is a system and generates answers to support decision-makers to understand the economic situations of the business organizations [23]. Conventionally, BI uses methodological models and numerical functionalities for analysis, used for mining valuable business information and data from basic information to help managers and decision-makers [31]. These business information mining processes and analysis procedures enhance forecasting and help decision-makers understand the progression and problems of any business organization [26].

2.2 Collective Intelligence and Crowdsourcing

General intelligence, as understood by psychologists is the (single) statistical factor that predicts variance in performance, when an individual performs some cognitive tasks (e.g., [11]); it includes an individuals capacity for logic, understanding, learning, reasoning, planning, creativity, critical thinking, problem-solving and many other aspects. When a group of individuals (human or machine) work together and use their individual intelligence, the aggregated intelligence of the group can be understood as CI.

In Information and Communications Technologies (ICT), CI has several definitions (for example, the most prominent ones are by Levy [11] and Malone [21]); each defines CI as having, three components: "individuals (with data/information/knowledge), coordination and collaboration activities (according to a predefined set of rules), and means/platform for real-time communication (viz., hardware/software)"–together these "enable intelligent behavior in groups or crowds" [29]. That said, that advent of the Internet has allowed for mobilization and harnessing of CI in truly novel ways, and this has enabled creation of web-based group discussion platforms that play a key role decision making today [28]. This has opened the gates to a wide variety of emerging research topics, including for example, research where scientists and academicians are trying to understand the influence of group discussion platforms on performance improvement in the quality, efficiency, and effectiveness of decision-making when using such platforms [24]. Some researchers are also focusing on how users behave, group members carry out activities, and knowledge that is generated on group discussion platform by a user and their groups. There also have been studies which focuses on collaborative IT solutions and group discussion systems (designed as web-based platforms), and aim to explain how BI is being used in business organizational context [8].

Another application of CI, that is gaining tremendous interest in research is crowdsourcing (defined as, process where a group of people work together and carry out a task, typically involving collection of data/information or building a solution; that was conventionally done by a single individual [7]. CI (also, crowdsourcing) involves group of people working together, but its key that the individual members of group are diverse [29]. Some researchers have expressed that the main aim of crowdsourcing is to distribute the task of one person to a group of people, and by doing this the overall workload can be divided and hence the task can be carried out effortlessly [6]. Such crowdsourcing activities are divided into three categories. First, directed crowdsourcing, where, a coordinator asks a specific question (with relevant explanation) oriented towards participants, and participants earn some kind of rewards or benefits to the effort and time they contribute. The second category is self-directed, where, participants contribute due to intrinsic motivations. Here participants comes to a common platform and discuss various topics according to their volition and try to come up with decisions or actions according based on the topic at hand. The third and final category is passive, where crowdsourcing is only a side effect of output produced by some action. Here, participant are not obliged to generate the output, or might not be even aware that are participating in a crowdsourced system [36].

A first popular example of crowdsourcing that has is often discussed in literature is the Goldcorp Inc.'s initiative from the year 2000, where they used crowdsourcing to identify gold mines in the Red Lake. The participants were awarded around 0.5 million, and Goldcorp agreed to share the information about the gold mines if they were able to find 6 million ounces of gold, from an identified site. Geologists and engineers from various counties started analysing the information provided by Goldcorp and the company started to receive replies (i.e., potential sites with gold) in a short amount of time. The results produced by participants were verified by a panel decided by Goldcorp, and the end of the competition the panel members were surprised by the both the creativity of the participants and the results produced by them. Goldcorp drilled at the best 5 locations suggested by participants, and found gold from at each of the locations. A key finding of the competition was that participants were able to find gold from all of these locations, without even the locations once. The competition also illustrated how intelligent individuals are, and that by utilizing humans (collective) intelligence combined with technology, organizations could come up with novel and innovative solutions (which could not be achieved conventionally) [5, 37].

2.3 Trust

Reputation and trust are considered key factors of a civilized society [12]. In CI systems too, trust is considered a key property [14, 29]. The success rate of a CI platform can be judged by measuring the trust and openness among the users [4, 14]. An easy method to assess the trustworthiness can be to just ask the users if they trust the source of information [32]. Dworken et al. [15] explained how trust perceived by organizations using examples from the news industry. They claimed that news coverage over the years has changed dramatically, and this is because users have started to analyze both the news and its source to check the reliability of the information [15]. Trust is also a key component in decision making as well as in collaborative working environments [12]. Trust is the belief that the trusted person or the organization will accomplish a particular task according to the task givers expectation [16]. BI applications provide trustable descriptions of various business situations and deliver numerous outcomes for understanding business organizational risks; whoever, as we eluded to earlier, even with the trustable nature of BI applications, trust and sensemaking still remain a challenge to some extent.

2.4 Reputation Systems

Reputation systems are mathematical functions used to calculate a user or objects trustworthiness or value as perceived by fellow users, and is calculated based on user feedback (which can represented using up-votes, stars, like etc.). Theses user score provided by fellow users can be used as a benchmark to identify the level of user trustworthy, and the aggregate of votes and feedback are considered as the reputation score. Literature indicates that theses votes/feedback and thus reputation score can often be violated, thus providing untruthful feedback to gain reputation (supporting non-worthy users) or to decrease the reputation of other users [27]. Reputation systems also face numerous other challenges [1], for instance, Sybil attacks, where attackers (or malicious users) create multiple fake accounts to up-vote their contributions in order to gain higher reputation score, or excessive use of self-promotion, or users with high negative reputations tend to delete old accounts and create new ones (this is referred to as whitewashing). A solution to whitewashing however is that the time duration it takes for an individual to gain reputation can be studied (as was done in [19]), as true good reputation typically only grows gradually. Another challenge to reputation systems is oscillation attack, where, the attacker creates a user account and behaves fairly to achieve good reputation, and then changes their behaviour, hence misleading noble users who trusting the reputation of the attacker [9]. This these challenges in mind, in this work, we aim to develop a novel reputation model that would attempt to tackle some of the challenges described above.

To summarize, this section presented a brief background of literature of BI, CI and reputation systems; this is critical as the review of the literature allowed us to illustrate the purpose of the study, the questions, limits and advantages. It also provides theoretical viewpoints, current views for identifying the study questions and a review of related experimental studies concerning the respective fields. The following section explains the proposed reputation model and how it is different from existing models and systems.

3 NOVEL REPUTATION MODEL

The study proposes a novel approach to the reputation system which aims to avoid the problems explained in the reputation system's literature review. The proposed approach follows a decentralized reputation system. To some extent, this model is similar to existing distributed reputation models like the one used by 'Stack Overflow'. The users give feedback through positive and negative votes (i.e., up-down votes). Whenever a user receives a vote, their reputation score is altered dynamically according to the received votes. In the proposed method, users would not getting the same score every time they receive an up/down vote; instead the amount of score that would added or reduced would depend on the reputation score of the user giving the vote. This means, if a user has a high reputation and they they give an up-vote to another user then the receiving user's reputation score would increase by a higher value, and if the user giving the up-vote does not have any reputation then the receiving user will get the minimal increase in their reputation score. In this approach, the overall score is not calculated while making the vote; but rather the votes are calculated with respect to the category tags (on the individuals profile, i.e., only the topics the user is familiar with) and points are also calculated according to these category tags.

In the proposed reputation model scores are calculated based on the category tags. Whenever a user casts their vote, the system first checks for the reputation score of that user according to the category. If the user has a reputation score, then the system divides that score (that will be added to the receivers reputation) using the total number of votes that the user has received for the particular category. If the calculated score is less than the minimum score, then the receiving user receives the minimum score else they receive the calculated score.

Take for example the following scenario, let us assume there are ten users:

$A_1, A_2, A_3, \ldots, A_{10}$, the minimum reputation score is 1 and we have three categories C_0, C_1, and C_2. At the starting time, T_0 everyone's reputation score is 1. If At a certain time T_1, 5 users are making positive votes for A_6 in respect to category C_0, then his or her reputation score will be $1 + (1 \times 5) = 6$. This score is the overall and $C_0's$ reputation score. At time T_2, if A_6 is casting a positive vote for A_7 with respect to category C_0, then A_7 will receive $1 + (6/5) = 2.2$, 1 is a minimum score of A_7. 6 is reputation score and 5 is the total number of votes for A_6 with respect to C_0. At time T_3, if A_6 is casting a positive vote for A_8 with respect to category C_1, then A_8 will get $1 + 1 = 2$. At time T_4, if A_6 is getting a positive vote from A_9 with respect to C_2, then $1 + 1 = 2$ is added to both category C_2 and the overall reputation score. At the time of T_4, the reputation scores of A_6 are, the overall reputation score is 8, category C_0 is 6, category C_1 is 1 and C_2 is 2.

When generating scores for negative votes (i.e., down votes) the exact same strategy is used, but with subtraction is used instead of addition.

To summarise, in this section, a novel reputation model has been described. The main advantage of the proposed reputation system is, that users can identify the expertise of every user by viewing an overall reputation score and separate score based on every category (the individual contributes/has contributed to). Now to validate this reputation model we have created an artifact, which we delve into in Section 4.

4 PROPOSED CI-FORUM

To study how sensemaking and trust can influence on user behaviour, and to evaluate the previously proposed reputation model here we present as discussion forum (named "CI-Forum"). The proposed artifact allows users to post questions and reply to the questions posted by other users. Users can the platform share knowledge and help other users to solve problems. Users can up-vote or down-vote other users comments and feedback, which in turn is used to calculate user reputation. Users can view posts by using filters, for example sorted based on the reputation scores of the user who posted the question/comment; and thus should be able to identify individuals experts (based on the best answers/comments). The primary notion behind the artifact is that such a CI based forum could potentially be used in line with BI strategies, and would allow organizations to use the collective intelligence of their employees when carrying out decision-making processes.

4.1 Coding and Implementation

The user interface for the artifact is designed using HTML, CSS and JavaScript. To send and receive data, AJAX POST method is used. The CI-Forum website communicates with the server and collects information in the form of JSON objects and files. To make the design process easier and to master coding, pages are used. On the server side, C# is used as the main programming language, together with a layered architecture. The application consists of four layers, i.e., a main project layer, a business logic layer, a data access layer, and a business object layer. The main project layer contains the '.aspx' files. The business logic layer provides all of the logical functionalities for the application. The layer works as a linking layer between the data access layer and the main project layer. The data access layer communicates with the business logic layer and collects data from the database. The business object layer contains objects and their values. Oracle 12C is used as the database.

4.2 System Features

The application has almost all functionality required by a question and answer (Q and A) forum. In addition to this, the application also shows overall and separate reputation scores for each category tag. This view helps the users to identify the best answer concerning the keywords and user. The main functionalities of the application are user creation, login, creating posts, viewing posts, viewing a single post with its answers, viewing reputation scores for every user and a user dashboard. The list of posts can be ordered in several ways, e.g., according to the latest posts, most viewed posts, most commented posts, or most favourites posts. The forum also has the feature to search posts by their titles and tags. The posts are listed in the form of a table, and each row consists of titles, contents, main category, and last three participant posts. Additionally,total number of comments to the post, the total number of viewers, date/time when the post was create are also visible to the users. Users can click on each participants name and view their basic information (including the name of the participant, when they joined the platform, overall reputation scores, reputation scores per category and achieved badges). These attributes were chosen so as to provide users with an overall idea of who their co-members are, thereby assisting in establishing a sort of trustworthiness among members of the community.

Users can click on each post, which then opens the post as a separate page. The post page shows users the posted question, their answers, comments, and edit-options for each post. Each post itself contains the contributor's name, the date when the post was created, its description, up and down vote options, its count and on option to mark the post as favourite. In addition to the question post, there are also options to create answers, make edits and add comments to the post. On the same page, users can see the basic information about the contributor by clicking on the contributor's name. To create a new post, users can select the 'Create New Post' option from the provided menus. Under the 'Create New Post' form, user can add the title, main category, subcategory, description and also upload relevant documents. The options to select tags is provided in the main and subcategory fields. Under the subcategory field, user can select multiple categories, as per their convenience.

To reiterate, a key advantage of designed artifact is that users can view the overall and individual reputation of all their co-members. This would helps users identify the best answers/contributions. The application also has the option to give votes to the other users based on the posts/contributions and behaviour. The code for the designed artifact and the associated database files are openly available as a repo on GitHub (https://github.com/ssijopious/CI-Forum). This is done so that the results presented in the work, can be reproduced and built upon by others.

5 EVALUATION

In this section we attempt to answer the questions we raised previously in this work. The first question, how to implement CI methods in the BI platform so as to solve business organization's decision-making problems related to trust and sensemaking in the process of decision making.

As we eluded to earlier, BI systems can help resolve issues and support in the process of business organizational sensemaking and trust, however it there is a need to create crowd-based platforms to make ensure data quality, flexibility and risk management. And maintaining data quality, requires that the source of the data are given higher priority. To make sure the integrity of the source, we can utilize the collective knowledge of humans using crowd-sourcing methods within BI systems. To entrust a source or user, would require time, and trustable users would need to contribute trustworthy information while also cooperating with other users of the system. The continuous interactions of the user would help develops trust in the platform. This accuracy of trust will have a high impact on the business organizational decision-making processes.

To solve the next question, this study proposes a new reputation model to identify the problems of the CI reputation model and support the BI system to make more trust and sensible decisions. To evaluate this artifact quantitative research method is used. A question and answer platform are created to implement this new reputation system (CI-Forum). A target group is selected for testing this platform and making the evaluations. In this evaluation, we tried to identify the target group's general understanding and habits of the reputation model. The target users are software engineers and IT specialists. Most of the participants have experience in using question and answer platform. The target group is from two different countries. To collect the evaluation, a questionnaire is created.

5.1 Experimental Procedure

To evaluate the designed artifact we conducted lab experiments with multiple users. The candidates for the experiments were identified through social media (primarily Facebook), by using snowballing. More that 50 potential candidates were identified and given presentation on how to use the platform. After the presentation, the candidates (i.e., participants or users) were provided the web address of the application (which was hosted online during the experiments). Each participant was asked to create separate user profiles, and were instructed to create multiple posts (questions, answers and comments). After this, the participants were asked to actively use the platform over the next two weeks. It is important to note here that all participants had a background in software development, hence they were asked to use the platform in the daily workflows. At the end of two weeks, more than 75 questions with multiple answers had been posted on the platform.

After this, all participants were forwarded survey questionnaires, and were given two days to fill in the same. In total, 68 questionnaires were collected at the end of the experiment. Only 45 valid opinions we found, and hence the remaining were 23 questionnaire responses were rejected.

5.2 Reputation Model

To assess the reputation model, the participants we asked questions related to identification of trustable users. This included three questions (given below), and participants were asked to score the questions through Likert scale ranging from (1) indicating 'Completely Disagree' to (5) indicating 'Completely Agree'. The results of the participants feedback is illustrated in Table and Figure .

- Did the CI-forum help the participant to identify the trustworthy user?
- Did the CI-forum help to analyze user expertise?
- Did the CI-Forum provide more overview of the users?

The participants feedback illustrates that the proposed reputation model helped users in identify trustful users. By showing a separate reputation for each category, users were able to identify the area of expertise of their co-members. The platform also helped users gain a better overview of their co-members overall. As indicated in Table and Figure for every question, most of the participants voted for 'Agree' and the average score was more than 3, so we conclude that the reputation model successfully assists users in making sensible decisions through the use of reputation score. The overall score of 3.6 indicates that all participants agreed with the new reputation model approach and were ready to accept the reputation scores. If a user had a high reputation score, then

Figure 1: A screenshot of list of posts as viewed by end-users on the proposed CI-Forum

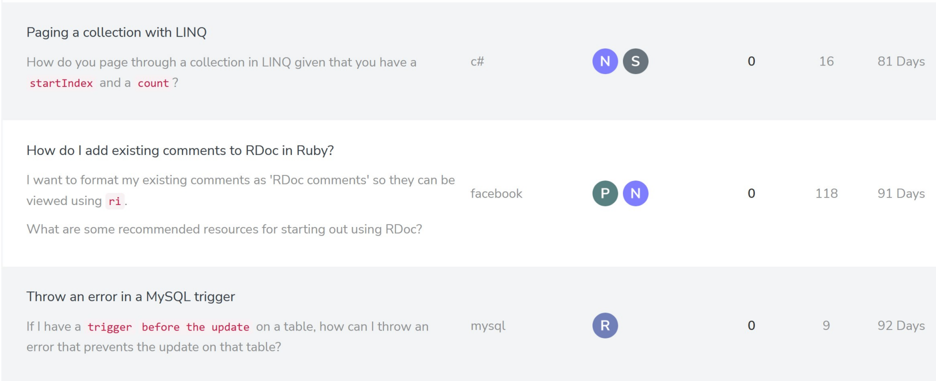


Table 1: User's assessment of the Reputation Model

Sub-factors		Level of Agreements 1	2	3	4	5	Mean	
Trust	*N*	1	5	11	18	10	3.7	Agree
	%	2.2	11.1	24.4	40	22.2		
User Expertise	*N*	1	4	11	16	14	3.9	Agree
	%	2.2	8.9	24.4	35.6	31.4		
Overview of Users	*N*	2	4	19	19	2	3.3	Agree
	%	4.4	8.9	42.2	42.2	2.2		
Total		*4*	*13*	*41*	*53*	*26*	*3.6*	*Agree*

their fellow users considered them as a trustworthy users and accepted their answers. These results also answer the second research question raised in this work. We can create a reputation model to solve the trust problem in BI by showing separate reputation score for each category, as this method benefits users by helping them identify the experts and helps users select the best inputs according to this information. This further aids BI to maintain data quality thereby assisting in sensible decision-making. We argue that this approach compels users to contribute consistently and mimics reputation as it exists in the real-world.

5.3 Usability of CI-Forum

To assess usability, the questionnaire (presented to the participants) contained four questions all revolving around the systems user interface and features. Answers to these provide us an overview of user interactions and the usability and ease-of-use of the designed CI-Forum. These questions again were supposed to be answered using a Likert scale ranging from (1) indicating 'Completely Disagree' to (5) indicating 'Completely Agree'.

- CI-Forum is easy to use or not?
- Are you willing to continue using the CI-Forum?
- Is CI-Forum having a clearer and easier operating interface?
- CI-Forum will be recommended to family and friends?

Table 2: User's feedback regarding the usability of the proposed CI-Forum

Sub-factors		Level of Agreements 1	2	3	4	5	Mean	
Easy to use	*N*	2	7	17	16	3	3.2	Agree
	%	4.4	15.6	37.8	35.6	6.7		
Will continue to use	*N*	0	7	12	20	6	3.6	Agree
	%	0.0	15.6	26.7	44.4	13.3		
Easier operating interface	*N*	0	7	15	18	6	3.6	Agree
	%	0.0	15.6	33.3	40.0	13.3		
Recommended to family and friends	*N*	1	6	17	12	9	3.5	Agree
	%	2.2	13.3	37.8	26.7	20.0		
Total		*3*	*27*	*61*	*66*	*24*	*3.6*	*Agree*

As Table and Figure indicate, the users found the system's interface easy-to-use and the forum in general usable. The users' interaction with CI-Forum were meaningful as they did not face any issues while using the application. Most of the users stated that they would to continue as well as recommended to their friends and family. The mean value of every question was more than 3. The average of the mean value was 3.5, which means that all users were satisfied with their interactions with the CI-Forum. Most users agreed that CI-Forum is useful for their purposes.

During the development phase of the CI-forum, additional feedback was gathered from industry experts, especially those working in the field of software development and testing. These feedback were used to enhance the systems functionalities and usability. Most feedback gathered during this process was positive, and although the experiments with participants was carried out at a smaller sample size, almost all participants simulated actual real-world end-users the CI-forum is oriented towards–as mostly confirmed by

Figure 2: A screenshot of a single posts as viewed by end-users on the proposed CI-Forum

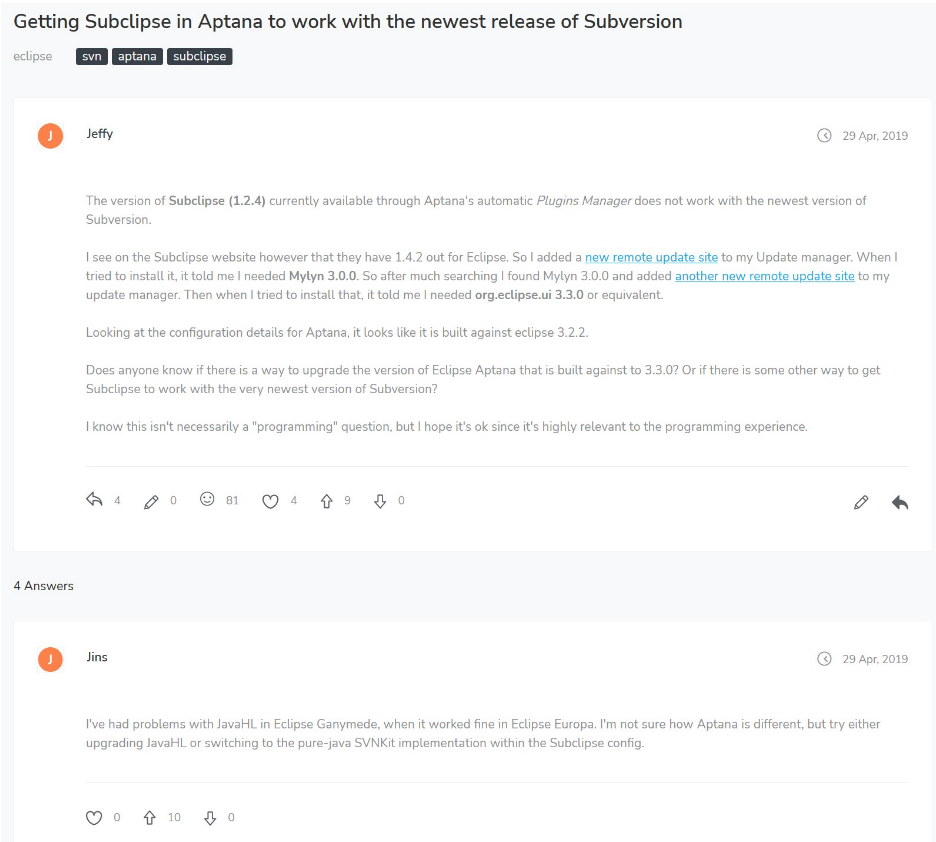
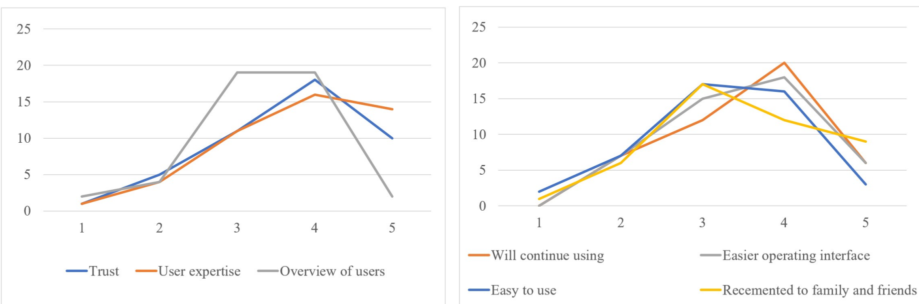


Figure 3: Users' evaluation of the proposed Reputation Model (left) and usability of CI-Forum (right)



the obtained results. The results of the experiments and its following quantitative analysis will be utilized in future to improve the CI-Forum further. The results of the above experiments are only limited by the number and homogeneity of the participant sample, and further user tests are required to develop more conclusive outcomes.

6 CONCLUSION

The overall aim of this work was two main challenges that are encountered when using BI strategies today, these are, sensemaking and trust. Given the critical nature of BI strategies in solving business organizational issues and in supporting organizational decision-making processes; we set out to solve the issues of sensemaking and trust by drawing influences from research in CI. We proposed a novel crowdsourcing approach to reputation models, built around a novel discussion-forum, with focus on organizational employees' perspective and helps establish trust among employees when using BI systems and strategies. By showing users separate reputation scores for each area of expertise, users were able to identify the experts among their fellow users. The idea, behind the approach was that if trustable users works together, the information and results generated by them would be by those organizations is more trustable and sensible to the organizations, specially when compared with non-expert/trustable employees.

The main challenges encountered in this work was that current technologies are still not well adapted to such scenarios. The evaluation of both the reputation model and the CI-forum were only carried out at a small scale, with limited number of participants. Hence the accumulated results only present a superficial view of the usability and usefulness of the proposed contributions. Further changes and fine-tuning is required to enhance the developed artifact. For now, the artifact allows users to identify users with expertise in specific tasks, however, for the next iteration of the forum we would like to develop it so that it can accommodate multi-organization scenarios. Also, as part of future work, we would like to investigate (on a larger scale) and understand the long-term effects of use of reputation scores within organizations and their BI systems.

ACKNOWLEDGEMENT

This work has been developed as part of the "ICT programme" project and was supported by the European Union through European Social Fund.

REFERENCES

[1] Mohammad Allahbakhsh, Boualem Benatallah, Aleksandar Ignjatovic, Hamid Reza Motahari-Nezhad, Elisa Bertino, and Schahram Dustdar. 2013. Quality Control in Crowdsourcing Systems: Issues and Directions. *IEEE Internet Computing* 17, 2 (2013), 76–81. https://doi.org/10.1109/MIC.2013.20

[2] B. Azvine, Z. Cui, D.D. Nauck, and B. Majeed. 2006. Real Time Business Intelligence for the Adaptive Enterprise. In *The 8th IEEE International Conference on E-Commerce Technology and The 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC/EEE'06)*. 29–29. https://doi.org/10.1109/CEC-EEE.2006.73

[3] Richard J. Boland. 2008. *Decision Making and Sensemaking*. Springer Berlin Heidelberg, Berlin, Heidelberg, 55–63. https://doi.org/10.1007/978-3-540-48713-5_3

[4] Efthimios Bothos, Dimitris Apostolou, and Gregoris Mentzas. 2012. Collective intelligence with web-based information aggregation markets: The role of market facilitation in idea management. *Expert Systems with Applications* 39, 1 (2012), 1333–1345. https://doi.org/10.1016/j.eswa.2011.08.014

[5] Daren C. Brabham. 2008. Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence* 14, 1 (2008), 75–90. https://doi.org/10.1177/1354856507084420

[6] Daren C. Brabham. 2013. *Crowdsourcing*. MIT Press. https://mitpress.mit.edu/books/crowdsourcing

[7] Thierry Buecheler, Jan Henrik Sieg, Rudolf Marcel Füchslin, and Rolf Pfeifer. 2010. Crowdsourcing, open innovation and collective intelligence in the scientific method: a research agenda and operational framework. In *The 12th International Conference on the Synthesis and Simulation of Living Systems, Odense, Denmark, 19–23 August 2010*. MIT Press, 679–686. https://doi.org/10.21256/zhaw-4094

[8] Hsinchun Chen, Roger H. L. Chiang, and Veda C. Storey. 2012. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly* 36, 4 (2012), 1165–1188. http://www.jstor.org/stable/41703503

[9] Xiaowen Chu, Xiaowei Chen, Kaiyong Zhao, and Jiangchuan Liu. 2010. Reputation and trust management in heterogeneous peer-to-peer networks. *Telecommunication Systems* 44, 3 (01 Aug 2010), 191–203. https://doi.org/10.1007/s11235-009-9259-5

[10] Thomas D. Clark, Mary C. Jones, and Curtis P. Armstrong. 2007. The Dynamic Structure of Management Support Systems: Theory Development, Research Focus, and Direction. *MIS Quarterly* 31, 3 (2007), 579–615. http://www.jstor.org/stable/25148808

[11] Ian J. Deary. 2012. Intelligence. *Annual Review of Psychology* 63, 1 (2012), 453–482. https://doi.org/10.1146/annurev-psych-120710-100353 arXiv:https://doi.org/10.1146/annurev-psych-120710-100353 PMID: 21943169.

[12] Pierpaolo Dondio and Luca Longo. 2011. Trust-Based Techniques for Collective Intelligence in Social Search Systems. In *Next Generation Data Technologies for Collective Computational Intelligence*, Nik Bessis and Fatos Xhafa (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 113–135. https://doi.org/10.1007/978-3-642-20344-2_5

[13] R. Duane Ireland and Justin W. Webb. 2007. Strategic entrepreneurship: Creating competitive advantage through streams of innovation. *Business Horizons* 50, 1 (2007), 49–59. https://doi.org/10.1016/j.bushor.2006.06.002

[14] Colette Dumas. 2010. Hosting Conversations for Effective Action. *Journal of Knowledge Globalization* 3, 1 (2010), 99 – 116. https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=51902286&site=eds-live

[15] Mark Dworkin, Lois Foreman-Wernet, and Brenda Dervin. 1999. Sense-Making and television news: An inquiry into audience interpretations. *The Electronic Journal of Communication* 9, 2 (1999). http://www.cios.org/EJCPUBLIC/009/2/009217.html

[16] Diego Gambetta et al. 2000. Can we trust trust. *Trust: Making and breaking cooperative relations* 13 (2000), 213–237.

[17] John Hancock and Roger Toren. 2006. *Practical Business Intelligence with Sql Server 2005* (first ed.). Addison-Wesley Professional.

[18] Borut Hočevar and Jurij Jaklič. 2010. Assessing benefits of business intelligence systems–a case study. *Management: journal of contemporary management issues* 15, 1 (2010), 87–119. https://hrcak.srce.hr/53609

[19] Kevin Hoffman, David Zage, and Cristina Nita-Rotaru. 2009. A Survey of Attack and Defense Techniques for Reputation Systems. *ACM Comput. Surv.* 42, 1, Article 1 (Dec. 2009), 31 pages. https://doi.org/10.1145/1592451.1592452

[20] Steve LaValle, Eric Lesser, Rebecca Shockley, Michael S Hopkins, and Nina Kruschwitz. 2011. Big data, analytics and the path from insights to value. *MIT sloan management review* 52, 2 (2011), 21–32. https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/

[21] Thomas W. Malone and Michael S. Bernstein (Eds.). 2015. *Handbook of Collective Intelligence*. The MIT Press, Cambridge, MA. 230 pages.

[22] Ojelanki K. Ngwenyama and Noel Bryson. 1999. Making the information systems outsourcing decision: A transaction cost approach to analyzing outsourcing decision problems. *European Journal of Operational Research* 115, 2 (1999), 351–367. https://doi.org/10.1016/S0377-2217(97)00171-9

[23] Muhmmad I. Nofal and Zawiyah M. Yusof. 2013. Integration of Business Intelligence and Enterprise Resource Planning within Organizations. *Procedia Technology* 11 (2013), 658–665. https://doi.org/10.1016/j.protcy.2013.12.242 4th International Conference on Electrical Engineering and Informatics, ICEEI 2013.

[24] O Pilli. 2014. LMS Vs. SNS: Can social networking sites act as a learning management systems. *American International Journal of Contemporary Research* 4, 5 (2014), 90–97. http://www.aijcrnet.com/journals/Vol_4_No_5_May_2014/9.pdf

[25] V. Pirttimaki. 2007. Conceptual analysis of business intelligence. *SA Journal of Information Management* 9, 2 (2007). https://doi.org/10.4102/sajim.v9i2.24

[26] Mahesh S Raisinghani. 2003. *Business Intelligence in the Digital Economy: Opportunities, Limitations and Risks: Opportunities, Limitations and Risks*. Idea Group Pub. https://books.google.ee/books?id=xKszZbc7RhYC

[27] Paul Resnick and Richard Zeckhauser. 2002. Trust among strangers in internet transactions: Empirical analysis of eBay' s reputation system. Advances in Applied Microeconomics, Vol. 11. Emerald Group Publishing Limited, 127–157. https://doi.org/10.1016/S0278-0984(02)11030-3

[28] J.P. Shim, Merrill Warkentin, James F. Courtney, Daniel J. Power, Ramesh Sharda, and Christer Carlsson. 2002. Past, present, and future of decision support technology. *Decision Support Systems* 33, 2 (2002), 111–126. https:

//doi.org/10.1016/S0167-9236(01)00139-7 Decision Support System: Directions for the Nest Decade.

[29] Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2020. Frameworks for Collective Intelligence. *Comput. Surveys* 53, 1 (may 2020), 1–36. https://doi.org/10.1145/3368986

[30] Shweta Suran, Vishwajeet Pattanaik, Sadok Ben Yahia, and Dirk Draheim. 2019. Exploratory Analysis of Collective Intelligence Projects Developed Within the EU-Horizon 2020 Framework. In *Computational Collective Intelligence*, Ngoc Thanh Nguyen, Richard Chbeir, Ernesto Exposito, Philippe Aniorté, and Bogdan Trawiński (Eds.). Springer International Publishing, Cham, 285–296.

[31] Carlo Vercellis. 2011. *Business intelligence: data mining and optimization for decision making*. Wiley Online Library. https://bit.ly/3FrU9fZ

[32] C. Nadine Wathen and Jacquelyn Burkell. 2002. Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology* 53, 2 (2002), 134–144. https://doi.org/10.1002/asi.10016

[33] Hugh J. Watson and Barbara H. Wixom. 2007. The Current State of Business Intelligence. *Computer* 40, 9 (2007), 96–99. https://doi.org/10.1109/MC.2007.331

[34] K.E. Weick. 2012. *Making Sense of the Organization, Volume 2: The Impermanent Organization*. Wiley. https://bit.ly/3oEHIY4

[35] K.E. Weick and K.E.W. Weick. 1995. *Sensemaking in Organizations*. SAGE Publications. https://bit.ly/3BkaKzT

[36] Michael Weiss. 2016. Crowdsourcing Literature Reviews in New Domains. *Technology Innovation Management Review* 6 (02/2016 2016), 5–14. https://doi.org/10.22215/timreview/963

[37] Sean Wise, Robert A. Paton, and Thomas Gegenhuber. 2012. Value co-creation through collective intelligence in the public sector. *VINE* 42, 2 (01 Jan 2012), 251–276. https://doi.org/10.1108/03055721211227273

[38] Barbara Wixom and Hugh Watson. 2010. The BI-Based Organization. *International Journal of Business Intelligence Research (IJBIR)* 1, 1 (2010), 13–28. https://doi.org/10.4018/jbir.2010071702

[39] Öykü Işık, Mary C. Jones, and Anna Sidorova. 2013. Business intelligence success: The roles of BI capabilities and decision environments. *Information & Management* 50, 1 (2013), 13–23. https://doi.org/10.1016/j.im.2012.12.001

Curriculum Vitae

1. Personal data

Name	Shweta Suran
Date and place of birth	19 March 1990, Delhi, India
Nationality	Indian

2. Contact information

Address	Tallinn University of Technology, School of Information Technologies, Department of Software Science, Akadeemia tee 15a, 12618 Tallinn, Estonia
Phone	+372 5803 7683
E-mail	shweta@taltech.ee

3. Education

2017–. . .	Tallinn University of Technology, School of Information Technologies, Computer Science, Doctor of Philosophy
2012–2014	SRM University, School of Computing, Knowledge Engineering, M.Tech
2007–2011	Dr. A.P.J. Abdul Kalam Technical University, Department of Computer Science, Computer Science Engineering, B.Tech

4. Language competence

Hindi	native
English	fluent
Sanskrit	basic level

5. Professional employment

2021–2021	Information Systems Group, Tallinn University of Technology, Analyst
2017–2021	Information Systems Group, Tallinn University of Technology, Early Stage Researcher
2015–2017	Krishna Engineering College (Dr. A.P.J. Abdul Kalam University), Assistant Professor
2014–2015	Cognizant Technology Solutions, Programmer Analyst
2011–2011	iYogi Technical Services Pvt. Ltd., Technical Specialist

6. Voluntary work

2012–2013	Worked on SRM - PURA (Providing Urban Amenities to Rural Areas)
2007–2010	Worked as a volunteer for Diabetes Health Awareness Camp

7. Computer skills

- Operating systems: MS Windows, GNU/Linux
- Document preparation: MS Office, Libre Office, Latex, VS Code

- Programming languages: Python, Bash, C, C++, Java
- Scientific packages: MATLAB, R

8. Honours and awards

- 2016, Received SAP Fellowship Award for being one of the Top 200 Performing Participants in “Use of ICT in Education for Online and Blended Learning” Faculty Development Program by IIT Bombay, India
- 2014, Received Gold Medal for securing First Rank with Honours in M. Tech - Knowledge Engineering, SRM Institute of Science and Technology (SRM University), India

9. Defended theses

- 2014, “Cerebral Vessel Delineation Based On Segmentation and Skeletonization”, M.Tech, supervisor Prof. Dr. D. Malathi, SRM Institute of Science and Technology (SRM University), School of Computing

10. Fields of research

- Collective Intelligence
- Digital Image Processing
- Knowledge Engineering

11. Scientific work
Papers

1. S. A. Peious, S. Suran, V. Pattanaik, and D. Draheim. Enabling sensemaking and trust in communities. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2021 – the 23rd International Conference on Information Integration and Web-based Applications and Services*, pages 1–9. ACM, 2021
2. S. Suran, V. Pattanaik, and D. Draheim. CommunityCare: Tackling mental health issues with the help of community. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2020 – the 22nd International Conference on Information Integration and Web-based Applications and Services*, pages 377–382. ACM, 2020
3. S. Suran, V. Pattanaik, and D. Draheim. Frameworks for collective intelligence: A systematic literature review. *ACM Computing Surveys*, 53(1):1–36, 2020
4. S. Suran, V. Pattanaik, S. B. Yahia, and D. Draheim. Exploratory analysis of collective intelligence projects developed within the EU Horizon 2020 framework. In N. T. Nguyen, R. Chbeir, E. Exposito, P. Aniorté, and B. Trawiński, editors, *Proceedings of ICCCI 2019 – the 11th International Conference on Computational Collective Intelligence*, Lecture Notes in Artificial Intelligence 11684, pages 285–296. Springer, 2019
5. V. Pattanaik, S. Suran, and D. Draheim. Enabling social information exchange via dynamically robust annotations. In M. Indrawan-Santiago, E. Pardede, I. L. Salvadori, M. Steinbauer, I. Khalil, and G. Anderst-Kotsis, editors, *Proceedings of iiWAS'2019 – the 21st International Conference on Information Integration and Web-based Applications and Services*, pages 176–184. ACM, 2019

6. S. Suran, V. Pattanaik, M. Singh, P. Gupta, and P. Gupta. Brain imaging procedures and surgery techniques: Past, present and future. *International Journal of Bio-Science and Bio-Technology*, 9(3):23–34, 2017

7. H. Tyagi, S. Suran, and V. Pattanaik. Weather – temperature pattern prediction and anomaly identification using artificial neural network. *International Journal of Computer Applications*, 140(3):15–21, 2016

8. S. Suran, V. Pattanaik, and D. Malathi. Discovering shortest path between points in cerebrovascular system. In *Proceedings of I-CARE'2014 – the 6th IBM Collaborative Academia Research Exchange Conference*. ACM Press, 2014

9. V. Pattanaik, S. Suran, and S. Prabakaran. Inducing human-like motion in robots. In *Proceedings of I-CARE 2014 – the 6th IBM Collaborative Academia Research Exchange Conference*. ACM Press, 2014

Elulookirjeldus

1. Isikuandmed

Nimi	Shweta Suran
Sünniaeg ja -koht	19. Märts 1990, Delhi, India
Kodakondsus	India

2. Kontaktandmed

Aadress	Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Tarkvarateaduse instituut, Akadeemia tee 15a, 12618 Tallinn, Estonia
Telefon	+372 5803 7683
E-post	shweta@taltech.ee

3. Haridus

2017–...	Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Informaatika, Filosoofiadoktor
2012–2014	SRM University, School of Computing, Knowledge Engineering, M.Tech
2007–2011	Dr. A.P.J. Abdul Kalam Technical University, Department of Computer Science, Computer Science Engineering, B.Tech

4. Keelteoskus

hindi keel	emakeel
inglise keel	kõrgtase
Sanskrit keel	põhitase

5. Teenistuskäik

2021–2021	Infosüsteemide Töörühm, Tallinna Tehnikaülikool, Analüütik
2017–2021	Infosüsteemide Töörühm, Tallinna Tehnikaülikool, Doktorant-nooremteadur
2015–2017	Krishna Engineering College (Dr. A.P.J. Abdul Kalam University), Assistant Professor
2014–2015	Cognizant Technology Solutions, Programmer Analyst
2011–2011	iYogi Technical Services Pvt. Ltd., Technical Specialist

6. Vabatahtlik töö

2012–2013	Töötanud PURA-projekti jaoks (maapiirkondade taristu arendamise projekt Indias)
2007–2010	Töötanud vabatahtlikuna diabeedialase teadlikkuse suurendamise laagris (Diabetes Health Awareness Camp)

7. Arvutioskused

- Operatsioonisüsteemid: MS Windows, GNU/Linux

- Kontoritarkvara: MS Office, Libre Office, Latex, VS Code
- Programmeerimiskeeled: Python, Bash, C, C++, Java
- Teadustarkvara paketid: MATLAB, R

8. Autasud

- 2016, SAP stipendium 200 parima osaleja hulka jõudmise eest IIT Bombay (India) õppejõudude arenguprogrammis „Info-ja kommunikatsioonitehnoloogia kasutamine hariduses veebipõhises ja kombineeritud õppe".
- 2014, kuldmedal magistriõppe kiitusega lõpetamise eest. MTech teadmistepõhise inseneriteaduse erialal, SRM Teadus- ja Tehnoloogiainstituut (SRM Ülikool), India.

9. Kaitstud lõputööd

- 2014, "Cerebral Vessel Delineation Based On Segmentation and Skeletonization", M.Tech, juhendaja Prof. Dr. D. Malathi, SRM Institute of Science and Technology (SRM University), School of Computing

10. Teadustöö põhisuunad

- Kollektiivne intelligentsus
- Digitaalne pilditöötlus
- Teadmistepõhine inseneriteadus

11. Teadustegevus
Teadusartiklite, konverentsiteeside ja konverentsiettekannete loetelu on toodud ingliskeelse elulookirjelduse juures.

ISSN 2585-6901 (PDF)
ISBN 978-9949-83-818-9 (PDF)