DOCTORAL THESIS

Digital Transformation: Artificial Intelligence Enablement in Public Services

Richard Michael Dreyling III

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2024

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
48/2024

Digital Transformation: Artificial Intelligence Enablement in Public Services

RICHARD MICHAEL DREYLING III

PRESS

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

The dissertation was accepted for the defense of the Doctor of Philosophy (Computer Science) degree on 10 September 2024

Supervisor: Associate Professor Dr. Ingrid Pappel,
Department of Software Science,
School of Information Technologies,
Tallinn University of Technology,
Tallinn, Estonia

Co-supervisor: Professor Dr. Tanel Tammet,
Department of Software Science,
School of Information Technologies,
Tallinn University of Technology,
Tallinn, Estonia

Opponents: Associate Professor Mihkel Solvak,
University of Tartu,
Tartu, Estonia

Professor Anis Yazidi,
Oslo Metropolitan University,
Oslo, Norway

Defence of the thesis: 25 September 2024, Tallinn, Estonia

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.

Richard Michael Dreyling III

signature

European Union
European Regional
Development Fund

Investing
in your future

Copyright: Richard Michael Dreyling III, 2024
ISSN 2585-6898 (publication)
ISBN 978-9916-80-198-7 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9916-80-199-4 (PDF)
DOI https://doi.org/10.23658/taltech.48/2024
Printed by EVG Print

Citation: Dreyling, R.M. (2024). *Digital Transformation: Artificial Intelligence Enablement in Public Services* [TalTech Press]. https://doi.org/10.23658/taltech.48/2024

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
48/2024

Digisiire ja tehisintellekti rakenduseeldused avalikes teenustes

RICHARD MICHAEL DREYLING III

TALTECH KIRJASTUS

Contents

List of Publications 7

Author's Contributions to the Publications 8

Other Publications 9

Abbreviations 10

Terms 11

1 Introduction 12

2 Focus and Aim 14
  2.1 Research Questions 14

3 Research Methodology 17

4 Related Works 20
  4.1 Theoretical Overview 20
    4.1.1 Technology Acceptance 20
    4.1.2 Change Management 21
    4.1.3 Organizational and Institutional Theories 21
  4.2 State of the Art 22
    4.2.1 E-Governance and the Digital State 22
    4.2.2 AI in the Public Sector 23
    4.2.3 AI Use in Public Services 24
    4.2.4 Artificial Intelligence Organizational Readiness and Maturity 25

5 Results 26
  5.1 Leveraging Digital Transformation For AI-Enabled Services 26
  5.2 Potential for AI-Enablement in Public Services[I] 26
  5.3 AI-Enablement Adoption for Existing Public Services[I] 27
  5.4 Key Steps and Challenges[I] 28
  5.5 A technical and organizational architecture based on social, legal, and strategic considerations[V] 30
  5.6 Organizational Maturity to adopt AI-enabled digital services[III] 31
  5.7 Ensuring ethical and responsible AI implementation for Digital public services[II, III, VI] 32

6 Limitations and Future Work 35

7 Conclusion 37

List of Figures 38

List of Tables 39

References 40

Acknowledgements 47

Abstract 49

Kokkuvõte 50

Publication I 51

Publication II 65

Publication III 75

Publication IV 93

Publication V 109
Publication VI 125
Curriculum Vitae 138
Elulookirjeldus 141

List of Publications

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

I R. Dreyling, T. Tammet, and I. Pappel, "Technology push in ai-enabled services: How to master technology integration in case of bürokratt," *SN Computer Science*, vol. Future Data Science Engineering, no. 5, p. 738, 2024

II R. Dreyling, E. B. Jackson, T. Tammet, A. Labanava, and I. Pappel, "Social, legal, and technical considerations for machine learning and artificial intelligence systems in government.," in *ICEIS (1)*, pp. 701–708, 2021

III R. Dreyling, J. Lemmik, T. Tammet, and I. Pappel, "An artificial intelligence maturity model for the public sector: Design science approach," *TalTech Journal of Social Sciences*, vol. 14, no. 2, p. 16, Forthcoming

IV R. M. Dreyling, T. Tammet, and I. Pappel, "Artificial intelligence use in e-government services: A systematic interdisciplinary literature review," in *International Conference on Future Data and Security Engineering*, pp. 547–559, Springer, 2022

V R. Dreyling, K. McBride, T. Tammet, and I. Pappel, "Navigating the AI maze: Lessons from Estonia's Bürokratt on public sector AI digital transformation," *SSRN*, 2024

VI R. Dreyling, T. Koppel, T. Tammet, and I. Pappel, "Challenges of genai chatbots in public services: An integrative review," *SSRN*, 2024

Author’s Contributions to the Publications

Contribution to the papers in this thesis are:

I The author of this thesis is the lead author of this article (first author and corresponding author), responsible for the majority of the article content, including data collection, data analysis, implementation and evaluation of the approach, and manuscript writing.

II The author of this thesis is the lead author of this article (first author and corresponding author), responsible for the majority of the article content, including data collection, data analysis, implementation and evaluation of the approach, and manuscript writing.

III The author of this thesis is the lead author of this article (first author and corresponding author), responsible for the majority of the article content, including reviewing all the articles, evaluating their relevance, methodology, findings, and manuscript writing.

IV The author of this thesis is the lead author of this article (first author and corresponding author), responsible for the majority of the article content, including reviewing all the articles, evaluating their relevance, methodology, findings, and manuscript writing.

V The author of this thesis is the lead author of this article (first author and corresponding author), responsible for the majority of the article content, including data collection, data analysis, implementation and evaluation of the approach, and manuscript writing.

VI The author of this thesis is the lead author of this article (first author and corresponding author), responsible for a large portion of the article content, including reviewing all the articles, evaluating their relevance, methodology, findings, and manuscript writing.

Other Publications

The author has contributed to other publications while studying at Tallinn University of Technology, which do not comprise the primary line of scientific inquiry and are not included in this thesis but may be cited where necessary.

VII R. Dreyling, E. Jackson, and I. Pappel, "Cyber security risk analysis for a virtual assistant g2c digital service using fair model," in *2021 Eighth International Conference on eDemocracy and eGovernment (ICEDEG)*, pp. 33–40, 2021

VIII R. Dreyling, R. Erlenheim, T. Tammet, and I. Pappel, "Ai readiness assessment for data-driven public service projects: Change management and human elements of procurement," *Human Factors, Business Management and Society*, vol. 97, no. 97, 2023

IX R. M. Dreyling III, T. Tammet, and I. Pappel, "Digital transformation insights from an ai solution in search of a problem," in *International Conference on Future Data and Security Engineering*, pp. 341–351, Springer, 2023

X E. Blake Jackson, R. Dreyling, and I. Pappel, "A historical analysis on interoperability in estonian data exchange architecture: Perspectives from the past and for the future," in *Proceedings of ICEGOV'21 – the 14th International Conference on Theory and Practice of Electronic Governance*, pp. 111–116, ACM, 2021

XI E. B. Jackson, R. Dreyling, and I. Pappel, "Challenges and implications of the who's digital cross-border covid-19 vaccine passport recognition pilot," in *2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG)*, pp. 88–94, IEEE, 2021

XII A. Labanava, R. M. Dreyling III, M. Mortati, I. Liiv, and I. Pappel, "Capacity building in government: Towards developing a standard for a functional specialist in ai for public services," in *International Conference on Future Data and Security Engineering*, pp. 503–516, Springer, 2022

XIII A. Labanava, R. M. Dreyling, and A. Norta, "Potential of smart contracts in the pharmaceutical supply chain of belarus," in *2022 IEEE 1st Global Emerging Technology Blockchain Forum: Blockchain & Beyond (iGETblockchain)*, pp. 1–6, IEEE, 2022

Abbreviations

AI	Artificial Intelligence
AIMM	Artificial Intelligence Maturity Model
CDO	Chief Data Officer
CMMI	Capability Maturity Model Integration
CTO	Chief Technology Officer
DT	Digital Transformation
eID	electronic Identification
GA	General Availability
GDPR	General Data Protection Law
GPT	Generative Pre-training Transformer
GUI	Graphical User Interface
HCAI	Human Centered Artificial Intelligence
HCD	Human-Centered Design
ICT	Information and Communication Technology
LLM	Large Language Model
MKM	Ministry of Economic Affairs and Communication/Majandus - ja Kommunikatsiooniministeerium
NIIS	Nordic Institute for Interoperability Studies
NLP	Natural Language Processing
OKR	Objectives and Key Results
PDPA	Personal Data Protection Law
REST API	Representational State Transfer Application Programming Interface
RIA	State Information Authority/Riigi Infosüsteemi Ameti
SLR	Systematic Literature Review
TARA	State Authentication Service/riigi autentimisteenuse
TAM	Technology Acceptance Model
TPB	Theory of Planned Behavior
TRA	Theory of Reasoned Action
UTAUT	Unified Theory of Acceptance and Use of Technology
XAI	Explainable Artificial Intelligence
YAML	Yet Another Markup Language

Terms

Artificial Intelligence System	According to NIST AI 100-1, "The AI RMF refers to an AI system as an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy (Adapted from: OECD Recommendation on AI:2019; ISO/IEC 22989:2022)." And according to the EU AI Act, "(1) 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments." data, information and/or evidence in the identity-proofing process.
Bürokratt	According to the Ministry of Economic Affairs and Communication in 2024, "The virtual assistant Bürokratt is an interoperable network of chatbots on the websites of public authorities that allows people to obtain information from these authorities through a chat window. It provides individuals, or users, with the opportunity to access direct public and information services using virtual assistants."
Bürokratt (original vision)	According to the Open Source Observatory (OSOR) in 2022, "The open source solution Bürokratt is a digital assistant that uses AI to establish an interoperable network of solutions, allowing citizens to perform various public service procedures through a single platform. The conception of Bürokratt followed the release of proposals on advancing the take-up of AI by the MKM. Legitimised by Estonia's national AI strategy, Bürokratt is envisioned as the way "public services should digitally work in the age of artificial intelligence (AI)"."
Digital transformation	"to emphasize the cultural, organizational, and relational changes that we highlight in the outcomes section in order to differentiate better between different forms of outcomes." [14]
Digitization	"to highlight the transition from analog to digital services with a 1:1 change in the delivery more and the addition of a technolo- gical channel of delivery." [14]
Digitalization	"to focus on potential changes in the processes beyond mere digitizing of existing processes and forms; " [14].

1 Introduction

Digital transformation (DT) is the adoption process of information technologies that causes recurrent feedback effects [14, 15]. Compared to digitalization or digitization, a digital transformation produces much more reflexive feedback [14]. Digital transformation is essential from a strategic level and, by its nature, means causing "significant changes" in an organization [15].

Artificial intelligence (AI) has a level of complexity tied to the nature of the technology, which uses algorithms and significant amounts of data to produce results [16], which are often hard or impossible to predict. The implementation of artificial intelligence adds additional levels of complexity [16] to this already complex process. Environmental pressures from internal and external sources [14] can cause the demand for digital transformations, which is currently occurring in governments across the world [17]. AI implementations promise efficiency and effectiveness gains for public administrations, but these can be difficult to capture [18].

When digital transformations in the public sector fail, they can have a large amount of waste [19] or potentially have a negative impact on the public [20]. Understanding how to adopt and implement artificial intelligence better can mitigate risk and support value creation for the public. This requires addressing technical, institutional, and organizational questions [21, 22, 23].

After the introduction of general-purpose technologies, it typically takes five to ten years to see gains in the total value created, as measured by economic techniques [21]. This time frame encompasses the development of accompanying enabling technologies, as well as the adoption and adaptation of the technology by organizations and institutions to ensure its successful implementation. This not only includes tactical and practical techniques, but it also requires significant institutional and organizational shifts [22, 23] to ensure the full integration of the technology into organizations' strategy and business processes within the institutional context.

By better reckoning with the complexity of AI implementations in the public sector and developing AI-enabled public services, public administrations can help bring the value of this relatively new general-purpose technology to residents and citizens. AI is becoming ubiquitous. As the technology diffuses across the world's populations [24], it affects the expectations and demands of those people toward governments, resulting in a form of external pressure in addition to the mimetic effect building internal pressure in institutions as more successful implementations occur [22] can become a demand for digital transformation [15]. Meanwhile, governments are attempting to manage the many demands posed by adopting technologies with many complicating factors to achieve efficiency and effectiveness gains [25] while reducing the potential adverse outcomes [16, 26].

When considering the many aspects of adopting AI in the context of public administration, various academic disciplines come into play. E-governance literature can provide lessons on implementing governance mechanisms through electronic means and creating public-facing digital services [27, 28]. Other insights include the importance of trust in the adoption of e-governance technologies by the public [29, 30]. Digital transformation-related theories from many disciplines are also required to understand the adoption of technologies that cause a recursive feedback mechanism within organizations and institutions. These can come from public administration literature [14, 15] change management literature [31, 32], institutional theory [22, 23], economics [33], and even business administration [34, 26].

Recent work [35] indicates that public administrations are adopting many pilots and implementations of artificial intelligence across the European Union. Many of these uses are related to chatbots and decision support systems. At the same time that these administrations are adopting AI-enabled chatbots, advances in other spheres, like the proliferation of generative AI technologies, can increase the expectations of humans who use the systems. The fast-paced advancement of technology makes keeping up with the standard technologies very difficult, as organizations face challenges maintaining their standard services while building new offerings [36]. Compared with private sector organizations, public institutions have an even more enormous challenge in facing the "innovator's dilemma," [36] due to legally defined organizational structures [37] that mirror machine organizations [38] and challenges with procurement and recruiting talent competent in emerging technologies like AI [**?**].

As a result of the first Estonian Strategy on AI put forth by the Minsitry of Economic Affairs and Communication[39] and later vision documents Estonia is developing and implementing a platform to become an AI-enabled digital public service. The *Bürokratt program* is currently a chatbot that can answer questions and is available on the web pages of the individual ministries that have chosen to adopt the program. However, the goal is that it will eventually be a network of interoperable agents able to not only execute government services through an AI-enabled chatbot [40], but it will ultimately be able to use voice-based virtual assistants to execute these public service transactions.

The main aim of this thesis is to investigate the digital transformation related to the development, implementation, and adoption of AI-enabled digital public services from various relevant perspectives. The focus is to gain an understanding that not only contributes academically to this rapidly expanding area of technology but also provides public decision-makers with insights into the challenges they may face when attempting to follow a similar path towards AI-enablement of digital public services. The thesis may be useful to decision-makers considering AI-enablement of digital public services or those private sector groups considering providing these services.

This thesis is arranged to make the progression of the argument and results clear to the reader. This includes the following chapters:

- **Chapter 2:** "Focus and Aim": This chapter expounds upon the focus and aim of the thesis and explains the research questions and why they were chosen.

- **Chapter 3:** "Methods and Research Design": This chapter delineates the research design and methodology of the thesis as a whole and all of the publications included in the thesis.

- **Chapter 4:** "Related Works": This chapter explains the academic literature context within which this thesis is contained in a theoretical overview and state of the art. The research gap is also addressed.

- **Chapter 5:** "Results": This chapter details some of the highlights of the contributions of the papers that answer the research questions.

- **Chapter 6:** "Limitations and Future Work": This chapter includes the limitations and future work.

- **Chapter 7:** "Conclusion": The final chapter summarizes the contributions and critical findings.

2 Focus and Aim

In Estonia, the improvement of their renowned e-governance took a notable step forward when a vision paper outlining AI-enabled services as the next evolution of their digital services infrastructure was released [40]. Around this time, the concept for this thesis and the research that comprises it came into being. The research's original purpose was to investigate virtual assistant providers' ability to complement and adhere to the then-named KrattAI concepts. However, initial research interviews with stakeholders in the program led the author to focus on more general concepts because the Estonian government was already in discussions with the most prominent providers under legal non-disclosure agreements. The necessity to ground the research in the more extensive scientific theory and empirical evidence also provided the impetus to shift toward analyzing the digital transformation process.

To this end, initial research focused on the many challenges associated with implementing AI in government from the technical, social, and legal perspectives. After gaining an understanding of some of the concepts that comprised this aspect of the research field, the author focused efforts on understanding digital transformation and the applicable scientific concepts on which it is based.

With this new knowledge, the author approached the planning, development, and implementation processes that have occurred and are currently happening in the Bürokratt program in Estonia from a technical and organizational perspective. Following this research, the author focused on instantiating and transferring this knowledge to public administrations looking to adopt AI-enabled digital public services through an Artificial Intelligence Maturity Model for the Public Sector.

2.1 Research Questions

The research questions focus on creating AI-enabled digital public services, including challenges and relevant considerations that will be illustrated based on the Estonian experience with the Bürokratt program. The main research question is a meta-question designed to be answered when the subsequent research questions are answered. Adopting AI-enabled services is a complicated topic, and thus, there are multiple research questions and subquestions to attempt to contribute to the topic in a meaningful way. Following the main research question are three research questions, the final two of which have subquestions.

- **MRQ**: How can public sector organizations leverage digital transformation to facilitate the adoption of AI-enabled services?

- **RQ.1**: What is the current state of digital transformation in the public sector organizations, and how does it affect the potential for AI enablement in public services?

- **RQ.2**: How would an existing public service adopt AI enablement?

  2.1: What are the key steps and challenges involved in this process?

  2.2: What social, legal, and strategic considerations influence the technical and organizational architecture of public sector organizations implementing AI-enabled e-services?

- **RQ.3**: How would one analyze the readiness of organizations to adopt AI enablement for e-services?

3.1: What existing frameworks exist for readiness evaluations of AI projects, and how can they be adapted to the public sector e-service application?

3.2: How can public sector organizations ensure ethical and responsible AI implementation in e-services?

When this research concept was conceived, the program, initially called KrattAI, went through a pilot stage and eventually into development and implementation. The pilot had a somewhat different architecture than that, which ultimately became the final architecture. However, this program's evolution and the ability to analyze the digital transformation as it occurred allowed the author to investigate the technological and organizational process considerations across time. The research questions were developed and iterated with this in mind.

From the literature and interviews with experts conducted while carrying out research, more on this is presented in Chapter 2. It also became apparent that there is a need to analyze organizations' maturity toward AI adoption. This became research questions 1.1 and 1.2.

These research questions combine theoretical and practical contributions that are important to the author and the field, which is why they were chosen. These research questions come partially from a research gap related to practical implementations of AI systems developed from a strategy initially and AI systems in the public sector directly relating to e-services or digital state services. The other consideration is the practical difficulty governments seem to reckon with AI implementation and adoption. This starts from the conception of pilots, considering the ability of public institutions to accomplish their objectives all the way through the digital transformation process. The following publications are a part of this thesis and seek to answer the research questions meaningfully.

Table 1: Correlation of the research publications to the research questions

AI-Enablement of Digital Public Services						
RQ No	**I**	**II**	**III**	**IV**	**V**	**VI**
MRQ	X	X		X	X	
RQ 1	X			X	X	
RQ 2	X					
RQ 2.1	X					
RQ 2.2		X				
RQ 3			X			
RQ 3.1			X			
RQ 3.2		X	X	X		X

Publication I, **"Technology Push in AI-Enabled Services: How to master Technology Integration in Case of Bürokratt"**, builds upon previous research in publications [IV, II] and addresses the main research question as well as research questions one and two through a qualitative case study about the planning and technical integration decisions and processes that the Estonian Government has taken while going from a strategy and vision to a usable product for residents and citizens.

Publication II, **"Social, Legal, and Technical Considerations for Machine Learning and Artificial Intelligence Systems in Government"**, addresses the first research question as well as sub-questions 2.2 and 3.2 by exploring the many ways in which governments must take into account social, legal, and technical ramifications of machine learning systems they design, develop, and deploy.

Publication III, **"An Artificial Intelligence Maturity Model for the Public Sector: Design Science Approach"** addresses research question three and the suborning sub-questions by introducing an artificial intelligence maturity model for the public sector, which distills many of the learnings of the research into an artifact which gives guidance on maturity in organizations related to AI adoption in the public sector.

Publication IV, **"Artificial Intelligence Use in e-Government Services: A Systematic Interdisciplinary Literature Review"**, addresses the main research question and research question three sub-question two. This early piece of research in the thesis helped to lay the groundwork for understanding what the state of the art was as it pertained to the use of AI in e-government services.

Publication V, **"Navigating the AI Maze: Lessons from Estonia's Bürokratt on Public Sector AI Digital Transformation"**, builds further the research conducted in (Publications I, II, and IV) and analyzes the digital transformation of Bürokratt from an institutional and organizational framework that aids in answering the main research question and research question one in a descriptive manner.

Publication VI, **"Challenges of Generative AI Chatbots in Public Services – An Integrative Review"**, takes into account recent developments in the area of artificial intelligence that has come with the expanded availability and adoption of large language models and addresses how governments may use generative AI chatbots in the public services through an integrative review of the literature in relevant fields giving a forward-looking answer to research question 3.2.

3 Research Methodology

This thesis uses mixed methods. Six relevant papers comprise this thesis. The thesis is considered mixed methods because not all of the studies that comprise the thesis use the same methodology [41]. The research design includes three qualitative methods based case studies of different variety [II, I, V], one systematic literature review [IV], one integrative review [VI], and a design science paper [III]. Figure 1 shows the research design.

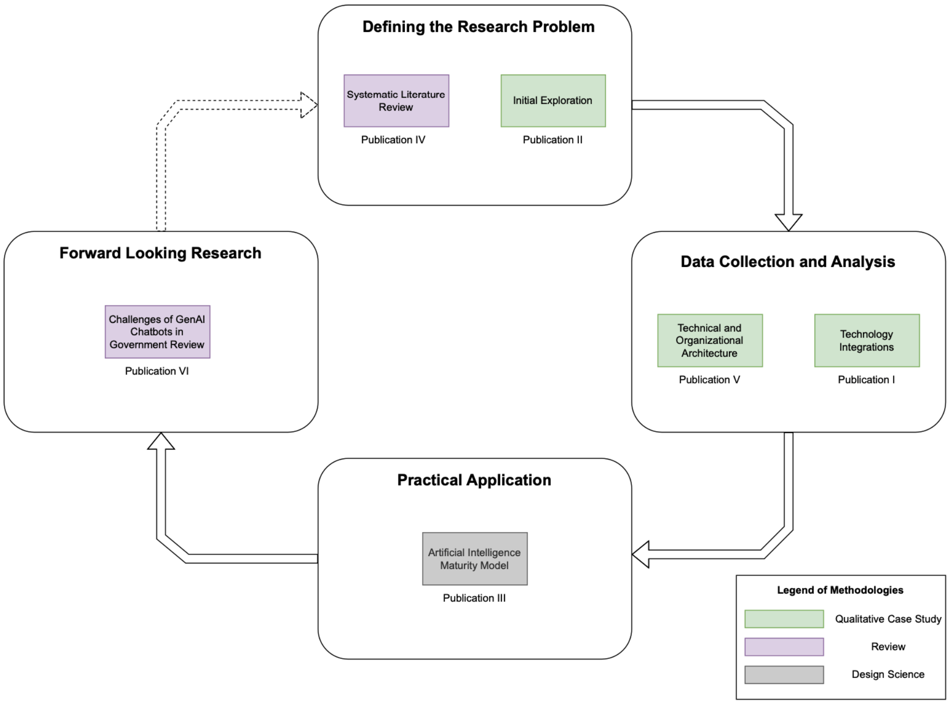

Figure 1: Research Design

The initial phase of the research is designed to understand the field as it stands at the beginning of the research. This includes publication IV and publication II. Publication IV is a systematic literature review attempting to understand the current state of digital transformation and the academic literature related to the use of AI to create digital state public services. Publication II is a qualitative study that used semi-structured interviews with stakeholders in the Estonian e-governance ecosystem. This included the Chief Data Officer (CDO) of Estonia and the Nordic Institute for Interoperability Studies (NIIS) Chief Technology Officer (CTO). This exploratory paper aimed to define the problem from a social, legal, and technical perspective.

The results of these initial studies provided feedback on the thesis's research questions and hypotheses and on where research efforts should be expended. Following this exploratory and research problem definition phase, two further studies were conducted to understand the phenomenon in a deeper way in its environment.

For the next phase of the research design, the emphasis was on qualitative case studies. This is because, according to [42], the qualitative case study is the best methodology to understand a complex phenomenon in its living environment. The interviews for these two papers were conducted over a year and a half. They included ten interviews

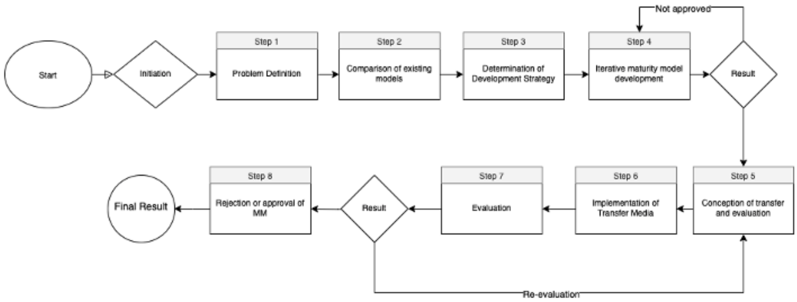

Figure 2: Design Science Research Design derived from [44, 45] [III]

with stakeholders from the Bürokratt program, including government decision-makers who had been present at the beginning of the program, government employees who participated in the program from various perspectives, and clients and consultants who were adopting the tool or analyzing it for research. Because the exploratory research had been conducted and the research questions refined by the point at which this research started, the interview guide reflected the increased understanding of the unit of analysis. The interviews were an average of an hour and asked questions related to the technical and organizational architecture, the technical integrations, and how organizations analyze readiness.

Publication I focuses on the technical integrations and processes through which a public entity can adopt AI-enabled public services, specifically from a technical and process lens. As stated above, this paper's methodology is a qualitative case study that triangulates the data collected from semi-structured interviews with secondary literature about AI strategies and technological integrations and academic literature about the types of market adoption in the diffusion of technological innovations. The coding in this case study is both deductive and inductive. The deductive codes are decided upon based on the influential literature in AI in the public sector [26], maturity categories elaborated upon in [43], and the digital transformation literature [14, 15, 22].

Publication V paper uses different questions in the interview schedule to derive the analysis and results. This subset of data that was used focuses primarily on the drivers of what specific drivers affected the technical and organizational architecture of the Estonian public entities conducting the design, development, and implementation of the Bürokratt program. Again, a qualitative thematic analysis was performed. However, this paper used an inductive strategy in addition to the deductive codes that the researchers produced from the coding process.

During the qualitative case studies phase of the research, the author attempted to understand the anecdotal real-world phenomenon of the procurement of pilots without understanding the organization's ability to make anything of it. A research question about the notes and codes was created to inform the factors that led to the maturity model. The methodology for III was based upon [44, 45]. The author followed the design science methodology for this paper. It began with the initial questions in the qualitative phase of the research and some coding in the cases for the other papers, which identified important points to build categories. The following [45] existing maturity models from [43, 34, 46, 47] were analyzed and repurposed with input from experts who suggested modifications along the way. The initial iterations used direct conversations with four experts who participated in the public sector and consultants in AI. During each iteration, the model was validated with experts through interviews and qualitative feedback. The final iteration was determined from the feedback in a quantitative validation questionnaire which presented the levels of maturity and the AI maturity model for the public sector. This Design Science portion of the research is not sole outcome of research, giving greater credence to the decision for this to be a mixed methods thesis rather than focusing purely on the outcome of publication III.

During the research there were twenty-five semi-structured interviews and workshops, with sixteen participants, two literature reviews, and two questionnaires sent to a total of twenty people with twelve responses.

4 Related Works

This section provides an overview of related works from multiple fields, primarily academic literature. It establishes the state of the art and the various perspectives and academic disciplines used in the publications to conceptualize and analyze the topic. The first subsection contains the theories which are used as foundations in the articles in the thesis. The second subsection contains a state-of-the-art. These related works have been gathered throughout the course of the thesis through the use primarily of the Scopus database, with supplemental use of Web of Science and Google Scholar to triangulate sources and understand changes in the field after the initial papers were written and submitted. The method of gathering these sources, with the exception of Publication IV, is a snowball sampling literature review with keywords relevant to the subject matter of each article.

4.1 Theoretical Overview

The primary focus of this thesis's investigation is Digital Transformation as it pertains to the adoption of AI in enabling public services. Several theoretical frameworks have been used throughout the thesis to ensure that the empirical evidence is grounded in a theoretical basis. Below is a brief overview of the theories, describing how each applies to digital transformation.

4.1.1 Technology Acceptance

Digital transformation is a topic that has to do with organizations adopting technologies that have some feedback on the organization [14, 15]. This can be because the technology is complex [16] or because it is important enough to the organization that the effect is significant enough to warrant a change in business processes.

The theory used in this thesis begins from what can be called a "micro" view, beginning with whether an individual person will accept or adopt a technology. This area of the academic literature is often called "Technology adoption" or "Technology acceptance" , derived from the names of the relevant theories. This area of scientific inquiry typically references the Theory of Planned Behavior (TPB) and its Theory of Reasoned Action (TRA) as one of the originating points. The theory of reasoned action, a social psychological construct put forth by Fishbein and Azjen, tries to explain how attitudes and behaviors produce human action [48]. The theory of planned behavior (TPB) continued the line of inquiry and extends the TRA to improve the prediction of whether someone would engage in a behavior and includes more determining factors than TRA [49].

The line of scientific inquiry then continued with Azjen's fellow researcher Davis and later Venkatesh to modify the TPB to attempt to predict whether a person would adopt or accept a technology. The resulting model was called The Technology Acceptance Model (TAM). TAM is a model that, similar to the TRA and TPB, uses factors like "perceived usefulness," "ease of use," and "external variables" to attempt to predict whether a person will adopt a technology [50]. Davis expanded upon many of the factors over time [51]. This eventually resulted in what is commonly referred to as TAM2 and TAM3, which include models that have extended factors [52, 53]. Eventually, the researchers tried to take a unified view that became the Unified Theory of Acceptance and Use of Technology [54].

Each of the theories spawned their own following and advocates and resulted in many papers analyzing the differences between them [55, 56] or using them for analysis in specific applications [57, 58]. Reviews of the multiple acceptance theories have also been written [59, 60, 61]. However, these theories are used the least in this thesis. They primarily inform the thesis as to potential key factors in user acceptance and adoption of technology at the individual level.

4.1.2 Change Management

Digital transformations are complicated in that they have to counteract much of the same resistance in the organization that is detailed in the change management literature on a larger scale that affects even the way the organization or institution is organized. The most referenced and discussed theories come from Lewin and Kotter [15, 62, 31, 32] and focus on the ways in which an organization needs to work together at multiple levels and have the people at multiple levels engaged in making the change. These models vary in the number of stages, with Lewin's original theory detailing three [32] and Kotter's detailing eight separate steps [62]. However, many of the included implied and explicitly detailed processes follow similar guidelines in different language with a slightly different focus [31]. Change management can occur at many levels, from changing a network switch in a campus's local area network to adopting software that completely shifts how the organization conducts business operations, like implementing a customer relationship management or enterprise resource planning system.

4.1.3 Organizational and Institutional Theories

Institutional and organizational theory can elucidate some critical factors that impact the ability to capture public value concerning new technologies. Economist Erik Brynjolfsson has studied this in terms of the overall economic value created in terms of economic gains provided by private firms during periods of rapid change that follow releases of general-purpose technologies [63]. The findings indicate that economic value creation does not occur immediately after the release of general-purpose technologies. Instead, there is an S-shaped curve with rapid increases in value creation after a time and then leveling off. The explanation that Brynjolfsson [64] has landed on is the concept that it takes time to combine with secondary and tertiary innovations that allow for firms to gain the ability to use the technology to an actual economic benefit. Although this discusses the concept from the view of economic analysis of various metrics related primarily to the private sector or economy as a whole, other innovations may be necessary to ensure that value can be created. This could have to do with the diffusion of the technological innovation [24] as well as with other innovations that create the ability for the technology to be applied. One example of this that Brynjolfsson points to is the idea that when electricity was created, there were many significant investments required to ensure that electricity reached homes and was generally available across the world [63, 33]. A recent example of this could be the creation of chatGPT, making one of the most significant Large Language Models (LLM) trained available through a chat interface that people outside the data science and artificial intelligence research professions could use. ChatGPT saw a large uptake in consumer use when it became available to everyone.

Somewhere between the abstractions of the organizational level and the macroeconomic level, institutional theory seeks to understand not only how radical change occurs at the organizational level but also at the field level and what drivers and barriers are present to aid in the shifting of the means of production associated with these radical changes. Unlike some of the more micro-focused economic theories and organizational theories, such as Mendelssohn's conception of the organization as an information processing set of business processes [65], which is determinate of their success or failure in the information technology market, the institutional theory takes into account the social considerations in sociotechnical systems [23]. The outcome of this is that it gives one possible explanation of the five to 10-year lag in value creation detailed in [63] and explains it rather than a symptom of there needing to be further peripheral technological innovations; it is instead seen as the gathering of institutional norms at a field level, which may explain some of the diffusion of this technology [22]. It is, of course, also possible that these theories explain processes that co-occur in the overall diffusion of technology discussed in [24].

This work, however, seeks to analyze the process through which an organization adopts artificial intelligence for a digital public service and what technical organizational architectural decisions are made. Institutional theory aids in explaining this at the organizational level because it acknowledges that there is an existing history of institutional and organizational norms [37] that govern not only the field but individual organizations and, over time, can be affected because of the mimetic phenomenon wherein organizations and institutions in the same field tend to view what others do that is successful and copy it. This also explains digital transformation at a more meta level than it is described in [15] and [14], which both view the phenomenon from the perspective of a single organization and its recursive feedback as it pertains to a digital transformation, whereas institutional theory refers to look not only at the social phenomena and norms that are detailed inside an organization but also across the field. Because of this, digital transformation is viewed from a multi-organizational perspective that shows the interactions with government stakeholders and other firms in the field attempting to institute a digital transformation rather than looking at digital transformation purely from the perspective of the internals in an organization.

The topic of digital transformation resonates because it is both a theoretical discussion of a practical and complex subject, including how organizations adopt technologies. Within digital transformation, organizations have to reckon implicitly or explicitly with concepts discussed in the sciences of organization and institutional theory, change management, as well as the individual technological and organizational challenges relevant to the specific technology they are adopting.

4.2 State of the Art

In addition to the theoretical basis used to analyze the topic included in this thesis, more application-related aspects of technologies and academic disciplines are detailed in the included publications and relevant to the research results. A brief overview of the state of the art is included below.

4.2.1 E-Governance and the Digital State

The topic of digital transformation in the public sector is closely tied to e-governance. E-Governance means "the use of Information and Communication Technology (ICT) in government in ways that lead to genuinely different structures or processes, a consequence of which may be the greater effectuation of or changes in norms and public values [66]." When a public sector organization is going to adopt technology to provide better services to their citizens and residents, there are changes that require addressing organizational and technological challenges. This means that a public sector organization attempting to adopt a form of e-governance will usually have to complete some digital transformation if the end state varies considerably compared with the beginning state.

E-governance as an academic discipline sits at the crossroads of the public administration and information sciences disciplines [67]. By its nature, this is an interdisciplinary branch of knowledge that focuses not only on the technology itself [68] but also on the ability of governments to adopt, implement, and manage these technologies [27]. In addition, research on citizen uptake of individual technologies explains the necessity of trust as a factor for citizen adoption [29, 69, 70]. Other research discusses specific methods that may have an impact on creating trust or confidence in e-governance systems [71]. This type of research also investigates frameworks that can be applied at various levels of government to aid in the successful adoption of e-governance technologies [27] or the building blocks of what can allow a government to move from e-governance and the digital state to a "postdigital" data-driven personalized government for citizens [72, 73]. Some research also indicates that the diffusion of e-governance technologies in terms of adoption by citizens follows a more linear distribution rather than the exponential curve that is typically seen among venture-backed hyper-scale [74] digital technologies [75].

As the body of knowledge has increased in e-governance, many research niches have been investigated, and the standard language used in the field has evolved [76]. Recently, rather than using the term e-government or e-governance, the standard terms are "digital public services" and "digital state." As stated above, some recent visions in Estonia are referred to as "postdigital" when referring to a new level of digital maturity in a society in which digital public services are the standard and the advanced technologies supporting them have become invisible [72]. Though the terms "digital state" and "digital public services" are used throughout this thesis, they have their roots in the e-governance field, the accompanying interdisciplinarity reflected above, and sociotechnical systems in general. In the spirit of [77] these interchangeable terms are not be defined in this thesis, as the definitions are not germane to a productive discussion on the use of AI to enable digital public services and the author could find no already complete work like [14] which clearly defines the differences between terms which are usually used in place of each other based on a large number of expert feedback.

4.2.2 AI in the Public Sector

Since 2019, artificial intelligence as an area of study within public administration and e-governance has drastically increased [35]. However, not all of the research on AI literature on AI in the public sector directly relates to the field of AI in public services.

The literature in the public administration concerns itself with significant questions, like governance of AI in general and how it will comply with legal and ethical guidelines in different domains [17]. The issues that bias can have in the public sector when bias is not accounted for and solved, or the scope of the system changed to shift the risk is well documented [20]. Bias can be related to many factors, and it is not always possible to correct it via weighting of more than one factor simultaneously [78].

A body of work covers the governance of AI and how it will affect the functioning of society in general [26, 79, 80]. This goes as far as attempting to predict the potential ramifications of AI in the public sector from an employment perspective, such as how governments respond to preempt large-scale AI-driven unemployment. Some researchers investigate how certain legal developments in Europe may affect the rest of the world, essentially arguing that similar to the general data protection law (GDPR), the "Brussels Effect" will likely mean that the EU AI Act will become the de facto standard or guidepost for governments around the world to regulate AI [81].

In addition, there is the challenge of ensuring that the developments of AI and their adoption in the public sector lead to additions in public value [82, 83]. This adds to previous research about co-creation and is an AI perspective on the creation of public value and the attempt to ensure that the adoption of AI benefits the populace.

Building on the idea of public value and the amelioration of the potential adverse effects of AI on the populace, another significant set of research papers is on the topic of creating “good” AI. This is not a scientific term. This is meant to give a meta-definition of many different approaches that have similarities. The authors are attempting to create AI in some way that is not a dangerous or malevolent force in human affairs. The broad issue at hand is the "alignment problem" discussed in the eponymous book [84], which seeks to find a way to create artificial intelligence that improves life for humans rather than having a detrimental effect. Acknowledging that these are all different in their perspective and consist of varying key points and levels of abstraction, this includes but is not limited to Trustworthy AI [85], Safe AI, Human Centered AI (HCAI)[86], Explainable AI (XAI)[87], and a Human-Centered Design (HCD)[88] approach to AI. These terms and their associated framework are instructive and necessary but outside the scope of the research covered in this thesis.

Estonia was one of the early movers in terms of national strategies [89]. The results of the working group and the development of a national strategy toward AI at an early point allowed for the creation of the vision papers which kicked off the Bürokratt program.

4.2.3 AI Use in Public Services

With the field being divided between heavily computer science-influenced applications of AI and the more social science or public administration literature concentrated on governance and bias, the author sought to understand the field from a public service perspective. Because of the inclination to separate the public services uses of AI from those that would impact the sphere of bias, the author sought to look for only articles related to the use of AI in e-government services. Again, the thesis aims to understand how AI could be used to create digital public services.

Part of the difficulty of a digital transformation related to the use of AI in public services speaks to challenges in many organizations with many technologies.

Public sector organizations face many challenges in competence and skills. These challenges range from the micro level, dealing with hiring and retaining competent personnel with technological competence, to the more macro digital transformation-related organizational competence. Even among some tech sector layoffs in the labor market, people with the technical capabilities to function in cross-functional teams developing, managing, maintaining, implementing, or adopting technology are in high demand. This is exemplified as it pertains to AI competence in the findings of [90] when they state that it is nearly impossible to find one person who fits all the suitable and necessary skills for a data scientist in the public sector. They suggest hiring teams of people who can meet all the requirements for public sector data science competence. The necessity for multiple disparate skill sets is brought into further relief in [12], where the results underline that human and social skills are essential in an environment governed by the law and complex human social dynamics. Organizational agility is a crucial factor in instantiating any change or transformation [91]. In the public sector, this can be called administrative agility [15, 92]. Organizational agility is tied to how organizations gain value from new technologies because the ability to create value from new technologies comes from the ability of organizations to be agile. This can be thought of as how an organization can adapt business processes and procedures due to feedback from within and without the organization. For example, in the case of Brynjolfsson et al. [33], the organization had to be able to adopt the new technology and integrate the technology in a co-pilot manner to help increase the productivity of customer service agents. It is entirely feasible that the technology could have been in a state where it was ready for adoption without actually being used in a manner that increased value. The simple existence of a technology does not mean that it will be used. This is why there is a body of literature that seeks to understand the diffusion of technology [24, 93] and the factors that determine and attempt to predict or understand the adoption of technology as mentioned above [29, 49, 51].

At the same time, when it comes to adopting technologies inside organizations, the organization needs to expend time and effort to be able to alter business processes and human resource strategy to not only change the way things are done in the organization to implement the technology better, but to hire individuals who have competences to bring that value to fruition. A consistent challenge for public sector organizations seeking to adopt artificial intelligence is the hiring and retention of people who are competent in the many skills necessary to do so. Even when a strategy of an organization is to procure most of the AI they will use, it is imperative to have enough internal competence in the organization to ensure that the promises of external providers are based on fact and completed in a way commensurate with the ethics, laws, and regulations of the organization implementing the AI [12].

4.2.4 Artificial Intelligence Organizational Readiness and Maturity

When it comes to the readiness of organizations to adopt artificial intelligence, the primary academically defensible method is through academic maturity models. This distinction must be made because there is a significant amount of grey literature on AI adoption in government [9]. However, much of this comes from the management consulting field. Readiness assessments and feasibility studies are helpful tools in the medical field. However, although these could be found in consultants' grey literature and various marketplaces for technological service organizations, the academic literature primarily had maturity models.

Maturity models have been used in the academic literature in the information systems and e-governance research to show the maturity of organizations to adopt a technology or meet specific criteria detailing how well the organization's systems and processes function [94, 95, 45].

Regarding the research questions posed in this thesis, the AI maturity model would be the artifact of choice. Artificial intelligence maturity models (AIMM) seek to place and understand organizations on a maturity matrix based on many factors. One consistent factor across many maturity models has to do with the organizational maturity for adopting AI and the competencies in the organization [43]. Many academic maturity models have been written in the information systems research field based on the design science in information systems guidelines in [44]. In [45] describes a system of repurposing maturity models to suit a new use. Although there are many AI-related maturity models in the literature, no holistic AIMM for public sector use has been based on the design science methodology detailed in [44, 45].

5 Results

The following section lists the results which are derived from a the publications which comprise this thesis. The results contained here are selections that aid in answering the research questions of the thesis. These answers are part of the contributions of the publications but not the entirety. Research questions are placed in the relevant section.

Table 2: Correlation of the research publications to the research questions

AI-Enablement of Digital Public Services						
RQ No	**I**	**II**	**III**	**IV**	**V**	**VI**
MRQ	X	X		X	X	
RQ 1	X			X	X	
RQ 2	X					
RQ 2.1	X					
RQ 2.2		X				
RQ 3			X			
RQ 3.1			X			
RQ 3.2		X	X	X		X

5.1 Leveraging Digital Transformation For AI-Enabled Services

- **MRQ**: How can public sector organizations leverage digital transformation to facilitate the adoption of AI-enabled services?

The papers chosen to comprise this thesis all have elements that relate to the main research question of this thesis. Publication I illustrates the technical challenges of developing, adopting, and implementing the first chatbot-based AI-enabled services building to virtual assistant-based services. Publication II explains some of the concerns related to the social, legal, and technical challenges governments face adopting AI-enabled services from an exploratory perspective. Publication IV is a systematic literature review analyzing the disciplines that comprise the academic literature on e-government services using AI. Publication V indicates the ways and choices for public entities to design their organizational and technical architectures. Publication III explains the ways the public administrations seeking to adopt AI can measure their maturity and gain an understanding of the road map to higher maturity and more effectiveness in AI endeavors. The main research question is answered in the following sections when considering the entirety of the results.

5.2 Potential for AI-Enablement in Public Services[I]

- **RQ.1**: What is the current state of digital transformation in public sector organizations, and how does it affect the potential for AI enablement in public services?

The awareness of the concepts of digital transformation cannot be generalized. Still, in the case of the Bürokratt initiative, it seems like the leaders who initiated the program and are currently developing it understand the reflexive feedback nature in digital transformations [I, II, III, V]. Estonia has an advantage in pursuing AI-enabled public services because they have an existing digital public service infrastructure [I]. The development, implementation, and adoption of an AI-enabled public service can certainly be done without this preexisting infrastructure. However, developing this from nothing takes more

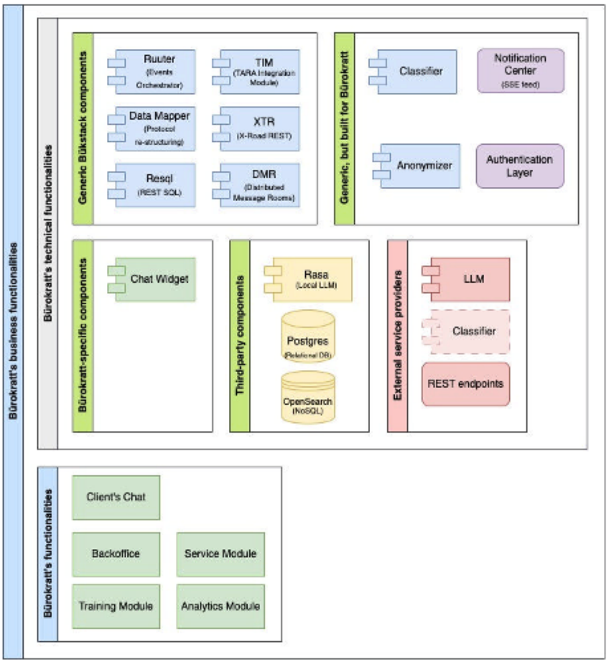

Figure 3: Bürokratt Technical integrations.

[I]

effort than integrating preexisting elements and adding additional parts. In the case of Bürokratt, the decision was made not only to use reusable building blocks of AI across the government but to integrate some already existing pieces that comprise the service architecture [I]. One of the most significant development pieces is the service module [I, V]. The concept of the service module is to make the 3000 e-services available in Estonia accessible through an AI-enabled channel with Bürokratt. This concept of prioritization is crucial in understanding the digital transformation of the Estonian government concerning AI-enabled services. Other countries that want to adopt AI-enabled digital services will have a different path to the result [V].

Another essential part of the Estonian case of AI-enablement adoption is the presence of leadership who understand the experimental nature of the programs [I]. From the proof of concept when the program was still called KrattAI to the current development, adoption, and implementation of the second version of Bürokratt, the experimental spirit of the strategy to functioning product and services process has been maintained. Suppose there is a problem that the current implementation does not do adequately. In that case, the team at the state information authority (RIA) has the bureaucratic agility and agile development and feedback mechanisms in place to requisition a new backlog item through their Github-based open procurement and development processes to address the challenge and solve the problem [I].

5.3 AI-Enablement Adoption for Existing Public Services[I]

- **RQ.2**: How would an existing public service adopt AI enablement?

The AI enablement for public services will vary greatly based on the existing services available in a geography [I]. The first thing is to have digital services available. For example, it would be challenging to take a service that is only available via paper forms and face-to-face interaction and make it available with AI-based tools. For the Estonian case, as mentioned above, there are over 3000 services available with X-road as the data exchange intermediary. X-road uses the Representational State Transfer Application Programming Interface (REST API) architecture with payload encryption over the public internet to exchange data between entities consisting of government organizations and private organizations [10]. This REST API architecture is one of the paradigms the Bürokratt program chose to implement in their development methods. The idea is that all of the Bürokratt functionalities are made possible through REST API calls. The first and most important thing is the business logic for the service. Then, it is put into a series of Yet Another Markup Language (YAML) files, which the Natural Language Processing (NLP) functions of Bürokratt can instantiate.

Additional requirements apply in any government implementation of services. Whenever an entity considers implementing AI enablement of public services, it is essential to ensure the privacy and safety of citizens and establish trust. Considering the security requirements of the AI-enablement platform [7] and the service provision element is necessary. This occurs in the Estonian ecosystem because the ministry offering services that adopt Bürokratt, and eventually the service module, signs an additional contract with the state authentication service (TARA). This allows residents and citizens to use the authentication infrastructure that includes multiple electronic identifications (eID) based authentication methods like the ID card, Smart-ID, and Mobile ID to authorize digital service transactions.

Legal requirements also play a role. When the service module is implemented, the ministries will sign up to allow citizens and residents to conduct services through the Bürokratt-enabled chat and voice channels. Additionally, the ministries will have to sign a contract with X-road. In Estonia, contracts are signed between service providers, service consumers, and other stakeholders. This contract fulfills a legal requirement to access the ability to use X-Road.

Interoperability and a method of authentication are also necessary for AI-enabled services if data is potentially exchanged between different service providers and consumers. The interoperability challenge has been solved by years of practice of having X-Road in use in Estonia, and another entity choosing to follow this path of development for AI-enabled services may find this an unexpected area that could be an issue.

5.4 Key Steps and Challenges[I]

- **RQ 2.1**: What are the key steps and challenges involved in this process?

The steps and processes in the Estonian journey from the KrattAI vision to Bürokratt's implementation did not go without challenges. Publication I primarily looks at this. Figure 4 shows the Estonian government's process in its strategically informed experimental path forward.

It is noticeable that there are multiple iterations of procurement throughout the process. This is because the top-down planning initiatives across many countries related to AI do not always correctly understand the capabilities of technical tools that may be integrated into the vision and strategy concepts. For example, when the KrattAI vision for the Next Generation Digital State was conceived, the thought was that the virtual assistants on phones were much more capable than, in fact, they were. Because of this, multiple iterations of looking at what is available on the market and how it would integrate into

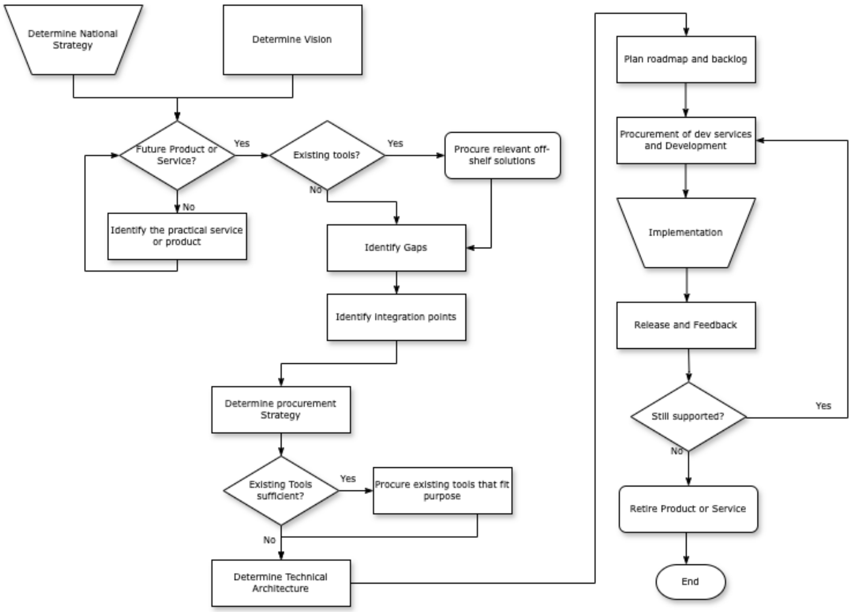

Figure 4: Planning process for technological development of strategic visions [I]

the architecture was necessary [I].

The previous section discussed the competitive advantage of the Estonian ecosystem in adopting AI-enablement because it has so many preexisting tools that can be integrated into an AI-enablement project. However, despite this, there were and continue to be many challenges that the RIA and the Ministry of Economic Affairs and Communication (MKM) teams have had to address throughout the process [I]. According to the program's architect, the ultimate goal is for Bürokratt to have a series of interchangeable elements, like protocol. In this way, if one piece does not function as well as expected, it can be replaced. In the soon-to-be-released version of Bürokratt, to help alleviate the challenge of correctly answering questions based on the knowledge base of a specific ministry, the team has chosen to implement an external LLM in addition to the Rasa intent recognition and GUI-based internal training tools [I]. The hope is that the NLP functionality will become more responsive to end-user requests. Another strategy will be used if this does not achieve the level the team or citizens want.

This is to say that even though there are many challenges going from a strategic vision to a deployed system, the ongoing developments and feedback continue to ensure that the system functions in a way that citizens and residents will want to use. The continuous improvement is never completed until the product becomes end-of-life and is no longer supported [I]. The diagram in Figure 4 represents the process as understood by interviews with decision-makers and personnel related to the Bürokratt program. In Publication I, this process flow is identified as a way for public administrations seeking to introduce an AI-enabled service to follow should they already have related compatible pieces of technology that they can integrate with the AI-enablement solution. The process flow implicitly contains the planning results of the stakeholders' challenges along the journey from a vision to a usable product or service. This process flow is meant to be a guideline as to history and cannot be assumed to be generalizable, although the experimental approach and paths determining whether to integrate existing tools or develop one's own are potentially useful to public administrations following the top-down technology push AI-enablement from strategy.

5.5 A technical and organizational architecture based on social, legal, and strategic considerations[V]

- **RQ 2.2**: What social, legal, and strategic considerations influence the technical and organizational architecture of public sector organizations implementing AI-enabled e-services?

The organizational architecture of a government is not a start-up. The government has a more machine-like structure [38]. The government organizational architecture is based on history and legal precedent that result from the social development of the region over time [37]. The history of a governmental organization cannot be removed from the history of the area completely. In Estonia, this results in a sort of decentralized dynamic inside of a centralized national government in the area of organizational architecture as well as the technical architecture in X-Road and Bürokratt. [V] largely deals with this question.

The goal of a large government organization cannot be purely to innovate or adopt AI-enabled services by the nature of it having already defined objectives that are part of its legal and historical position as a part of the government. At the same time, a government organization cannot have the sole goal of existing as an information processing unit or spin-off startup to bring to fruition projects that do not fit the primary functions of a business as proposed by[36]. Because the more flexible organizational architectures define private sector principles in the information economy but still want to capture the spark of innovation, the idea has to blend the old organizational paradigm with the best practices of the new [V].

The technical and organizational design results from the government's strategic goals. The MKM and RIA teams creating the Bürokratt program face the challenge of building the "new" while maintaining the services for which they are already responsible. To this end, the choice was made to create a special team responsible for Bürokratt within RIA. Additional teams within MKM handle certain functions, like the service module and the natural language processing (NLP) elements. However, the primary team members include the architect and product manager, who are organizationally inside of RIA. As a part of the organizational challenge to ensure that there is communication across various groups in the government who are working on AI-related topics, a project manager for AI was created as an overlay position who reports directly to the Chief Data Officer (CDO) of Estonia. This is not only to help ensure that there is no duplication of development effort but also to help ensure knowledge sharing across the government about ongoing efforts to adopt AI and AI-enabled services [V].

Similar to organizational architecture, some technical architectural decisions can also be connected to the historical achievements in the relevant geography. For example, Publication V cites how the architecture of Bürokratt takes heavily from X-Road. Not only does Bürokratt integrate with X-Road for the purposes of executing digital public service and using a REST API-based architecture, but the development process of having a "decentralized monolith" [10] is followed. The technical architecture is developed with the principles of X-Road in that it has some centralized components that provide services to distributed systems. The development is conducted through public-private partnerships with the responsibility for the objectives and key results (OKRs) [96] being primarily placed on one team.

This technical and organizational architecture is similar to X-Road in that it is also necessary to evangelize and get other ministries to adopt the technology. During the development of X-Road, they had personnel who specialized in the cross-government entity sales process [10]. Similarly, part of the OKRs of the Bürokratt team included getting other ministries to adopt the functionality [V]. The objective is to fulfill the strategic goals, and the organizations participating have to work within the legal and social parameters posed by the mode of operation of the government body responsible for executing the strategy.

5.6 Organizational Maturity to adopt AI-enabled digital services[III]

- **RQ.3**: How would one analyze the readiness of organizations to adopt AI enablement for e-services?

  3.1: What existing frameworks exist for readiness evaluations of AI projects, and how can they be adapted to the public sector e-service application?

During the thesis process, it became apparent that few or no public entities conduct a readiness assessment or maturity evaluation before embarking on the process of adopting AI-related projects. In the AI Watch project, it became apparent that in some cases, pilots failed due to a lack of prerequisites, namely data, that could have been understood at the beginning.

During the investigation phase of the research that became publication III, the author investigated existing frameworks and their adaptation to the e-service application. Unfortunately, the main frameworks to measure readiness and maturity for AI adoption in the public sector not only did not have a specific e-service component but also did not have any design science method-based framework for adopting AI in the public sector. Many readiness assessments and maturity models were from consulting organizations [III]. The problem is that they are not based on academic literature, even if they are based on theoretical or practical experience. Another challenge with these is the published white papers that publicize the consultancy's perspective on the topic; the information contained in these documents is meant as a lead generation tool for governments to acquire their services, which is a different aim than the academic literature. Because of this, the gray literature was only used to inform the author what areas of concern are potentially important in the case of the adoption maturity of government entities toward AI.

In addition, from interviews that provide the basis of the "transformation" of the existing maturity models to the new use of adoption of AI in the public sector, it became apparent that it is rare for anyone to consider the maturity of the organization before choosing to adopt an AI project. One interview explained that when it is done, it is usually completed by the data science group only and reflects their opinion of their own capabilities. The benefit of conducting a maturity evaluation is partially to get an assembly of cross-functional stakeholders to have a frank discussion about the organization's maturity. It would take leadership to achieve any significant change, and this aids in gaining leadership approval and engagement [8]. The AIMM, detailed in publication III, explains the large categories that public sector entities ought to take into account prior to implementing an AI project. The categories are broad enough to include AI projects, not only AI-enabled public services. Part of this is due to the theoretical nature of maturity models. Unlike many of the academic maturity models included in [43] and investigated in the theoretical part of publication III, this model attempts to more concretely define the levels of maturity than just the titles of the boxes included in the maturity model itself. The maturity model levels are derived from the Capability Maturity Model Integration Institute (CMMI) maturity levels, similar to those in the [34]. The maturity model and the levels were iterated upon four times. The majority of the respondents to the final questionnaire

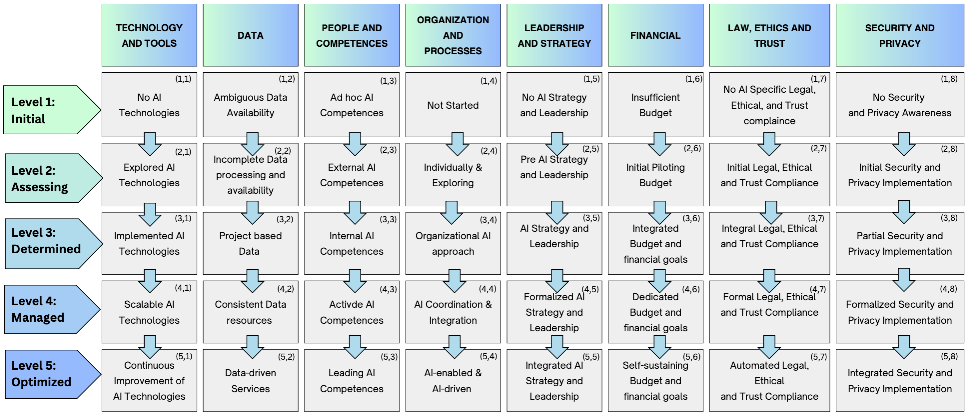

Figure 5: Artificial Intelligence Maturity Model for the Public Sector

[III]

on the model agreed that it fit the criteria for fulfilling its purpose as a maturity model for AI adoption in the public sector.

5.7 Ensuring ethical and responsible AI implementation for Digital public services[II, III, VI]

- **RQ 3.2**: How can public sector organizations ensure ethical and responsible AI implementation in e-services?

Ethical and responsible AI are prevalent topics in the public administration literature, e-governance, and information systems research. Bias comes in many different forms in the field of AI research. The decision was made early in the research to focus this thesis on enabling digital public services through AI. The current way this is done is through the use of chatbots or virtual assistants to open a new digital channel through which the citizen or resident can conduct digital public service transactions. This inherently limits the ways that bias can affect people because of the limiting of the scope. In this case, the primary potential challenge is non-native speaker bias. In Estonia, this has some ramifications. The Bürokratt program primarily uses Estonian. Interviews late in the thesis research indicate that English and Russian capabilities are further backlog items in the road map. As detailed in Publication V, the Estonian first development is because the Estonian language is currently underrepresented in the technologies provided by large tech providers. This includes the iPhone. One interview mentioned specifically that they use their iPhone set to German or English because there is no Estonian language capability. The government wants to provide the native language option for Estonians to access services from their own government and ensure a lasting digital footprint for the Estonian language. To this end, the government created various initiatives to place a large emphasis on labeling language data. At the same time, there is not much non-native bias research in languages other than English. The majority of the research questions in this thesis are focused on NLP enablement of the digital service processes. Because of the lack of decision support systems or other AI systems that would likely fall under the purview of high-risk systems in the EU AI Act, there is not much direct contribution in the research that comprises this thesis in the area of risk and bias.

However, [III] addresses the topic of risk and bias in a more open manner on the topic of applications of bias and risk. The reason for this is that the lack of an AIMM that is directly applicable to the public sector gave the author reason to create an AIMM for the public sector adoption of AI and not an AIMM for the adoption of digital public service enablement. To this end, the AIMM included in the publication and validated through the questionnaire-based feedback as well as interviews where the respondent was willing and available has two relevant categories that implicitly and explicitly discuss topics related to bias and ethics. The author and the resulting AIMM do not ascribe to a particular labeled version of an approach to the topic. The AIMM is written with the decomposition of these elements written into the categories of “Law, Ethics, and Trust” and “Privacy and Security.” In the Publication III AIMM levels, there are measures of the organizational requirements that would be made organization-wide and standardized as maturity increases.

Sometimes, technology-based decisions can increase the ability to be legally and ethically compliant and increase trust. As explained in Publication II, the early part of the KrattAI pilot, to adhere to the Estonian Personal Data Protection Act (PDPA), the government chose to make a network of chatbots rather than a single chatbot attached to a data lake [II]. This makes developing and training multiple chatbots much more complex due to the multiple knowledge bases and training data sets that must be developed and maintained for each relevant chatbot associated with a ministry database [II]. This does not even include the mechanism to pass the user to the relevant chatbot and move them during the middle of a conversation or a change of topic [I]. However, this decision was made explicitly to ensure an ethical implementation. Publication I explains that in the recent version of Bürokratt, which has not yet gone into production due to ongoing security testing, the team made the determination that it is necessary to add some sort of LLM-based functionality at least to potentially increase the responsiveness and capability of the chatbot. In the technical functioning, the team made the decision to take away all personally identifiable information and metadata before sending a query to the external LLM due to privacy concerns [I]. This solves one of the challenges of implementing GenAI chatbots in government covered in [VI]. This is done to ensure that citizen privacy and data protection are not compromised in the case of a data leak by the external provider. This is similar to the quantitative risk analysis conducted in [7], but the paper gave an example of an external virtual assistant provider. In the actual Bürokratt implementation for version 2, the external system in question is an external LLM. Quantitative risk analysis programs may also help ameliorate security challenges by allowing organizations to focus on their most costly or likely potential attack avenues. Another example of a technological decision meant to enhance trust is that the original pilot version of Bürokratt did not have X-Road integration. To ensure legal compliance, any sort of tool in the Estonian ecosystem that will eventually be able to instantiate government services transactions must use X-Road [I]. The original KrattAI pilot did not use this because of a challenge related to the synchronous message traffic required by X-Road detailed in the Next Generation Digital State Architecture vision paper [97]. During the pilot’s time, the necessary decision was made to use Apache Kafka as detailed in Publication II [II]. By the time the production version of Bürokratt was created, the creation and integration of distributed messaging rooms as detailed in Publication I was able to solve this problem [I].

Sometimes, when using AI to develop or contribute to digital public services, decisions need to be made regarding how to optimize a service. In Fourth Publication, there is an example of a situation detailed in the literature [98] [IV]stating how the government could have made a service more profitable through optimization. However, this change would have affected a vulnerable population, so the decision was made not to change the service. This type of consideration is in line with the ethical application of these technologies

The examples mentioned do not aim to predict every way that AI will develop technologically. Instead, they try to offer a set of practical actions for entities to take while developing, implementing, and continuously improving their AI systems. These actions should guide the entities in ensuring the government's role in protecting those whose lives it is responsible for improving.

6 Limitations and Future Work

The rapid advancement of technology and the internal and external pressures for digital transformation were key elements explored in this thesis. It is worth noting that during the research period, ChatGPT became generally available (GA). The paid version featured access to Generative Pre-training Transformer (GPT) 4, while the free version utilized GPT 3.5 as its underlying model. The Bürokratt program currently uses a system called Rasa to do the intent analysis paired with a Graphical User Interface (GUI) based tool to limit the competencies necessary for ministries to adopt the technology. One challenge related to the intersection of the fast-moving technology set and the changing demands, both internal and external to the organization, is that the citizenry's expectations are likely to have shifted. There is no data on this authoritatively. However, there is an indication that as technology progresses the amount of time people will give websites less of a chance to load prior to abandoning the site and, most times, not coming back. Similarly, it could be hypothesized that the changing expectations toward the responsiveness and perceived intelligence of virtual agents are likely to cause pressure for further digital transformations oriented at any chatbot or virtual assistant to ensure that their perceived level of usefulness is commensurate with that of the commercially available products. This would be a useful potential area of future work.

The GA of ChatGPT and, more importantly, the rapid adoption of the tool among the worldwide user base made this work more relevant to the discussion because it put the topic at the forefront of public administration's mind. However, at the same time, the primary use case is a chatbot that was planned and developed prior to this phenomenon and is attempting to add to the capabilities with LLM integration rather than completely redesigning and refactoring the tool. This goes hand in hand with the architectural conception of Bürokratt by its architect, which aims to become a set of protocols that people can use. This thesis deals primarily with traditionally developed chatbots in the majority of the included publications, with the exception of [VI]. This is a limitation because of the seeming ubiquity of GenAI and LLM systems.

The use of a single case is a difficult choice. This is primarily because it was the only case that had the expressed inclination of not just creating a chatbot through which citizens and residents can get answers to queries but ultimately using this to be able to instantiate government services.

The keywords used in the systematic literature review [IV] represented the author's best knowledge at the time. However, with hindsight, it would potentially yield more detailed results with the addition of keywords that acknowledge the shift in terminology from "e-government" and "e-governance" to the newer "digital state" and "digital public services." The literature review used the term "public service," but more articles may have appeared. The proliferation of articles about AI in the years since the SLR was conducted would also garner different results if conducted now with the same exact keywords. However, since the primary goal of conducting the SLR was to understand how the topic was investigated at the time, the paper is sufficient for the task.

The process of writing this thesis also demonstrated that a paper purely investigating the differences in the definitions of the keywords in the previous paragraph might be useful for someone.

The AIMM itself is not without its limitations. One of the respondents, who is from a very high position in government and an international digital government consultancy, thought that the best form would be a readiness assessment that had simple yes and no boxes to be filled in by the relevant authorities. The author agrees. Initially, the goal was to ensure a relatively lightweight assessment tool to gauge readiness among the public entities seeking to adopt AI. However, when this intention was met with academic literature, it became clear that the only academically justifiable way to create a tool to help this process was through the design science methodology in the form of a maturity model. This maturity model is quite complex. A work followed Publication III's AIMM that validated a form of self-assessment for a public authority received feedback as such [99]. Ultimately, due to the limited response to the validation mechanisms the author considers this a matter for consultancies to work with public authorities. Once engaged the public administration pays them money and will then have "skin in the game" [100] and increase the likelihood that the group participating would take the exercise seriously and engage with the subject matter.

One significant limitation to studying an ongoing digital transformation in a program like Bürokratt is that the digital transformation process is not complete. Many parts that will be important to the overall success or failure of the program are still in motion and have not been released. For example, it is unclear if the GUI-based training system meant to decrease the necessary technical competencies of adopting ministries by replacing two full-time data scientists with customer service representatives will be successful. It will be a challenge to take a traditionally developed chatbot and have it implemented in multiple ministries and have the agents able to answer difficult questions necessary to create the effectiveness and efficiency gains public administrations expect.

Thus, there is much left to be learned and studied in the ongoing Bürokratt development implementation, including but not limited to how they are able to handle the training in ministries to create chatbots that help citizens and how the expectations and awareness of AI have shifted among the Estonian populace with multiple LLMs reaching general availability.

Finally, it's important to note the underlying normative judgment in this work, which is a bias. It asserts that a government service utilizing AI to enhance the efficiency and effectiveness of providing services to citizens and residents is beneficial. While in the past this might have been obvious, in today's era, it seems necessary to explicitly state this viewpoint as a normative bias.

7 Conclusion

The topic of this thesis is somewhat difficult to organize coherently because it involves a phenomenon that is not only a newer technology with fast-moving developments but also has technical, social, legal, and practical implications. However, even with this complexity, several themes arose in the course of the research.

Because of the complexity of the digital transformation process when it includes AI adoption, it is necessary to include the clients and, if possible, the end users as early in the process as possible, given the constraint that many of the implementations of AI in public services are of a technology push market dynamic by nature. This is why it is important to have some assurance that the populace of the country in question will be able to use the technology and choose to do so. The Estonian government attempted to do this by commissioning a consultancy to understand the willingness of Estonian citizens to use services that were enabled with AI [101]. The feedback mechanisms that are in the process diagram from publication I tied to the feedback mechanism in [14] in that it shows from a process

The results also underscore the importance of competencies within the organizational human resources. This is per the discussions by [12]. The AIMM reflects this importance by giving people and processes an entire category, which includes a discussion of the maturity of human resources strategy.

In addition, comparing a traditionally developed chatbot like Bürokratt to an LLM-based chat function, which would nominally be considered the next iteration of NLP technology, provides additional challenges to tailoring the technology to citizens' and residents' user experience desires.

Digital transformations are complex and have a history of a lot of academic literature that reaches back from attempts to understand the modern definition in public administration literature [14] antecedents and dynamics of digital transformations [15] to the diffusion of recent general purpose technologies from the perspective of macroeconomic analysis, [64] all the way back to organizational and institutional theory on digital transformation and radical change [22]. Some researchers [102, 16] believe that any digital transformation that involves AI has added complexity due to the technology. This thesis attempts to understand the planning, development, and implementation of an AI-enabled service in a specific case and extract potentially generalizable knowledge to relevant situations. In addition, to learn to cohere to scientifically and practically relevant maturity models that can potentially help guide public administrations seeking to adopt AI.

List of Figures

1 Research Design 17
2 Design Science Research Design derived from [44, 45] [III] 18
3 Bürokratt Technical integrations. 27
4 Planning process for technological development of strategic visions [I] 29
5 Artificial Intelligence Maturity Model for the Public Sector 32

List of Tables

1 Correlation of the research publications to the research questions 15
2 Correlation of the research publications to the research questions 26

References

[1] R. Dreyling, , T. Tammet, and I. Pappel, "Technology push in ai-enabled services: How to master technology integration in case of bürokratt," *SN Computer Science*, vol. Future Data Science Engineering, no. 5, p. 738, 2024.

[2] R. Dreyling, E. B. Jackson, T. Tammet, A. Labanava, and I. Pappel, "Social, legal, and technical considerations for machine learning and artificial intelligence systems in government.," in *ICEIS (1)*, pp. 701–708, 2021.

[3] R. Dreyling, J. Lemmik, T. Tammet, and I. Pappel, "An artificial intelligence maturity model for the public sector: Design science approach," *TalTech Journal of Social Sciences*, vol. 14, no. 2, p. 16, Forthcoming.

[4] R. M. Dreyling, T. Tammet, and I. Pappel, "Artificial intelligence use in e-government services: A systematic interdisciplinary literature review," in *International Conference on Future Data and Security Engineering*, pp. 547–559, Springer, 2022.

[5] R. Dreyling, K. McBride, T. Tammet, and I. Pappel, "Navigating the AI maze: Lessons from Estonia's Bürokratt on public sector AI digital transformation," *SSRN*, 2024.

[6] R. Dreyling, T. Koppel, T. Tammet, and I. Pappel, "Challenges of genai chatbots in public services: An integrative review," *SSRN*, 2024.

[7] R. Dreyling, E. Jackson, and I. Pappel, "Cyber security risk analysis for a virtual assistant g2c digital service using fair model," in *2021 Eighth International Conference on eDemocracy and eGovernment (ICEDEG)*, pp. 33–40, 2021.

[8] R. Dreyling, R. Erlenheim, T. Tammet, and I. Pappel, "Ai readiness assessment for data-driven public service projects: Change management and human elements of procurement," *Human Factors, Business Management and Society*, vol. 97, no. 97, 2023.

[9] R. M. Dreyling III, T. Tammet, and I. Pappel, "Digital transformation insights from an ai solution in search of a problem," in *International Conference on Future Data and Security Engineering*, pp. 341–351, Springer, 2023.

[10] E. Blake Jackson, R. Dreyling, and I. Pappel, "A historical analysis on interoperability in estonian data exchange architecture: Perspectives from the past and for the future," in *Proceedings of ICEGOV'21 – the 14th International Conference on Theory and Practice of Electronic Governance*, pp. 111–116, ACM, 2021.

[11] E. B. Jackson, R. Dreyling, and I. Pappel, "Challenges and implications of the who's digital cross-border covid-19 vaccine passport recognition pilot," in *2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG)*, pp. 88–94, IEEE, 2021.

[12] A. Labanava, R. M. Dreyling III, M. Mortati, I. Liiv, and I. Pappel, "Capacity building in government: Towards developing a standard for a functional specialist in ai for public services," in *International Conference on Future Data and Security Engineering*, pp. 503–516, Springer, 2022.

[13] A. Labanava, R. M. Dreyling, and A. Norta, "Potential of smart contracts in the pharmaceutical supply chain of belarus," in *2022 IEEE 1st Global Emerging Technology Blockchain Forum: Blockchain & Beyond (iGETblockchain)*, pp. 1–6, IEEE, 2022.

[14] I. Mergel, N. Edelmann, and N. Haug, "Defining digital transformation: Results from expert interviews," *Government Information Quarterly*, vol. 36, no. 4, p. 101385, 2019.

[15] G. Vial, "Understanding digital transformation: A review and a research agenda," *The Journal of Strategic Information Systems*, vol. 28, no. 2, pp. 118–144, 2019.

[16] J. Jöhnk, M. Weißert, and K. Wyrtki, "Ready or not, AI comes – an interview study of organizational AI readiness factors," *Business & Information Systems Engineering*, vol. 63, no. 1, pp. 5–20, 2021.

[17] G. Bell, J. Burgess, J. Thomas, and S. Shadiq, "Rapid response information report: Generative ai-language models (llms) and multimodal foundation models (mfms)," tech. rep., Australian Council of Learned Academies (ACOLA), 2023.

[18] J. Bareis and C. Katzenbach, "Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics," *Science, Technology, & Human Values*, vol. 47, no. 5, pp. 855–881, 2022.

[19] J. Randall and J. Alter, "Factors That Contributed to NLARS Problems: Special Review February 14, 2019," tech. rep., State of Minnesota, Office of the Legislative Auditor, 02 2019.

[20] S. Bekker, "Fundamental rights in digital welfare states: The case of syri in the netherlands," *Netherlands Yearbook of International Law 2019: Yearbooks in International Law: History, Function and Future*, pp. 289–307, 2021.

[21] E. Brynjolfsson and L. M. Hitt, "Beyond computation: Information technology, organizational transformation and business performance," *Journal of Economic perspectives*, vol. 14, no. 4, pp. 23–48, 2000.

[22] B. Hinings, T. Gegenhuber, and R. Greenwood, "Digital innovation and transformation: An institutional perspective," *Information and Organization*, vol. 28, no. 1, pp. 52–61, 2018.

[23] J. Koppenjan and J. Groenewegen, "Institutional design for complex technological systems," *International Journal of Technology, Policy and Management*, vol. 5, no. 3, pp. 240–257, 2005.

[24] E. M. Rogers, A. Singhal, and M. M. Quinlan, "Diffusion of innovations," in *An integrated approach to communication theory and research*, pp. 432–448, Routledge, 2014.

[25] A. Androutsopoulou, N. Karacapilidis, E. Loukis, and Y. Charalabidis, "Transforming the communication between citizens and government through ai-guided chatbots," *Government information quarterly*, vol. 36, no. 2, pp. 358–367, 2019.

[26] B. W. Wirtz, J. C. Weyerer, and I. Kehl, "Governance of artificial intelligence: A risk and guideline-based integrative framework," *Government Information Quarterly*, vol. 39, no. 4, p. 101685, 2022.

[27] I. Pappel, V. Tsap, and D. Draheim, "The e-locgov model for introducing e-governance into local governments: an estonian case study," *IEEE transactions on emerging topics in computing*, vol. 9, no. 2, pp. 597–611, 2019.

[28] M. Kupi and K. McBride, "Agile development for digital government services: Challenges and success factors," in *Proceedings of ePart'2021 – the 13th IFIP WG 8.5 International Conference on Electronic Participation* (N. Edelmann, C. Csáki, S. Hofmann, T. J. Lampoltshammer, L. Alcaide Muñoz, P. Parycek, G. Schwabe, and E. Tambouris, eds.), vol. 12849 of *Lecture Notes in Computer Science*, (Cham), pp. 139–150, Springer International Publishing, 2021.

[29] L. Carter and F. Bélanger, "The utilization of e-government services: citizen trust, innovation and acceptance factors," *Information Systems Journal*, vol. 15, no. 1, pp. 5–25, 2005.

[30] S. E. Colesca, "Understanding trust in e-government," *Engineering Economics*, vol. 63, no. 3, 2009.

[31] B. H. Sarayreh, H. Khudair, and E. A. Barakat, "Comparative study: The kurt lewin of change management," *International Journal of Computer and Information Technology*, vol. 2, no. 4, pp. 626–629, 2013.

[32] B. Burnes, "The origins of lewin's three-step model of change," *The Journal of Applied Behavioral Science*, vol. 56, no. 1, pp. 32–59, 2020.

[33] E. Brynjolfsson, D. Li, and L. R. Raymond, "Generative ai at work," tech. rep., National Bureau of Economic Research, 2023.

[34] P. Fukas, J. Rebstadt, F. Remark, and O. Thomas, "Developing an artificial intelligence maturity model for auditing.," in *ECIS*, 2021.

[35] C. van Noordt and G. Misuraca, "Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union," *Government Information Quarterly*, vol. 39, no. 3, p. 101714, 2022.

[36] C. M. Christensen, *The innovator's dilemma: when new technologies cause great firms to fail*. Harvard Business Review Press, 2013.

[37] D. Draheim, R. Krimmer, and T. Tammet, "On state-level architecture of digital government ecosystems: From ict-driven to data-centric," in *Transactions on Large-Scale Data-and Knowledge-Centered Systems XLVIII: Special Issue In Memory of Univ. Prof. Dr. Roland Wagner*, pp. 165–195, Springer, 2021.

[38] H. Mintzberg, *Mintzberg on management: Inside our strange world of organizations*. Free Press, 1989.

[39] M. of Economic Affairs and C. (MKM), "Estonia's national artificial intelligence strategy 2019-2021," 2019.

[40] S. Sikkut, O. Velsberg, and V. K., *#KrattAI: the Next Stage of Digital Services in #eEstonia*. Republic of Estonia GCIO Office, 2020.

[41] J. W. Creswell and C. N. Poth, *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications, 2016.

[42] R. K. Yin, *Case Study Research and Applications: Design and Methods*. SAGE, 2017. 6th edition.

[43] R. Sadiq, N. Safie, A. Abd Rahman, and S. Goudarzi, "Artificial intelligence maturity model: a systematic literature review," *PeerJ Computer Science*, vol. 7, no. e661, 2021.

[44] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS quarterly*, pp. 75–105, 2004.

[45] J. Becker, R. Knackstedt, and J. Pöppelbuß, "Developing maturity models for it management: A procedure model and its application," *Business & Information Systems Engineering*, vol. 1, pp. 213–222, 2009.

[46] P. Fukas, J. Rebstadt, L. Menzel, and O. Thomas, "Towards explainable artificial intelligence in financial fraud detection: Using shapley additive explanations to explore feature importance," in *International Conference on Advanced Information Systems Engineering*, pp. 109–126, Springer, 2022.

[47] S. D. Das, P. K. Bala, and A. N. Mishra, "Towards defining a trustworthy artificial intelligence system development maturity model," *Journal of Computer Information Systems*, pp. 1–22, 2023.

[48] M. Fishbein and I. Ajzen, "The theory of reasoned action as applied to moral behaviour: A confirmatory analysis," 1975.

[49] I. Ajzen, "The theory of planned behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, 1991.

[50] F. D. Davis, R. Bagozzi, and P. Warshaw, "Technology acceptance model," *J Manag Sci*, vol. 35, no. 8, pp. 982–1003, 1989.

[51] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, pp. 319–340, 1989.

[52] V. Venkatesh and F. Davis, "Extension of the technology acceptance model (tam2)," 2000.

[53] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decision sciences*, vol. 39, no. 2, pp. 273–315, 2008.

[54] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS quarterly*, pp. 425–478, 2003.

[55] P. C. Lai, "The literature review of technology adoption models and theories for the novelty technology," *JISTEM-Journal of Information Systems and Technology Management*, vol. 14, no. 1, pp. 21–38, 2017.

[56] R. J. Vallerand, P. Deshaies, J.-P. Cuerrier, L. G. Pelletier, and C. Mongeau, "Ajzen and fishbein's theory of reasoned action as applied to moral behavior: A confirmatory analysis.," *Journal of personality and social psychology*, vol. 62, no. 1, p. 98, 1992.

[57] D. Ozag and B. Duguma, "The relationship between cognitive processes and perceived usefulness: An extension of tam2," in *Proceedings of 23rd Annual Organizational Systems Research Association Conference. Pittsburgh: Pennsylvania*, Citeseer, 2004.

[58] L. Carter and F. Bélanger, "The utilization of e-government services: citizen trust, innovation and acceptance factors," *Information systems journal*, vol. 15, no. 1, pp. 5–25, 2005.

[59] S. H. Alshammari and M. S. Rosli, "A review of technology acceptance models and theories," *Innovative Teaching and Learning Journal (ITLJ)*, vol. 4, no. 2, pp. 12–22, 2020.

[60] F. J. Rondan-Cataluña, J. Arenas-Gaitán, and P. E. Ramírez-Correa, "A comparison of the different versions of popular technology acceptance models: A non-linear perspective," *Kybernetes*, vol. 44, no. 5, pp. 788–805, 2015.

[61] W. R. King and J. He, "A meta-analysis of the technology acceptance model," *Information & management*, vol. 43, no. 6, pp. 740–755, 2006.

[62] J. P. Kotter, "Why transformation efforts fail," *Harvard business review*, 1996.

[63] E. Brynjolfsson, D. Rock, and C. Syverson, "Artificial intelligence and the modern productivity paradox," *The economics of artificial intelligence: An agenda*, vol. 23, pp. 23–57, 2019.

[64] E. Brynjolfsson and A. McAfee, *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company, 2014.

[65] H. Mendelson, "Organizational architecture and success in the information technology industry," *Management Science*, vol. 46, no. 4, pp. 513–529, 2000.

[66] F. Bannister and R. Connolly, "Defining e-governance," *E-Service Journal: A Journal of Electronic Services in the Public and Private Sectors*, vol. 8, no. 2, pp. 3–25, 2012.

[67] K. McBride, "Open government data co-created public services," 2020.

[68] J. P. Gibson, R. Krimmer, V. Teague, and J. Pomares, "A review of e-voting: the past, present and future," *Annals of Telecommunications*, vol. 71, pp. 279–286, 2016.

[69] S. E. Colesca, "Increasing e-trust: A solution to minimize risk in e-government adoption.," *Journal of applied quantitative methods*, vol. 4, no. 1, 2009.

[70] M. Weck, T. Gelashvili, I. Pappel, and F. Ferreira, "Supporting collaboration and knowledge sharing in building sles for ageing well: Using cognitive mapping in kms design," *Knowledge Management Research & Practice*, vol. 20, no. 6, pp. 865–877, 2022.

[71] M. Solvak, "Does vote verification work: usage and impact of confidence building technology in internet voting," in *Electronic Voting: 5th International Joint Conference, E-Vote-ID 2020, Bregenz, Austria, October 6–9, 2020, Proceedings 5*, pp. 213–228, Springer, 2020.

[72] M. Solvak and A. Lauringson, "A case study of the public sector digital ecosystem in estonia," *Computer*, vol. 57, no. 5, pp. 44–49, 2024.

[73] M. Maksimova, M. Solvak, and R. Krimmer, "Data-driven personalized e-government services: Literature review and case study," in *International Conference on Electronic Participation*, pp. 151–165, Springer, 2021.

[74] A. Azhar, *Exponential: Order and Chaos in an Age of Accelerating Technology*. Random House, 2021.

[75] M. Solvak, T. Unt, D. Rozgonjuk, A. Võrk, M. Veskimäe, and K. Vassil, "E-governance diffusion: Population level e-service adoption rates and usage patterns," *Telematics and Informatics*, vol. 36, pp. 39–54, 2019.

[76] S. Marche and J. D. McNiven, "E-government and e-governance: the future isn't what it used to be," *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, vol. 20, no. 1, pp. 74–86, 2003.

[77] K. Popper, "The myth of the framework," in *Rational changes in science: Essays on scientific reasoning*, pp. 35–62, Springer, 1976.

[78] D. F. Engstrom and D. E. Ho, "Algorithmic accountability in the administrative state," *Yale Journal on Regulation*, vol. 37, p. 800, 2020.

[79] B. W. Wirtz, P. F. Langer, and C. Fenner, "Artificial intelligence in the public sector-a research agenda," *International Journal of Public Administration*, vol. 44, no. 13, pp. 1103–1128, 2021.

[80] Z. Engin and P. Treleaven, "Algorithmic government: Automating public services and supporting civil servants in using data science technologies," *The Computer Journal*, vol. 62, no. 3, pp. 448–460, 2019.

[81] C. Siegmann and M. Anderljung, "The brussels effect and artificial intelligence: How eu regulation will impact the global ai market," *arXiv preprint arXiv:2208.12645*, 2022.

[82] C. van Noordt, G. Misuraca, and I. Mergel, "Analysis of driving public values of ai initiatives in government in europe," in *Research Handbook on Public Management and Artificial Intelligence*, pp. 226–244, Edward Elgar Publishing, 2024.

[83] C. Wilson, "Public engagement and ai: A values analysis of national strategies," *Government Information Quarterly*, vol. 39, no. 1, p. 101652, 2022.

[84] B. Christian, *The alignment problem: How can machines learn human values?* Atlantic Books, 2021.

[85] B. Li, P. Qi, B. Liu, S. Di, J. Liu, J. Pei, J. Yi, and B. Zhou, "Trustworthy ai: From principles to practices," *ACM Computing Surveys*, vol. 55, no. 9, pp. 1–46, 2023.

[86] B. Shneiderman, *Human-centered AI*. Oxford University Press, 2022.

[87] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, "Xai—explainable artificial intelligence," *Science robotics*, vol. 4, no. 37, p. eaay7120, 2019.

[88] J. Auernhammer, "Human-centered ai: The role of human-centered design research in the development of ai," 2020.

[89] A. Watch, "National strategies on artificial intelligence," 2022.

[90] S. Baškarada and A. Koronios, "Unicorn data scientist: the rarest of breeds," *Program*, vol. 51, no. 1, pp. 65–74, 2017.

[91] Y. Gong, J. Yang, and X. Shi, “Towards a comprehensive understanding of digital transformation in government: Analysis of flexibility and enterprise architecture,” *Government Information Quarterly*, vol. 37, no. 3, p. 101487, 2020.

[92] B. W. Wirtz, J. C. Weyerer, and C. Geyer, “Artificial intelligence and the public sector—applications and challenges,” *International Journal of Public Administration*, vol. 42, no. 7, pp. 596–615, 2019.

[93] G. A. Moore and R. McKenna, *Crossing the chasm*. Capstone Oxford, 1999.

[94] S. Khanra and R. P. Joseph, “E-governance maturity models: a meta-ethnographic study,” *The International Technology Management Review*, vol. 8, no. 1, pp. 1–9, 2019.

[95] I. Pappel, T. Gelashvili, and I. Pappel, “Maturity model for automatization of service provision and decision-making processes in municipalities,” in *Proceedings of Sixth International Congress on Information and Communication Technology: ICICT 2021, London, Volume 3*, pp. 399–409, Springer, 2022.

[96] P. R. Niven and B. Lamorte, *Objectives and key results: Driving focus, alignment, and engagement with OKRs*. John Wiley & Sons, 2016.

[97] K. Vaher, *Next Generation Digital Government Architecture*. Republic of Estonia GCIO Office, 2020.

[98] S. Hong, Y. Kim, and J. Park, “Big data and smat city planning: The case of owl bus in seoul,” in *2018 IEEE International Conference on Big Data (Big Data)*, pp. 4492–4500, IEEE, 2018.

[99] J. Lemmik, *Operationalising an Artificial Intelligence Maturity Model in the Public Sector*. PhD thesis, Tallinn University of Technology, 2024.

[100] N. N. Taleb, *Skin in the game: Hidden asymmetries in daily life*. Random House, 2018.

[101] Civita, “Awareness and opinions of Estonian residents about artificial intelligence,” tech. rep., Ministry of Economic Affiars and Communication, 10 2023.

[102] J. Holmström, “From AI to digital transformation: The AI readiness framework,” *Business Horizons*, vol. 65, no. 3, pp. 329–339, 2022.

Acknowledgements

From the bottom of my heart, I would like to thank my supervisor, Ingrid Pappel, for her consistent support over the past four years and all the challenges that have come as a part of the Ph.D. process and simultaneously raising two children. I would also like to thank my co-supervisor, Tanel Tammet, for being there to meet and discuss to help my work be more academically and technically cogent. Thank you to my colleagues who were in the boat for the journey.

This work has been partially conducted in the project "ICT programme" which was supported by the European Union through the European Social Fund.

Although it is not often acknowledged in American culture, whenever a person has any successes, there is often a village responsible. So now I will thank some people from my personal life who I believe are responsible for me being in a position to complete this study program and thesis.

Alena, my wife, and partner, it is a pleasure and a joy to raise two children with you and build our empire. "The days are long and the years are short," it may be a platitude by now, but I feel it in my heart every day. From sick kids and weeks in the hospital to four years of mostly seven-day work weeks, this road hasn't been easy or smooth. But the best is yet to come. There is hard work and the accompanying reward ahead. No good work goes unpunished, and we're only getting started.

To my two children, you were not here before I started this, but I am sure glad you came along. You give me a focus and a will to positively impact the world around me that can't be overestimated. To my family: I've changed a lot on this journey, mostly because I want to be alive, share the time with you as long as I can, and set a positive example while doing so. I love you all.

Mom and Dad, thanks for getting me out of the hood and instilling good Midwestern values, common decency, and determination in me. Mom, thanks for moving into that trailer in a good school district. Even though, to quote a kid who didn't understand what he was talking about, you were "wasting money on a depreciating asset." I am 100 percent sure that the depreciation was far surpassed by the education we received in the new town. Dad, you have always been an inspiration and a big reason I have become a "jack of all trades." To you both, thanks for trusting in me when I continued to set bigger and bigger goals, including the insane idea of moving across the world to continue my education. To my brother Aaron, thank you for being willing to listen when no one else will. My friends, Richard, Rich, Josh, Josh, and Mell, thank you for always being a source of strength, support, and a sounding board from all across the world.

My grandparents had a huge effect on me, but only one is still around. Grandpa Rich (Sr.) I want to thank you for embodying what it means to have skin in the game. Your discipline and work ethic were on full display when I got the chance to work with you in the summers during high school. I will never forget how you would drop me off at home at eleven or midnight, then be back at six the next morning in a giant loaded flatbed truck; somehow having managed to sleep, wake up, get to the yard, load a truck, and drive said truck back to my parents' place in time for me to roll out of bed and meet you on the road. The work ethic you have managed to keep into your old age inspires me, along with your entrepreneurial spirit. I will never let this go.

Lastly, I feel very lucky in life. I feel blessed to be able to come here and take the time to learn so much in a land so far from my own. However, on the flip side of Luck, as Morgan Housel writes, is Risk. There are far too many people to mention here who accompanied me on the road of life who are no longer with us. Through war, war at home, and inside. There are so many people with whom I served or grew up who died along the way.

To all of you who are no longer here, who have not been blessed with the opportunities I have had, please know that I take no opportunity for granted, appreciate every day, and carry you ever forward with me in my heart.

I will never forget you or your sacrifices. Never above you, never below you, always by your side.

With Love and Gratitude,
Richard Dreyling III
2024

Abstract

Digital Transformation: Artificial Intelligence Enablement in Public Services

This Ph.D. thesis researches the complex topic of sociotechnical systems involving the digital transformation process of adopting artificial intelligence enabled digital public services. The research aims to understand and synthesize the technical, institutional, organizational, and social qualitative data from researching the Estonian Bürokratt program.

The research comprises of six publications which represent four primary phases of research. In the first phase, "defining the research problem" the author conducted exploratory qualitative case study research about the KrattAI pilot and the social, legal, and technical considerations when choosing to adopt artificial intelligence and machine learning in the public sector. In this phase the researcher conducted an interdisciplinary systematic literature review to understand the academic literature relating to the use of artificial intelligence in e-government services. This yielded an understanding of the field and knowledge gaps.

The second phase of the research in the thesis called, "data collection and analysis" the researcher conducted two descriptive qualitative case studies with different aims. One of the publications that contains these results is focused on the technical integrations and planning processes required to go from a strategic vision of AI-enabled services to a usable product for citizens and residents. The second focused on the organizational and institutional factors that inform the technical and organizational architecture.

Phase three is called "practical application," and the end result of this phase is an artificial intelligence maturity model for use in the public sector, which uses the knowledge gained in the research up to this point as well as the design science methodology to create an artifact which may be of use to public sector organizations adopting artificial intelligence projects.

The final phase of the research is forward-looking and an integrative review of the challenges facing public entities that would like to adopt GenAI chatbots.

This thesis’ primary contribution is an in-depth research on the complex and difficult topic of AI-enabled digital public service adoption using one of the only existing government programs attempting to build this. The thesis uses interdisciplinarity to create aspects that may be helpful to researchers from many fields and public administrations looking for practical work based on multidisciplinary theory to aid in implementation of AI digital public services.

Kokkuvõte

Digisiire ja tehisintellekti rakenduseeldused avalikes teenustes

Käesolev doktoritöö uurib sotsiaaltehniliste süsteemide keerulist teemat, mis hõlmab tehisintellekti toega digitaalsete avalike teenuste kasutuselevõtu digitaalset transformatsiooni. Töö eesmärk on mõista ja sünteesida Eesti Bürokratt programmi uurimisel saadud tehnilisi, institutsionaalseid, organisatsioonilisi ja sotsiaalseid kvalitatiivseid andmeid.

Doktoritöö koosneb kuuest publikatsioonist, mis esindavad uurimistöö nelja peamist etappi. Esimeses etapis, üurimisprobleemi määratlemine", viis autor läbi kvalitatiivse juhtumiuuringu KrattAI pilootprojekti kohta ning analüüsis tehisintellekti ja masinõppe kasutuselevõtu sotsiaalseid, juriidilisi ja tehnilisi kaalutlusi avalikus sektoris. Selles faasis viis autor läbi interdistsiplinaarse süstemaatilise kirjanduse ülevaate, et mõista tehisintellekti kasutamist digitaalsed avalikud teenused käsitlevat akadeemilist kirjandust, mis võimaldas mõista valdkonda ja tuvastada teadmiste lüngad.

Teises etapis, mida nimetatakse ändmete kogumiseks ja analüüsiks", viis autor läbi kaks, erinevate eesmärkidega, kirjeldavat kvalitatiivset juhtumiuuringut. Üks neid tulemusi sisaldavatest väljaannetest keskendus tehnilistele integratsioonidele ja planeerimisprotsessidele, mis on vajalikud tehisintellekti toega teenuste strateegilisest visioonist kasutatava tooteni jõudmiseks kodanike ja elanike jaoks. Teine keskendus organisatsioonilistele ja institutsionaalsetele teguritele, mis mõjutavad tehnilist ja organisatsioonilist arhitektuuri.

Kolmandat etappi nimetatakse "praktiliseks rakendamiseks"ja selle faasi lõpptulemuseks on tehisintellekti küpsuse mudel, mis kasutab seni uurimistöö käigus omandatud teadmisi ning disainiteaduse (design science) metoodikat, et luua artefakt, mis võib olla kasulik tehisintellekti projekte rakendavatele avaliku sektori organisatsioonidele.

Uurimise viimane etapp on tulevikku suunatud ja integreeriv ülevaade väljakutsetest, millega seisavad silmitsi avalik-õiguslikud üksused, kes soovivad kasutusele võtta GenAI vestlusroboteid.

Käesoleva doktoritöö peamine panus on põhjalik uurimus tehisintellekti toega digitaalsete avalike teenuste kasutuselevõtu keerulisel teemal, kasutades ühte vähestest olemasolevatest valitsusprogrammidest, mis seda üritavad luua. Doktoritöös kasutatakse interdistsiplinaarsust, et, olla abiks teadlastele ja riigiasutustele paljudest valdkondadest, kes otsivad multidistsiplinaarsel teoorial põhinevat praktilist tööd tehisintellekti digitaalsete avalike teenuste rakendamisel.

Publications (Article I - IV)

Publication I

R. Dreyling, , T. Tammet, and I. Pappel, “Technology push in ai-enabled services: How to master technology integration in case of bürokratt,” *SN Computer Science*, vol. Future Data Science Engineering, no. 5, p. 738, 2024

SN Computer Science (2024) 5:738
https://doi.org/10.1007/s42979-024-03064-0

ORIGINAL RESEARCH

Technology Push in AI-Enabled Services: How to Master Technology Integration in Case of Bürokratt

Richard Dreyling III[1] · Tanel Tammet[1] · Ingrid Pappel[1]

Received: 2 May 2024 / Accepted: 18 June 2024
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd. 2024

Abstract
The Estonian program Bürokratt is meant to implement AI-enabled digital public services, beginning as a chatbot and moving toward virtual assistant-based transformation of access to services. This study is meant to explore the planning process and technical integrations required for implementing AI-enabled services for this program. The paper is a qualitative case study based on the triangulation of academic literature, secondary document reviews and semi-structured qualitative interviews. The research findings include the importance of including stakeholders in the design phase, the challenges of complexity in AI and strategic use of open-source components and procurement as it applies to the Bürokratt program. This paper seeks to contribute to the literature on digital transformation and AI in the public sector by examining a detailed case study of the approach to developing and implementing Bürokratt. Future research could include deeper studies on digital transformation processes in different sectors or further development of AI maturity models.

Keywords Digital transformation · Bürokratt · Virtual assistants · Open-source components · Public sector AI · Technology integration · AI maturity models · e-Governance

Introduction

As governments continue to keep up with changes in technologies, they are increasingly forced to consider adopting new and less proven technology sets. Sometimes the technology on the bleeding edge, like artificial intelligence, can have an issue of the solution in search of a problem resulting from beginning with a top-down vision and then attempting to translate that vision to a usable service [1]. Within government, technology push can cause agencies to focus on digital transformation projects that do not have a proven source of public value [2].

One way to decrease this risk is to follow a planning process when going from vision to developed system and implementation that includes stakeholders, especially users, in the design phase of a new technology to ensure that there is the need for the end product [3] and that it is designed to fit the specific purposes of users [4]. When this is not possible, it may be helpful to have organizational flexibility in digital transformations [5].

AI has a challenge of complexity that comes embedded in the technology's nature [6]. Organizations considering AI adoption must consider numerous factors that can affect the readiness of the organization to adopt AI projects and pilots as well as the importance and ramifications of the project for public value [6, 7].

In Estonia the Bürokratt program is being developed and implemented as a response to the government AI strategy and the vision for what the digital state will look like when citizens are able to access digital state services through AI-enabled channels like virtual assistants and chatbots. The current program is a chatbot that is being implemented by government ministries to answer questions from citizens but is developing the capability to execute services.

This paper attempts to contribute to the literature by giving a study of a government entity whose job is to develop, evangelize, and implement an AI enabled technology that is derived from a vision, describes the planning process and technical requirements that the relevant entity uses to implement the proposed AI-enabled e-service. The researchers also present a suggestion for a process through which similar organizations can develop their own AI-enabled

Richard Dreyling III
richard.iii@taltech.ee

1 Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086, Estonia

Published online: 29 July 2024

public service projects distilled from the challenges and opportunities which the Estonian authorities have been and are managing while going from vision to a usable system. This study endeavours to do this by answering the research questions, "How would an existing public service adopt AI enablement, and what are the key steps involved in this process when going from vision to usable system?" and "What technical developments or integrations would AI enablement require?" The methodology employed is a qualitative case study in which the researchers triangulate the academic literature, secondary document review, and semi-structured qualitative interviews.

Case Description

Estonia, as a small country, has continually according to former Google CEO Eric Schmidt "punched above its weight" [8] in the area of the use of technology in public service provision. Estonia has been one of the early countries to create a national Artificial Intelligence strategy. The UN E-government Survey in addition to academic publications have expressed similar sentiments discussing the maturity of the Estonian data exchange platform "X-Road" which has even been implemented in other countries [9–11]. The country has piloted and is currently developing and implementing an application called Bürokratt, which at its core is a part of a visionary new way to access government services through the use of virtual assistants. The government seeks to create a personalized virtual assistant for every citizen. The organization within the government of Estonia responsible for making "reusable building blocks" for AI use in government is the Ministry of Economic Affairs and Communication (MKM). As a part of MKM, the State Information Authority (RIA) has a team tasked with the development and implementation of Bürokratt. In accordance with the Estonian AI strategy, other departments and ministries are also able to make or procure their own AI systems, but the chatbot and virtual assistant functions are primarily found in the Bürokratt development.

The Estonian Government has participated in the formation of EU policy regarding Artificial Intelligence and has a stake in following the guidelines through which the EU attempts to protect residents and citizens. The Estonian national artificial intelligence strategy was originally implemented in 2019–2021. Since the beginning of that period, Estonia has greatly increased the number of implementations of AI in the government context. The current Estonian ecosystem for AI uses open-source components to build an "interoperable network of AI applications" [12]. Even though they can be developed by different organizations within the government with various public-private partnerships, the main goal in the end is to create many AI enabled services [13]. There will be the open-source reusable components for government entities to use in order to build their own AI enabled projects and any service that is currently using the Estonian data exchange platform X-Road will be able to be accessed via AI-enabled digital channels.

In accordance with the legal framework in Estonia at the current moment there are no purely machine decisions. It follows that any system that would be deployed in the country would have to adhere to principles for trustworthy and human centered AI. One of the chief goals of the AI strategy is to aid accessibility and usability of e-Governance services. Architecturally speaking, the changes defined as necessary in Estonian CTO Kristo Vaher's vision paper [14] for the next generation of digital government would still need to be made. For example, when the Bürokratt chatbot POC was conducted, the developers used Apache Kafka to avoid the need to have an architectural change to the data exchange platform X-Road.

Bürokratt represents the future version of e-Governance services that will eventually be accessible from virtual assistants. In the aforementioned step toward this goal, the Estonian government piloted Bürokratt as a proof-of-concept system in 2020 and 2021 that was designed as a network of chatbots. It did this to comply with the Estonian Personal Data Protection Act (PDPA) which, in combination with the guidance of the "once only principle" adhered to by the government, requires that data must reside where it is collected. Prior to a chatbot system, this would result in a log file being left on the Estonian Government's data exchange platform X-Tee or X-road. This allows for citizens to see all of the queries against their personal data. For the chatbot proof of concept, when a query given to the chatbot is not answerable by a chatbot which is associated with a particular government authority and its databases, the query is passed to other chatbots until the intent is understood and the query is able to be answered. This would be done on the back end and the person asking the question would not have seen the shuffling of responding back-end database associated chat interfaces.

In 2021, the Estonian government was able to fund the Bürokratt initiative which intends to take the proof of concept from the pilot chatbot and bring it to fruition. However, there are challenges as it pertains to building the system in a way that complies with all the appropriate laws and is able to deliver a product in addition to the challenges posed by the lack of off the shelf components able to be used to fulfill certain required functions.

The currently available components to the architecture of Bürokratt include a translation engine that supports seven languages, a text analytics tool, a Speech synthesis tool that uses neural networks, a speech recognition tool, and the chatbot [12]. Development and release of improvements and implementation of Bürokratt is ongoing. The Estonian government's approach to envisioning and planning for a digital state with virtual assistant-enabled services may offer valuable insights to public administrations seeking to do the same. Their process of evaluating capabilities, creating a strategic roadmap, and addressing challenges while adhering to their own technical infrastructure and history presents a useful model for researchers and policymakers to consider when implementing AI-enabled digital public services.

Background

Digital Transformation

Digital transformation is a complex topic and action that in practice requires stakeholders from many siloes to collaborate to solve a problem related to the sociotechnical system of the organization [15]. The practical definition of DT as digitalization of a process [2] is informative and the phenomenon of multistakeholder engagement also applies to those digital transformations in which organizations are attempting to adopt artificial intelligence to digitalize a process or service inside of an organization. It is important to also consider that Mergel goes further and defines a digital transformation as something that is a technological implementation that has ramifications across the organization which cause feedback between parts of the organization and even the customer base, in this case that of the citizenry [16]. Although the current case under examination is a chatbot and some evidence indicates that a chatbot in itself does not cause large amounts of feedback across the organization [17] none of the chatbots which were analysed for the study were part of a larger program that is meant to create a method for citizens to access digital public services with AI powered tools.

There exists in the adoption of AI what Jöhnk referred to as inherent complexity [6]. If a routine digital transformation requires internal competence, it is reasonable to assume that it is at least as important in AI related projects to adhere to Luciano's recommendations of IT and business alignment and competence [2]. There are many academically published maturity models that look at various aspects of maturity evaluate the organization in many categories [7]. The factors identified as frequently appearing factors in the studied list of AI maturity models by Sadiq et al., are "Data, Analytics, Technology and Tools, Intelligent Automation, Governance, People, and Organization" [2021]. The successful implementation and public value creation from digital transformation initiatives can be dependent on the existence of certain skills inside an organization among other factors [2]. One of the challenges of hiring professionals for AI and data science purposes into public administrations is that there are many parts of the required skills that are found in disparate areas of competences [18]. A person who is excellent at the technical parts of AI may not have certain business and management skills required to be able to complete all the necessary tasks of a public sector AI functional specialist [19]. It is important to have people inside the organization who have enough competence in the technology to be able to understand if potential private sector partners are overpromising or have goals and aims that are at odds with the public administration adopting the technology [20].

Instantiating change in any organization comes across resistance [21, 22]. Digital transformations have at their heart a more complex change due to the feedback in the organization [16] from a technical and organizational perspective. And the changes discussed in this paper have an even more challenging type of technology associated with them [6]. When facing something this complex and these many factors of complexity adding to one another, it may be useful to know how a specific organization approached such a task.

Technology Push and a Solution in Search of a Problem

Two primary strategies are currently considered to be most in favour when it comes to digital products among private sector companies. These are the technology push and demand-pull perspectives [1]. The concept of demand pull is fairly clear. This entails finding a problem in the market that has not been met but has a significant source of acknowledgement of the need and then a company produces a product that fills the need [1]. Technology push on the other hand is when there is not a significant awareness of the demand, the firm is tasked with pushing the technology in the market to drive adoption and ensure that the diffusion of technology gets to a point that "crosses the chasm" [23].

The concept of a solution in search of a problem is mentioned frequently in the technology sector when engineering heavy teams develop something without validating it first with consumers. This can sometime lead to a product with an unknown market. In the academic literature, from an economic perspective, a solution in search of a problem can also considered to exist when "there is no evidence of a significant market failure that needs fixing" in the case of government interventions [24].

Empirical evidence shows that although the demand market pull paradigm of innovation in the market works better for digital start-ups, it is also possible for technology push to achieve success. Some evidence points to the combination for the paradigms to potentially lead to less optimal results [25]. Although governments may not have the same success metrics, it is reasonable to believe that if they are going to create a technological innovation the ideal result would be that the citizenry chose to uptake the technology and use it.

By the nature of an innovation starting in R&D centres or a national strategy and vision for the public sector to adopt AI enabled public services the resulting product is in a position of technology push. This necessarily means that the technology has facets that have not been put into the market in the same context. This has ramifications for the process of finding relevant technology sets to integrate because no off the shelf solution is likely to exist. The process for selecting technologies to integrate and develop becomes slightly more complicated than it would be were there a set of existing technologies the government could easily adopt. Multiple processes become a part of the larger strategy-to-product planning processes whether in government or business [26].

The use of AI in the Public Sector

The public sector across the world has been trying to reckon with the adoption of AI in recent times. The amount of academic literature about the topic of the use of AI in public services and the public sector has expanded greatly since 2019 [27, 28]. The focus of the literature can vary in many cases but there tends to be a focus on different aspects depending upon the disciplines of the authors. Governance is a key aspect of the challenges faced by the public sector when addressing AI [29]. This field of inquiry concerns more than the uses of AI in the public sector alone, for example governments must reckon with the potential for AI to affect the labour market among other difficult choices that will have to be made by governments.

Bias can be of concern when discussing use of decision support systems (DSS) which have an effect on the lives of citizens [30]. However, when looking at AI-enabled public services, when the primary part of the system that is AI enabled is purely the channel through which citizens and residents can access services, then the main bias type which is germane to the discussion is the native speaker bias [31]. Machine language systems have been found to have difficulty when detecting speakers who are non -native speakers of the language [32]. In addition, there has been research related to the mitigating the biases of non-native English speakers [33, 34]. The case in this research is a system designed to, in the opening phases work for Estonian speakers due partially to the lack of Estonian language inclusion in mobile phones and virtual assistant providers. Although the literature primarily discusses these biases as they pertain to the English language the potential for challenges related to this bias is possible when local residents whose native language is not Estonian begin to use the technology.

Trustworthiness and AI in Digital Public Services

Even with the complexity, adoption of AI in government continues. The EU has been one of the governmental organisations which has sought to get ahead of the challenges posed by AI and ameliorate some of them through regulation. There have been an increasing number of Artificial Intelligence adoption cases in the EU. The EU is also a leader in attempting to ensure that the use of AI is conducted in a trustworthy manner. As of 2021 there were 143 AI use cases in the public sector available to be investigated on the AI Watch Github [35]. This does not mean that these are public services, but it demonstrates a large adoption rate nonetheless.

The use of AI-enabled public services through chatbots expands upon the maturity of e-government services in the country. Layne and Lee [36] discuss stages of development of these services. This model was expanded and the use of a successful program for the implementing these AI-enabled public services would fight the highest level of integration according to [37].

Trustworthiness is an important factor in the adoption of e-government services [38]. This effect also likely to transfer to the adoption and acceptance of AI technologies. Avoiding black box reasoning and ensuring that the system has a high degree of transparency and causability are important for increasing acceptance among users through increasing trust [39]. The transparency of decision-making increases trust among the users and those who work with the AI system. In addition, the increased trust will aid optimization with the human operators of the system because they will know better when they should defer to the machine or when the decision could be faulty or biased.

Methodology

This piece of research is an explorative qualitative case study and part of research aimed at investigating the digital transformation process surrounding the Bürokratt program and its development and adoption by various government entities. The qualitative case study is considered an appropriate methodology when the research questions begin with “what” and “how” and the goal is to explore a topic of research in its environment [40]. This research focuses on the process through which the team chose to address the vision and bring it to fruition as well as how they did this.

The primary data collection techniques are semi-structured interviews, workshops, and document review.

Semi-structured interviews are a good way to have the benefits of flexibility in the interview, avoiding the pitfalls of structured interviews while still following a interview schedule [41].

The research team conducted workshops and interviews with ten experts from academia, the private sector as well as the MKM and RIA personnel related to the envisioning, development and implementation of Bürokratt as well as some employees from stakeholder organizations. This sample was selected using the snowball method and asking interviewees for recommendations who to talk to next. Interview sampling stopped when it was deemed that little increased knowledge would be gained from further interviews. The interviews and workshops were an average time of an hour and fifteen minutes. The interviews occurred from 2020 when the pilot was being planned and conducted until 2024 during the development and implementation of the Bürokratt was well under way with the majority of the interviews having taken place in 2023. Interviews were recorded via a device or when interviews occurred over a video conference software, through the software itself.

Interviewee	Title	Participation
Interviewee 1	Former CIO of Estonia, Managing partner of Consulting firm	Interview
Interviewee 2	Project Manager for AI	Two Interviews, Workshop
Interviewee 3	Product Manager for Bürokratt	Two Interviews, Workshop
Interviewee 4	Bürokratt Architect	Three Interviews
Interviewee 5	Project Manager for Bürokratt Services	Interview, Workshop
Interviewee 6	AI Researcher and Consultant	Interview
Interviewee 7	Head of Machine Learning and NLP	Interview
Interviewee 8	Team Lead – Client Organization	Interview
Interviewee 9	Chief Data Officer - Estonia	Interview
Interviewee 10	Chief Technology Officer – Nordic Institute for Interoperability Solutions	Interview

The team then transcribed the interview audio files using the machine learning speech to text programs Otter.ai and Microsoft Teams where applicable. After this was complete the team listened to the interviews and corrected the machine transcriptions then qualitatively coded the files with Atlas.ti computer aided qualitative data analysis (CAQDAS) tool to extract the relevant themes and global themes to inform the results. The coding technique used was a combination of inductive and deductive coding. The deductive codes were derived from influential papers on digital transformation, and AI maturity models, and the use of AI in government [6, 7, 35]. The research team agreed that the deductive coding schema would be appropriate to ensure that the literature frameworks were represented in the analysis of this exploratory study. These themes cover the generalized topic matter of digital transformation, AI maturity, and the use of AI in the public sector. This allows for the derivation of global themes that ensure the analysis is relevant to these lenses, which the research team thought would inform the meta-analysis of the data to help establish global themes. After the initial deductive coding process, inductive coding was used. When this was complete, the team conducted a qualitative analysis and extracted quotations which elucidated the points that answered the research questions.

Results

Adopting AI Enablement for Existing Services

When considering introducing AI enabled digital public services there are several options that face public administrators. There are three options that these boil down to. The first, is to develop the entire system and mechanisms for completing digital services from nothing. The second is to create AI enabled digital services by integrating systems that may not have a direct role in current digital state services. The third is to integrate systems which already conduct services to an AI-enabled component or set of components which allow for an AI-enabled channel to instantiate the digital public services.

The level of effort, organization and technical transformation required varies based partly upon which of these three paths are available to the decision makers deciding strategy. One of the first challenges of a digital public service is that the service itself has to already be digitalized to adopt an AI enablement option. If this infrastructure is not in place, the organization will have to create some sort of mechanism and the prerequisite items to be able to securely conduct services. Developing and implementing all of this from no existing systems is logically more resource intensive than integrating existing systems to achieve the same outcome. However, with this method there the potential that there is no or less reckoning with technical debt that may exist in the organizational IT and software infrastructure.

Another option is to potentially take systems that have no direct role in public services but are operational and useful in their context and integrate them to a digital channel. Another paper [42] explored the potential for integrating a decision support system used in the Estonian Unemployment Insurance Fund (EUIF) [42] with Bürokratt to create a new AI-enabled digital public service that would allow for residents and citizens to be able to access the decision support system's output through a chatbot. The DSS predicts the odds of the person being unemployed after 180 days. This idea did not achieve any traction though because of three separate issues. Prioritization is the first reason; the existing digital services are planned to be accessible via Bürokratt before any other integrations are planned. In addition, the Estonian government's policy of human in the loop, human centric-AI dictates that the decision on whether to use the predictive output lies with the unemployment counsellor at their sole discretion. There is also the issue of form factor, in which some services do not have ideal formats to accomplish via virtual assistant or chatbot, including for example the signing of long contracts that are better read on a browser or pdf than through a chatbot output. Although the example above proposed to stakeholders was immediately set aside, the concept of integrating existing systems is possible. However, it also comes with certain caveats and considerations which make it a somewhat complex option even if the level of effort for development and implementation is theoretically less than a greenfield process would be.

In the case of the Bürokratt vision, this means that the AI enablement would be a digital public service accessible through a digital channel that has an AI component through which the citizen or resident can access the service, like a chatbot or a virtual assistant. One significant competitive advantage that Estonia has in this line of inquiry is that they already have existing public services available through digital channels with all of the legally mandated requirements standardized. This includes the national data exchange platform X-road which can exchange data between individuals, the public sector, and private sector entities. In addition, the security infrastructure and eID are in place for residents and citizens to be able to authorize and sign transactions digitally. The existing infrastructure allows for time stamping and hashing of documents to be able to legally verify when a digital transaction was signed, and what version of a document was signed. The Bürokratt architect illustrates the heart of this idea when he states:

> "I always say that the goal is to make sure it doesn't matter if we have one, ten, 3,000, or any number of the combination of these e-services. But these can be done. And in a way that we don't create any of the e-services in-house, we provide a technical solution and very good graphical user interface to make sure that anybody can create e-services in a really simple way" (Interviewee 4).

This speaks to the manner in which the Bürokratt team is able to approach building the AI-enablement from the existing service base and legally compliant technical infrastructure to create a module that will be able to complete the services when a potential citizen or resident is using a chatbot or virtual assistant to access the service.

Key Steps and Challenges to Overcome in the Process

The Estonian case is instructive in the manner in which they approached the problem of AI enablement. This section addresses the process with which the government approached and thought about the processes as extracted from the qualitative steps. The Estonian Ministry of Economic Affairs and Communication (MKM) and State Information Authority (RIA) began with the national AI strategy. This led indirectly to the vision put forth by the former CIO of Estonia [13] of e-governance and digital state services in a world in which virtual assistants are ubiquitous. From this, the Chief Technology Officer (CTO) of Estonia described the technical changes which he could foresee being necessary to make an AI enabled public service possible given the current state of the existing technical infrastructure [14]. Interviewee 1 states, "The white paper [13] was really basically put the flesh to the bone saying hey, 'So what it's going to be like.' And secondly, the Vision Paper [14] was also to kick delivery in motion."

One challenge faced with adopting public services as happens often when discussing digital technologies, the anticipated changes were not yet fully anticipated. Interviewee 1 states, "But clearly the implementation goes with differently than we thought. When we put the vision paper out, we sketched out some initial steps. We knew it was going to take work, but this was always been meant to be, not like an actual road map of delivery, but more or less saying, these are some of the streams of work." In actuality, after the idea of the future product that would enable services was completed, the evaluation of currently available tools did go as expected. "We assumed when we were writing the paper for example, 'linking up to Siri is going to be like quite an easy thing, right?' What came out in the work is that actually the way that the Siris, the Google Assistant, the others are built up, they were nowhere near ready to what Bürokratt was meant to offer." This outcome was expected in the way that the government viewed the initial steps of this process as an experiment. This became the pilot and this mentality continued into the development process.

As a part of the overall process, the government identified gaps in what was necessary to build the vision into a usable product. Then they identified the integration points where they could use existing tools and technologies to help build the end product as well as evaluating what the current systems could and could not do. This happened during and after the pilot experimentation phase.

The next step was to determine the procurement strategy. This took place after the pilot but before the development began in earnest. In the case of MKM, the decision was made to centralize development leadership in RIA. However, due to consistent challenges in the Estonian government related to the number of available developers, the decision was made to use public private partnerships and to have external developers execute on the technical architecture the architect designed. "But we have a really small team which procures. And we can do it in a way that we have can have any number of developers from any number of partners. We have no developers in house" (Interviewee 4). In this case due to considerations beyond technical ones, this decision was made. In addition, the MKM and RIA teams chose to conduct all of the procurements in public on the Bürokratt GitHub.

After this the next step was to determine the technical architecture. Consideration to the importance of keeping competence requirements as low as possible clearly shows when Interviewee 4 states, "In a new procurement, we define the requirements that the developers who will be creating new services with DSLs have to be junior developers. Because we have a problem that developers, when they on board, they say oh my god, it's just writing (text) files. I don't want to do it. It's too boring. And they I get this. I get this, but it's a happy problem for me. " This mindful simplifying of the code came from the interaction of the requirements with the procurement strategy and helps to prevent vendor lock-in because more developers are able to participate because the level of sophistication is less.

The next steps occur in a cycle which will be familiar to many who ascribe to the agile methodologies or cyclical processes in general. The cyclical process from the technical architecture is shown in Fig. 1. The team plans a tentative roadmap including backlog. Then the team procures and develops the product. The team then implements the technology while continuing evangelizing the technology to create more implementation opportunities. Then the product becomes generally available to clients and end users. The feedback derived from these people then contributes to the

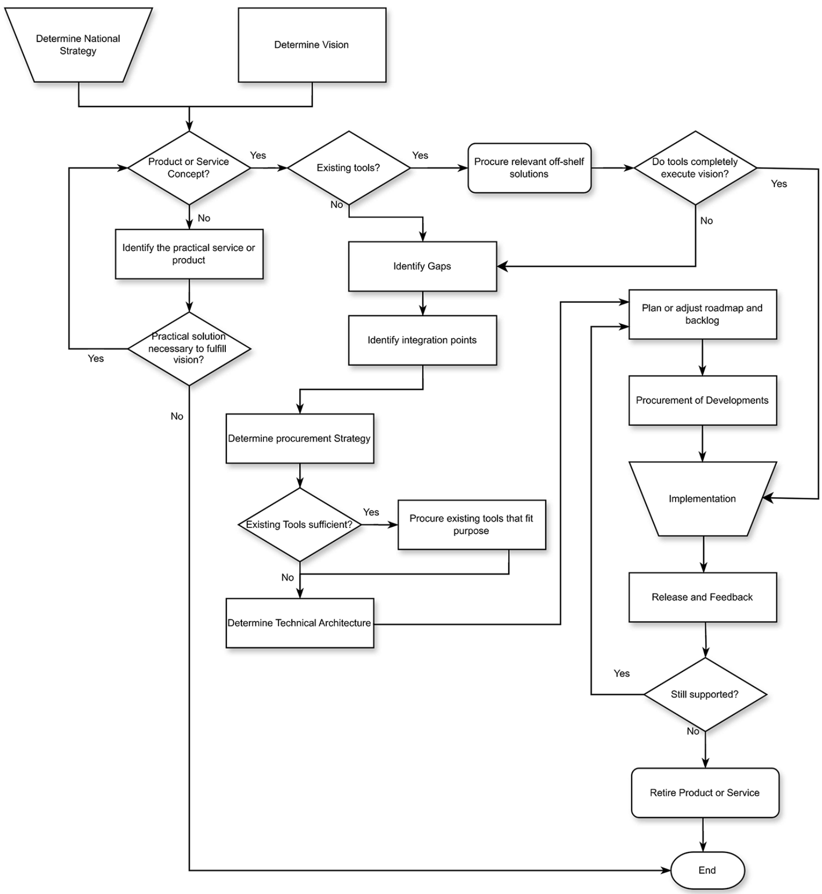

Fig. 1 Planning Process Flow Followed by the Estonian Government in the Vision to Product for Bürokratt

roadmap and backlog. Feedback also occurs in the form of user data which can also be useful for continuous improvement. And so on until the relevant authorities choose to make the product and associated digital channels end of life or retired.

This series of steps is best described by Interviewee 4, “But right now we have this practice of, we have this full backlog on epic levels. And the business side knows what to do. And we have this technical grooming with technical partners, and everybody is on the same page, understand that you want it, why you want it and set these requirements as user stories. It’s not just that I want it. It’s you have to know why you want it. And then it makes sense.” In this way, the RIA team managing the development and implementation of Bürokratt follows agile best practices for this cycle of developments.

What to Develop or Integrate

The difficult answer is that the technical developments and integrations depend on the situation for each public entity that is interested in enabling public services with AI. The Bürokratt case had different perceptions of what the developments and integrations would be. Because of their planning process and the part in which they considered high-level technical integrations, they had the knowledge that they would most likely have to integrate third party systems into the solution for Bürokratt. The quote above stated as much. However, this section focuses on the solutions with which Bürokratt in the second version uses for integrations, after the initial experimentation phase.

With Bürokratt, the goal is to have an Estonian language capable chatbot, a language not supported by any of the large virtual assistant providers, which will allow for eligible individuals to complete government services. The method of accomplishing this is to develop open-source reusable components, combined with some off-the-shelf solutions to be used in combination with a service module that allows for the use of Bürokratt to execute government services using X-Road REST APIs. In the roadmap, the plan is to develop the speech to text capability for voice functionality, among other things. To complete this, the RIA team procured development for Ruuter, Distributed Messaging Rooms, and a Service Module.

The tools and integrations are also illustrative as to the options a public entity may have at their disposal. When it comes to the components, in some cases it is possible to use off the shelf solutions. In the case of Bürokratt, the database solutions are off the shelf, at the current moment PostgreSQL and OpenSearch by default. Rasa gives some chat and intent detection functionality. For complex queries there is in version two, an integration with an external LLM provider. All queries sent to the LLM provider are stripped of identifying information prior to being sent to the provider. The team ties to solve for security and privacy concerns. Figure 2shows a list of the integrations currently used for the version two of the software. The section labelled “Bürokratt’s Functionalities” are the elements the user will have access to and see. The technical functionalities are divided. Bükstack components, with the exception of the Distributed Message Rooms (DMR) are open-source reusable components developed for general use. The only tool developed purely for Bürokratt is the chat widget. Third-party components consist of the software developed by third parties that is deployed locally.

Discussion

Planning processes have many challenges and even different methodologies of approaching technological developments, integrations, and running IT infrastructure. When considering a topic as complex as going from a national strategy and vision to AI-enabled digital public services, there are many potential ways to go about it. This paper’s goal is to extract and analyse how the Estonian government did this with the Bürokratt program. One of the key challenges is that by nature of beginning with a top-down vision and strategy the resulting product can be a solution in search of a problem. This means that the product is subject to the technology push dynamic when considering how people will adopt the technology. In addition, the products that initially inspired the vision may not be able to complete the designated function in a practical product. This can be seen in the inspiration of the Bürokratt vision coming from mobile based virtual assistants that neither had the capability to use the Estonian language nor to fulfil the similar function for the purposes of the government. Procurement can also vary based on the capabilities, human resources and legal guidelines of the organization which is developing and implementing the technology.

In the case of the Estonian government, the plan was to take off-the-shelf products when it was possible. Although it was not possible to integrate with the virtual assistant programs on many mobile phones, they were able to use some component parts that had been developed by outside entities. When the government had to develop components especially in AI related developments, they chose to make them functional building blocks that would be able to be used by any of the government entities who with similar requirements. Even in the case of Bürokratt, the majority of the components which are labelled as Bükstack by the architect are actually developed as these building blocks. Some were even developed prior to the official start of the Bürokratt implementation. Part of this is the goal to make Bürokratt

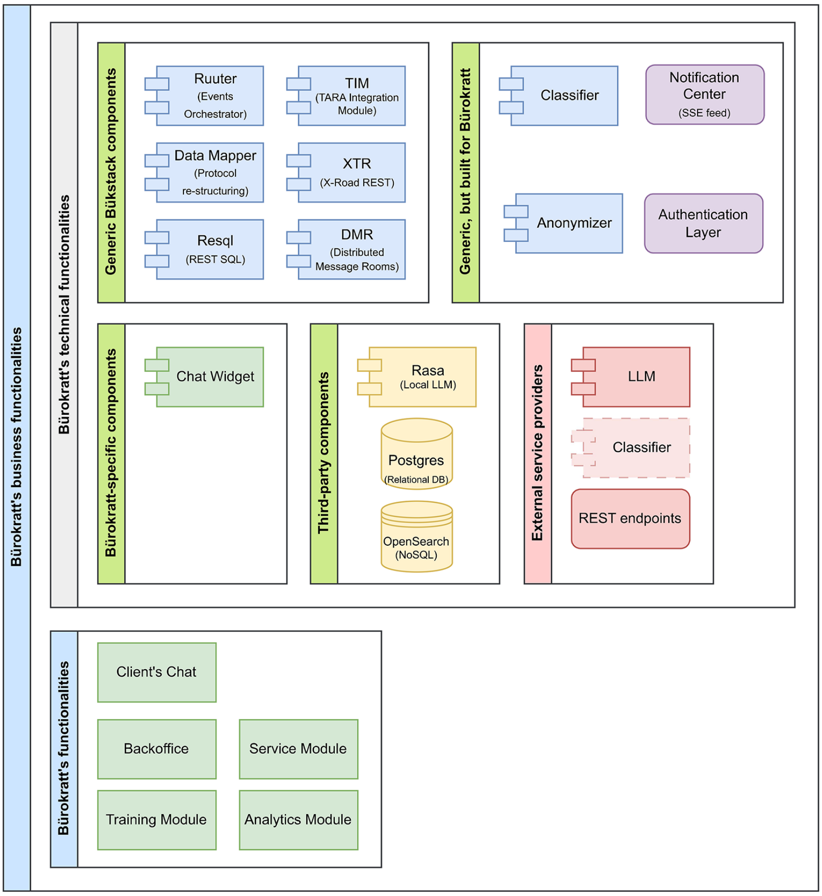

Fig. 2 Bürokratt Business Functionalities and Integrations. (Source: Bürokratt Architect)

into a type of protocol over the long term, with interchangeable parts that fulfil functions. One of the examples of this is that the NoSQL database that was used for Bürokratt was ElasticSearch [43] Since that time, OpenSearch has been chosen as a replacement [44]. In addition, although Rasa has until this point been used as the primary intent detection tool, the government is investigating and has implemented in the upcoming new version, an additional function is to be added that will allow intents and more complex questions to be handled by an external LLM integration [45]. The LLM integration is a way to handle the challenges posed by changing expectations in the population by the large-scale adoption in general availability of LLMs that occurred after the initial development of Bürokratt. This is a form of external pressure detailed in [5]. The flexibility of the architecture should allow for the LLM integration to help ameliorate the challenges posed by traditional chatbot development and the difficulty of getting appropriate responses without heavily involved data scientists in the training process for each government entity. In accordance with the security and privacy concerns related to use of data for AI with external providers [29] the Bürokratt system is set up to strip all of the personal information from the query as it goes to the external providers including the current LLM integration, future LLMs or external intent recognition providers.

The multiple changes in the integrations in the initiative exemplify organizational and technical agility discussed in [5]. In this case, it is not just the flexibility of the organization itself to adopt to new technologies, but the ability to change vendors and even procurement styles when the results and deliverables are not meeting expectations or the service can be improved. From a planning perspective though, the technical development was planned in a way that would give the lowest possible necessary competence requirements following the understanding of the challenges which are presented by [19]. Through allowing GUI based training, simplified code, and a government cloud implementation of the IT infrastructure, the team lowers the necessary technical competence to ease adoption. Even with the technological integrations that have been made in the program, there is still the matter of how to ensure that the adopting ministries will have chatbots that are able to answer difficult and valuable citizen requests.

One significant challenge with solutions in search of a problems is the potential that the technology will not diffuse past innovators and early adopters [24]. In start-up and other private sector organizations the way to do ameliorate this concern is to begin with a problem or potentially have heavy involvement from end users at the beginning of the design phase, often called co-creation. In the case of Bürokratt, the government understood this challenge. The group of stakeholders as defined by Interviewee 5 does not include end users, it is groups that can have an effect on the direction of the product. However, the government did commission a study in which they got citizen feedback on the topic of AI enabled public services, virtual assistants, and chatbots [46]. Through this, they hope to garner enough feedback add qualitatively to the development, implementation and citizen evangelization processes. This would be expected to aid in the diffusion of the technology to help ensure that enough citizens are using the system to be able to warrant continued investment.

Limitations and Future work

The researchers chose the methodology and adhered to it in an attempt to ameliorate bias concerns and ensure a repeatable process that would give similar results given the same process. However, all people including academics, have some level of bias. Other sources of bias in this study could be the set of experts who were willing to participate in giving feedback and interviews to the research team. Bias in the sample selection is a possibility because of the snowball sampling choice. This means that although all the respondents whose opinions were considered had the requisite expertise, they also were from similar disciplinary or at least network backgrounds.

Future research could focus on going deeper into understanding the digital transformation of the ministries involved in adopting AI through the Bürokratt initiative across the government, AI enabled services transformation in other contexts or the further development of an AI maturity model for the public sector.

Conclusion

The complexity of AI projects and their adoption inside of governmental organizations is a relatively new field of study. This paper explains the design of an example of a product that is explicitly derived from a vision which means it is implicitly a "technology push" innovation in the market. Especially in an emerging area like artificial intelligence, such an example is worthwhile and pertinent to the academic community and public administration because it shows a topic that not many administrations have done but are considering or are currently planning or developing. The insights from the planning, development and implementation process investigated for this research paper indicate the necessity to consider in a holistic way the technical architecture, potential integrations and procurement approaches to achieving AI-enabled services.

From a practical perspective, ensuring that a system has a function and design that would add utility to people, especially in the public sector, has its own utility. In private companies, these control mechanisms reflect the constancy of the omnipresent profit and loss statement and other quantitative value measures that directly pertain to the adoption and development or scrapping of an idea. As discussed above, with the public sector, sometimes newer technology projects are adopted because a government wants to show that they can do new things and without a real conception of whether the project has the ability to come to fruition through thoughtful and considered planning, development, implementation and adoption processes it is hoped that the resulting product can add to public value.

Funding ICT Programme support of corresponding author as determined in FP7 in form of scholarship.

Data Availability Interview transcripts available on request via email to corresponding author.

Declarations

Ethical Approval This declaration is not applicable.

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Di Stefano G, Gambardella A, Verona G. Technology push and demand pull perspectives in innovation studies: current findings and future research directions. Res Policy. 2012;41(8):1283–95. https://doi.org/10.1016/j.respol.2012.03.021.
2. Luciano E, Wiedenhöft M. G C 2020 The role of organizational citizenship behavior and strategic alignment in increasing the generation of public value through digital

transformation. Proc 13th Int Conf Theory Pract Electron Gov https://doi.org/10.1145/3428502.3428577.
3. Virkar S, Alexopoulos C, Tsekeridou S, Novak AS. A user-centred analysis of decision support requirements in legal informatics. Government Inform Q. 2022;39(3). https://doi.org/10.1016/j.giq.2022.101713.
4. Distel B. Bringing light into the shadows: a qualitative interview study on citizens' NonAdoption of eGovernment. Electron J e-Government. 2018;16(2):98–105.
5. Vial G. Understanding digital transformation: a review and a research agenda. J Strateg Inf Syst. 2019;28(2):118–44. https://doi.org/10.1016/j.jsis.2019.01.003.
6. Jöhnk J, Weißert M, Wyrtki K. Ready or not, AI Comes— An interview study of organizational AI readiness factors. Bus Inform Syst Eng. 2020;63(1):5–20. https://doi.org/10.1007/s12599-020-00676-7.
7. Sadiq RB, Safie N, Rahman A, A.H. and, Goudarzi S. Artificial intelligence maturity model: a systematic literature review. PeerJ Comput Sci. 2021;7:e661.
8. Sikkut S, Eric S. Our Competitive Advantage: Trust and Democracy in the Era of AI + Q&A, Tallinn Digital Summit. 2022. https://www.youtube.com/watch?v=nBzlyikG0Wg. Accessed 20 April, 2024.
9. Saputro R, Pappel I, Vainsalu H, Lips S, Draheim D. Prerequisites for the Adoption of the X - Road Interoperability and Data Exchange Framework: A Comparative Study, 2020 Seventh International Conference on eDemocracy & eGovernment (ICEDEG), Buenos Aires, Argentina, 2020, pp. 216–222, https://doi.org/10.1109/ICEDEG48599.2020.9096704. (2020).
10. Pappel I, Tsap V, Draheim D. The e-LocGov Model for Introducing e-Governance into local governments: an Estonian case study. IEEE Trans Emerg Top Comput. 2021;9(2):597–611. https://doi.org/10.1109/TETC.2019.2910199. 1 April-June 2021.
11. UN E-Government survey. Department of Economic and Social Affiairs, Division for Public Institutions and Digital Government. https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2020. 2020. Accessed April 23, 2024.
12. Lopes Gonçalves D. Digital Public Services based on open source: Case study on Bürokratt. Joinup European Commission. [Online]. Available: https://joinup.ec.europa.eu/collection/open-source-observatory-osor/document/digital-public-services-based-open-source-case-study-burokratt. (2022).
13. SikkutS,VelsbergO,VaherK.*KrattAIthenextstageofdigitalservices in Estonia. https://*f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_b44333364c1e941c9a5a7633f7555bdd*fpdf.*
14. Vaher K. Next Generation Digital Government Architecture. Republic of Estonia GCIO Office; 2020.
15. Breaugh J, Rackwitz M, Hammerschmid G. Leadership and institutional design in collaborative government digitalisation: evidence from Belgium, Denmark, Estonia, Germany, and the UK. Government Inform Q. 2023;40(2):p101788. https://doi.org/10.1016/j.giq.2022.101788.
16. Mergel I, Edelmann N, Haug N. Defining digital transformation: results from expert interviews. Government Inform Q. 2019. https://doi.org/10.1016/j.giq.2019.06.002.
17. Van Noordt C, Misuraca G. Exploratory insights on Artificial Intelligence for Government in Europe. Social Sci Comput Rev. 2020;089443932098044. https://doi.org/10.1177/0894439320980449.
18. Baškarada S, Koronios A. Unicorn data scientist: the rarest of breeds. Program. 2017;51(1):65–74. https://doi.org/10.1108/prog-07-2016-0053.
19. Labanava A, Dreyling RM, Mortati M, Liiv I, Pappel I. Capacity Building in Government: Towards Developing a Standard for a Functional Specialist in AI for Public Services. FDSE. 2022; https://doi.org/10.1007/978-981-19-8069-5_34.
20. Fukas P, Rebstadt J, Remark F, Thomas O. Developing an Artificial Intelligence Maturity Model for Auditing (2021). ECIS 2021 Research Papers. 133. https://aisel.aisnet.org/ecis2021_rp/133.
21. Burnes B. The origins of Lewin's three-step model of change. J Appl Behav Sci. 2020;56(1):32–59.
22. Kotter JP. Why transformation efforts fail. Harvard Business Rev. 1995;73(2):59–67.
23. Moore GA, McKenna R. Crossing the chasm:Marketing and Selling High-tech Products to Mainstream Customers. HarperBusiness. 1999.
24. Reddy B, Evans DS. Government Preferences for Promoting Open-Source Software: A Solution in Search of a ProblemMay 21, (2002). Available at SSRN:https://ssrn.com/abstract=313202or https://doi.org/10.2139/ssrn.313202.
25. Guo H, Wang C, Su Z, Wang D. Technology push or market pull? Strategic Orientation in Business Model Design and Digital Startup Performance. J Prod Innov Manage. 2020. https://doi.org/10.1111/jpim.12526.
26. Brem A, Voigt KI. Integration of market pull and technology push in the corporate front end and innovation management—insights from the German software industry. Technovation. 2009;29(5):351–67.
27. van Noordt C, Misuraca G. Evaluating the impact of artificial intelligence technologies in public services: towards an assessment framework. InProceedings of the 13th international conference on theory and practice of electronic governance. 2020 Sep 23 (pp. 8–16).https://doi.org/10.1145/3428502.3428504.
28. Dreyling RM, Tammet T, Pappel I. Artificial Intelligence Use in e-Government Services: A Systematic Interdisciplinary Literature Review. In International Conference on Future Data and Security Engineering 2022 (pp. 547–559). Springer, Singapore.
29. Wirtz BW, Weyerer JC, Kehl I. Governance of artificial intelligence: a risk and guideline-based integrative framework. Government Inform Q. 2022;39(4):101685. https://doi.org/10.1080/01900692.2020.1749851.
30. Appelman N, Fathaigh RÒ, van Hoboken JS, Welfare. Risk profiling and Fundamental rights: the case of SyRI in the Netherlands. J. Intell. Prop. Info. Tech. & Elec. Com L. 2021;12:257.
31. Choi LJ. Interrogating structural bias in language technology: focusing on the case of voice chatbots in South Korea. Sustainability. 2022;14(20):13177. https://doi.org/10.3390/su142013177.
32. Hutiri WT, Ding AY. Bias in automated speaker recognition. In proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency 2022 Jun 21 (pp. 230–247). https://doi.org/10.1145/3531146.3533089.
33. Zhiltsova A, Caton S, Mulwa C. Mitigation of unintended biases against non-native English texts in sentiment analysis.
34. Zhang Y, Zhang Y, Halpern BM, Patel T, Scharenborg O. Mitigating bias against non-native accents. InInterspeech 2022 (pp. 3168–3172).
35. Hernandez L. 2021. Dataset with cases of Artificial Intelligence usage in the public sector available as Open data. Joinup European Commission. https://joinup.ec.europa.eu/collection/elise-european-location-interoperability-solutions-e-government/news/143-ai-cases-public-sector-are-available-open-data. (2021).
36. Layne K, Lee J. Developing fully functional E-government: a four stage model. Government Inform Q. 2001;18(2):122–36.
37. Lemke F, Taveter K, Erlenheim R, Pappel I, Draheim D, Janssen M. Stage models for moving from E-Government to Smart Government. In: Chugunov A, Khodachek I, Misnikov Y, Trutnev D, editors. Electronic governance and open Society: challenges in Eurasia. EGOSE 2019. Communications in Computer and Information Science. Volume 1135. Cham: Springer; 2020. https://doi.org/10.1007/978-3-030-39296-3_12.
38. Bélanger F, Carter L. Trust and risk in e-government adoption. J Strateg Inf Syst. 2008;17(2):165–76.

39. Shin D. The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. Int J Hum Comput Stud. 2021;146:p102551.
40. Yin RK. *Case Study Research Design and Methods. (5th ed.)*. Thousand Oaks, CA., Sage (2014).
41. Wirtz BW, Weyerer JC, Geyer C. Artificial intelligence and the public sector—applications and challenges. Int J Public Adm. 2019;42(7):596–615. https://doi.org/10.1080/01900692.2018.1498103.
42. Dreyling RM III, Tammet T, Pappel I. Digital Transformation Insights from an AI Solution in Search of a Problem. InInternational Conference on Future Data and Security Engineering 2023 Nov 17 (pp. 341–351). Singapore: Springer Nature Singapore.
43. Ministry of Economic Affairs and Communication (MKM). Decision support of the unemployment fund OTT. https://www.kratid.ee/kasutuslood. (2020).
44. National Code repository. Bürokratt Architecture. High level flow https://koodivaramu.eesti.ee/buerokratt/architecture/buerokratt-high-level.
45. Buerokratt Architecture Github. https://github.com/buerokratt/Architecture/tree/main/docs.
46. Open Data Portal. Minsitry of Economic Affiars and Communication. Awareness and opinions of Estonian residents about artificial intelligence survey https://avaandmed.eesti.ee/datasets/%22eesti-elanike-teadlikkus-ja-arvamused-tehisintellektist%22-uuringu-alusandmed.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Publication II

R. Dreyling, E. B. Jackson, T. Tammet, A. Labanava, and I. Pappel, "Social, legal, and technical considerations for machine learning and artificial intelligence systems in government.," in *ICEIS (1)*, pp. 701–708, 2021

Social, Legal, and Technical Considerations for Machine Learning and Artificial Intelligence Systems in Government

Richard Dreyling[1][a], Eric Jackson[1][b] Tanel Tammet[1][c] Alena Labanava[1][d] and Ingrid Pappel[1][e]
[1]*Department of Software Science, School of Information Technologies, Tallinn University of Technology, Akadeemia Tee 15a, Tallinn, Estonia*
richard.iii@taltech.ee, eric.jackson@taltech.ee

Keywords: Artificial Intelligence, Machine Learning, Government Services, E-Government, E-Governance, AI Ethics, ML Ethics, AI Bias, ML Bias, Estonia.

Abstract: Expansion of technology has led to governments increasingly reconciling with advanced technologies like machine learning and artificial intelligence. Research has covered the ethical considerations of AI as well as legal and technical aspects of the operation of these systems within the framework of government. This research is an introduction to the topic in the Estonian context which uses a multidisciplinary inquiry based in the theoretical framework of technology adoption and getting citizens to use these services for their benefit. (Suggest that there are first results as well)

1 INTRODUCTION

The twenty-first century has brought with it the expansion of digital transformation in the public and private sectors. Information and communications technologies have been used by the public and private sectors to enhance efficiency and service delivery. Since the introduction of the microchip in 1971, the technological revolution has changed the way businesses conduct affairs as well as the ways in which governments handle governance tasks (Perez, 2002, 2010). The advent of the internet and the information technology boom has changed not only the ways that bureaucrats can govern, but also the items which must be governed. Expansion of technology provides new ways for businesses and citizens to push against laws in ways that governments could not have imagined at the advent of the microchip.

Governments have adopted E-government methodologies and platforms to be able to use information and communications technologies to streamline the business processes of government and deliver services to citizens in a more efficient manner. One country that has developed a reputation for the use of ICTs in service provision is Estonia. The small Baltic country has put a lot of effort into digitizing many government services. They offer many services online with the ability for citizens to accomplish the majority of their interactions with the government through authentication through various forms of electronic ID. The country has worked to minimize its digital divide, ranked as the twelfth most inclusive country in the world in a recent index (Economist Intelligence Unit, 2020). The combination of a tech savvy populace that also trusts its government has helped these efforts be successful. Since the 2000's Estonia has offered increasing government service offerings online with electronic identification (eID) and data exchange between government entities in a secure and tracked manner. They have even been successful in bringing e-Government to local municipalities and attracting people to virtual residency through their e-Residency program (Pappel et. al., 2015) (Kimmo et. al., 2018).

The expansion of computing power since the early 2010s, driven by graphics processing units has allowed for the expansion of artificial intelligence and machine learning research. Governments across the world have begun to use AI and ML in the conduct of

[a] https://orcid.org/0000-0001-7269-6236
[b] https://orcid.org/0000-0003-3726-5814
[c] https://orcid.org/0000-0003-4414-3874
[d] https://orcid.org/0000-0001-6318-7704
[e] https://orcid.org/0000-0002-5148-9841

government business and governance to try to better deliver services to citizens and in some cases control them.

However, Estonia would like to go further in using technology to help make life better for its citizens. In March of 2020, the Chief Technology Officer of Estonia launched a Next Generation Digital State Architecture Vision Paper. In this document the CTO discusses the concept of AI enabled virtual assistants to help achieve easier access to government services (Vaher, 2020). In Estonia, the public sector has a history of cooperating with academia in the country to ensure that the public officials were following the best available science at the time. Because of this cooperation, the research began after the release of the paper to investigate and support the topics laid out by the CTO through academic research.

This introductory research seeks to find the answer to the main research question that asks, "How can virtual assistant systems affect eGovernance services in Estonia?" This multidisciplinary paper will address the ways in which virtual assistant systems can enable government services in Estonia, what particular challenges are inherent to the general practice of using AI and machine learning in government and the specific case. This paper seeks to introduce this research topic as well as formalize the research gaps involved and lay out a roadmap and preliminary results regarding automation of government services and enablement through Next Generation Digital Government Architecture (NGDA) initiative. This paper will be an overview of the challenges that AI and ML enabled programs in government face from a legal, technical, and social perspective and how stakeholders in current active pilot programs in Estonia intend to contend with these challenges.

2 STATE OF THE ART

2.1 Introduction to Estonian E-Government Systems

The Estonian government has used technology as a way to ameliorate the issues caused by having a small population from which they can hire government employees. The Estonian government now has almost all services able to be completed by eID validated transactions online. The key building blocks necessary for this infrastructure from a technological perspective are the electronic ID and the Estonian implementation of a data exchange layer they call "X-Tee" or "X-Road" in English. All the official identification cards have a cryptographic chip capable of electronic authentication and giving signatures to documents. This enables use of a public key infrastructure (PKI) that enables encryption and digital signing of documents and transactions that are secure and legally binding. The X-Road acts as a data exchange layer. Developed in the early 2000's. X-Road uses security servers to authorize service clients and service providers. Any transaction, to include making changes to data or accessing data, registers with the time-stamping server and leaves a trace. Through this architecture, they ensure authentication, authorization, and accounting (Vaher, 2020). The time stamping server leaves a time hack on any transaction, which must be accompanied by an eID signature. Estonia ensured at the time that these innovations came into use that they included the social aspects, legal framework, and technical aspects of the solution all were primed in order to encourage use of the solution. The state subsidized the purchase of the ID cards containing the eID signing ability, as well as partnered with banks to make the IDs useful for logging into internet banking and completing transactions. The country also chose the best technical solution for eID, and has continued to handle any technical or security issues that have arisen from the non-compliance to best practices by contractors (Lips et. al., 2018). This enhances trust among the citizenry which is a likely factor in the strong adoption of the Estonian population of e-services.

Similar to other contexts, when a country is an early adopter of new technologies, technical debt and other phenomena can make further innovation a difficult task. The vision paper released by the Chief Technology Officer (CTO) of Estonia proposes methods to continue the path of innovation in the area of public sector service implementation. Some of these initiatives primarily focus on updating the technology currently in use in the Estonian eGovernance architecture. These include moving from monolithic applications toward an event driven microservices architecture. More than simply discussing some architectural changes, this paper outlines a vision that would have Estonians conducting government services through virtual assistants.

As outlined in the NGDA paper the uses for artificial intelligence and machine learning in government are called "Kratt." This name is based on an entity from Estonian mythology (Scholl & Velsberg, 2020). KrattAI "is first a vision of how public services should digitally work in the age of artificial intelligence" (Sikkut et. al., 2020). When the Estonian Government refers to a "Kratt" this specifies a use of AI or ML, whereas the specific signifier “KrattAI” is the initiative that focuses on the aforementioned provision of government services that use the human computer interaction method of virtual assistants or chatbots (Scholl & Velsberg, 2020).

2.2 Technology Adoption Theories

One area of research has tried to codify the factors which can help to predict whether a citizen or employee will adopt a piece of technology. The area of technology adoption models began with the Theory of Reasoned Action (TRA) in 1975, which focused primarily on a social psychological explanation of people’s perceptions and norms (Fishbein and Ajzen, 1975). Fishbein and Ajzen then expanded TRA into the Theory of Planned Behavior (TPB). From these, the research expanded into many different theories related to the adoption of technology in different contexts. Some of these include the Technology Acceptance Model (TAM), the expansions of TAM, including TAM2 and TAM3, as well as The Unified Theory of Acceptance and Use of Technology (UTAUT), and (UTAUT2). Each of these have various identified ontologies of factors which the researchers believed would affect technology adoption. Some of these theories have similarities that help to show the importance of factors that would encourage successful execution of projects containing machine learning and artificial intelligence. For example, in the Technology Acceptance Model’s third version (TAM3) some of the determinants include the perceived ease of use of a piece of technology. These factors are “computer anxiety,” “perceived enjoyment,” “objective usability,” as well as “perceived usefulness” from earlier TAM models (Venkatesh and Bala). In the Unified Theory of Acceptance and Use of Technology (UTAUT) the determinants of “effort expectancy,” and “performance expectancy” are relevant to the specific challenges of AI and ML based systems in government, even though this model originally considered the corporate sphere (Venkatesh et al., 2003). These factors from a theoretical perspective can be considered proxies for the general concepts of effectiveness, usefulness, and usability. These concepts show the reasons that practitioners in the government would want to ensure that a tool that uses AI and ML are useful, effective, and usable by everyday citizens. In further research conducted on technology adoption shows trust to be an important factor in the use of e-government services (Grimsley & Meehan, 2007), (Colesca 2005, pp.39), (Carter & Bélanger, 2005). In addition, further research stated that trust is one of the most important factors related to “behaviour intention” (Alharbi et. al., 2016, pp. 1). For the solution to be successfully adopted in the populace, trust could be a key factor. The theories regarding technology adoption also apply to adoption of artificial intelligence and machine learning in government. Specific factors in the areas social, technical, and legal concerns will have an effect on the success of the Estonian Next Generation Digital Government Architecture (NGDGA) and its artificial intelligence related proposals

2.3 Social Perspective

Specific social challenges exist related to the effectiveness, usefulness, and usability of machine learning and artificial intelligence initiatives in government. One of the main challenges to AI and ML initiatives is that these will end up enhancing current disparities through the digital divide, and bias.

One issue that causes concerning social factors is research related to bias in AI and ML. A report called Government by Algorithm suggests that three findings became apparent in their investigation of the literature. They found that “the potential for machine learning to encode bias is significant” (Freeman Engstrom, et al., 2020). The researchers used the example of criminal risk assessment scores in the United States that have different rates of false positives for those of different ethnic groups (Freeman Engstrom, et al., 2020). The reasons for this are that AI can become biased due to programming or training, based on the data inputted to train the model, which can have the effect of making bias integral to the decision making of the AI (Mehr 2017)(Center for Public Impact, 2017). In addition, proposed methods of keeping machine learning fair can potentially not co-exist if these methods must have more than one definition of “fairness” (Freeman Engstrom, et al., 2020). If considering multiple groups of people who have multiple differences in race or gender it is impossible to ensure that all possible key performance metrics are equal across the groups (Freeman Engstrom, et al., 2020). The report also pointed out the necessity to consider how human and AI-assisted decisions correlate with one another because the bias in the AI and ML decisions comes from the human decision making (Freeman Engstrom, et al., 2020).

The context of the above review of the literature was the United States. However, the European Parliamentary Research Service has also considered bias in these issues. They explain a resolution adopted by the European Parliament in 2019. The report states, "'any AI model deployed should have ethics by design'. The resolution specifically mentions four sets of issues in relation to the ethical discussion: 1) human-centric technology; 2) embedded values in technology – ethical-by-design; 3) decision-making – limits to the autonomy of artificial intelligence and robotics and 4) transparency, bias and explainability of algorithms (pp. 9). The European Parliament guidance on these systems recommends that any AI or ML based system does not perpetuate bias by ensuring ethical behavior integration in systems. When taken into account this in a practical sense puts the responsibility of making sure that bias and lack of ethics do not perpetuate current disparities.

2.4 Legal Considerations

Any Estonian implementation using AI for government purposes should comply with Estonian and European Law with regard to automated decision making and data protection. In the European Union at the moment there are competing existing frameworks for adopting AI. One assessment suggested that, "a common EU framework on ethics has the potential to bring the European Union €294.9 billion in additional GDP and 4.6 million additional jobs by 2030" (Evas, 2020 pp. 1). Beyond the general approach to data protection brought by the GDPR, Europe does not have specific legislation dictating how member states can implement AI in their countries. However, Estonia has a law that may impact the ability for AI to achieve what could be considered its full potential.

The Personal Data Protection act passed in 2018 has provisions that give specific purposes and criteria that need to be met for data processing which could mean that organizations other than the one which collected the data are unable to use AI or ML applications to provide services (Personal Data Protection Act, 2018). This law also provides specific criteria that must be met for automated decision making. According to some legal experts, one of these criteria means that the only two state registers which would qualify are the land register and company register because they are "considered having legal effect" (Kerikmäe & Pärn-Lee, 2020 pp. 6). In practice this means leads to the hypothesis that that any automated capability would be used more as a decision support system for a human decision maker. This law also has ramifications for technical best practices that will be discussed in the following section. In addition, the cross-border aspect of the data sovereignty requirements put in place by GDPR, the US CLOUD Act and the Estonian PDPA may make integration with the large virtual assistant providers complicated (Varughese, 2020).

2.4 Technical Concerns

The vision for a next generation digital government architecture must overcome technical challenges to ensure success. Although chatbots originated in private sector use cases, researchers have studied chatbots as a method of allowing consumers to directly speak through an AI mediated platform to government entities to assist in completing tasks (Akkaya & Krcmar, 2019) (Freeman Engstrom & Ho, 2020) (Androutsopoulou et. al., 2018) (Mehr, 2017). A chatbot is a system that has to accomplish several tasks. The chatbot must use natural language processing be able to interpret intent of a customer or citizen. After understanding intent, the bot should be able to complete the required tasks or connect the citizen with the relevant stakeholders to help assist them in completing the task. A chatbot may use supervised learning and when properly trained will improve its ability to operate the more it is used.

Data is a key factor in the accuracy of machine learning and artificial intelligence systems. Estonia has had over twenty years of e-government service experience. Because of this, they have accumulated massive amounts of data and have done a better job than some other countries of ensuring this data is machine readable (Scholl & Velsberg, 2020). The way the Estonian PDPA has been put into practice makes one legal challenge into a technical challenge. Estonia follows the "once only principle," which means that data is stored where it is collected and the citizen should not have to provide it to other government authorities. For example, if the police would like to know a person's address, they should query the population registry database. This leaves a signature through X-Road, the data exchange layer. When discussing an AI system though, even though the Estonian government may have more data available it is in various databases around the country. Researchers have attempted to ameliorate some of the organizational issues related to data, quality, and formatting in Estonia (Tepandi et. al., 2017). Because of this, there is no massive data pool from which the chatbots could be trained. This theoretically would make it difficult for the chatbot and virtual assistant programs to be able to gain the accuracy necessary to achieve instant citizen uptake. Although, they could get better as time continues if the proper training and feedback mechanisms were implemented into the workflows of the system.

The NGDGA document elaborates on a vision in which chatbots would move beyond a single instance on a website toward a virtual assistant model. One of the options could be to integrate the Estonian government's hypothetical chatbot with the large virtual assistant providers to provide a more robust experience for the citizen (Vaher, 2020). This poses an issue because the Estonian language does not have support in the large virtual assistant providers or the existing translation APIs are not sufficient in quality. The language issue and the method of integration with virtual assistant providers are issues that must be solved.

A report regarding the United States Federal Government's adoption of AI and ML mentions the concept of internal and external competencies (Freeman Engstrom, et al., 2020). They found that some of the most successful implementations were created by employees of the government who were hired in a capacity such as lawyers and then developed their own machine learning and artificial intelligence capacity on their own time. They recommended to government procurement personnel in the US context to not simply outsource the development of AI and ML projects to private sector contractors. They found that the in-house developed solutions solved some of the issues with data access and source code access that outsourced projects experienced. In the United States the private sector has the advantage when it comes to AI and ML experience. However, Estonia has shown in recent years a propensity to use public private partnerships (PPP) to procure technological expertise that leads to successful projects when the need arises (Paide et. Al., 2018).

Harvard researchers identified five potential use cases for chatbots in the public sector which included, "(i) answering citizens' questions, com- plaints and inquiries through automated AI-based customer support systems, (ii) searching in documents (including legal ones) and providing guidelines to citizens on filling forms, (iii) getting citizens' input and routing them to the responsible public administration office, (iv) translating governmental information, and (v) drafting documents with answers to citizens' questions" (Mehr, 2017) (Androutsopoulou et. al., 2018). The vision put forth by the Estonian government goes further than this and calls for the virtual assistant technology to be able to help the citizen complete tasks (Vaher, 2020). The Mehr report quotes, CEO of Synthesis Corp. Ari Wallach, "'Imagine having direct and constant access to a high-level government concierge that is constantly learning and improving" (2017, pp. 10). This entails having a system that can constantly learn through supervised learning across data sets and stepping into territory which governments have not tread before at scale.

3 METHODOLOGY

To better investigate the current and future states of eGovernance with AI and ML enabled virtual assistants, qualitative methods were used. A review of recent literature served to get preliminary information. In addition, two workshops were conducted to elicit feedback from groups of experts who are stakeholders in the Estonian eGovernance context. Qualitative research has the inherent issue of bias. However, the workshop format and its semi-structured nature gives the participants the ability to express themselves freely and to communicate the way they perceive the issues at hand (Yin, 2014). Due to the early investigatory nature of the research at hand, the qualitative methods have the largest amount of flexibility to gather information to determine the future path of research. This methodology allows for the researcher to get the maximum amount of information from the experts in the field rather than have them conform to already existing theories and phenomena (Gioia et. al., 2012). This represents the best way to ensure that the researchers would not ask leading questions that bias responses when discussing the topics with experts and stakeholders in workshops. The workshops included stakeholders from the Nordic Institute for Interoperability Solutions (NIIS), stakeholders from the Ministry of Economic and Social Affairs of Estonia (MKM) as well as the software development company that is developing the KrattAI chatbot proof of concept (POC).

4 DISCUSSION AND RESULTS

Artificial Intelligence use can be considered to be controversial. Apart from the popular culture depictions of artificial intelligence as an antagonist force toward humanity, there exists a lot of literature on the topic. In section two, a review of some of the social, legal and technical concerns explored some of the issues that a government implementation of AI and ML would have to avoid.

The workshop led to a discussion of these topics and how the Estonian government plans to ameliorate some of the issues presented in section two. The

Estonian vision of may be considered one of the more recent developments in government services due to the initiation of the chatbot proof of concept to eventually directly provide services to citizens. Estonia is working right now to traverse the challenges and barriers which have been pointed out above. From the workshops with stakeholders the researcher gained insights into how the social. legal, and technical challenges have shaped the pilot programs in Estonia. Many of these are interrelated and will be presented in a manner which acknowledges this factor. These methods can inform the ways that other governments may shape their programs to help ameliorate some of the difficult points concerning AI and ML based initiatives.

From a social perspective, getting feedback from users both inside and outside of the government is important for the stakeholders in the various AI and ML programs. This concerns the theoretical grounding of technology adoption in a practical manner. One thing that a stakeholder observed was that though the team tried their best to make the instructions and all relevant materials in as clear language as possible, they got the feedback that some of the directions were too complex for those not already embedded in the IT world. This allowed them to ensure that by the time the services roll out to citizens and ordinary government workers, the likelihood of adoption will increase because they can iterate until usability has increased. They look at usability not only of the end user but of all the stakeholders in the chain who will be using

During the discussions, stakeholders acknowledged the potential for machine learning and artificial intelligence derived bias. However, they pointed out that the Estonian government has signed onto and helped shape the European Parliament's suggestions relating to ethical AI and controls against bias. And in the areas in which there are no standards that are universally accepted, the people in the Estonian government who manage AI suggest them to governing bodies. This helped to shape the way the Estonian government set up the chatbot POC that is the initial step toward the KrattAI vision as well as other Kratts. They decided from the beginning that whenever an AI or ML enabled decision support system would have a decision point that directly affects a citizen's service provision, in accordance with the Estonian law on automation, that a human decisionmaker would be there to make the final decision in some cases. Kerikmäe & Pärn-Lee summarized the guidelines dictating the law in practice as follows, "Human interaction should take place only if the algorithm result turns out negative or if the subject of the administrative decision disputes" (2020 pp. 6). This still does not completely solve the issue of bias due to human decisionmakers over time causing the bias, but it does take steps toward preventing hardcoded bias. Deference of human decisionmakers to automated decision systems is another potential source of problems in this area (Freeman Engstrom, et al., 2020). The stakeholders in this situation use the predictive, prioritization, and optimization abilities from AI to help in areas that the citizen and the government benefit from, not as a punitive function like using AI imagery analysis to determine subsidy compliance based on whether farmers have mowed their land or not. Instead of fining a farmer based on the results, the government would contact the farmer to ask the situation. Sometimes the farmer would have mowed the farm earlier in the year or be ready to do it. This saves government resources from doing on the spot investigation of each farm and farmers appreciate the ability to discuss with officials (Scholl & Velsberg, 2020).

In addition, there are some useful capabilities inside the government which can use AI and automate items that have no decision impact on the citizen but increase the ability for government responsiveness to the citizen. An example of this is internal email forwarding. The Estonian government had a massive problem with citizens emailing officials, employees, or department email addresses requesting information on where to direct their inquiries. One stakeholder mentioned specifically that in addition to normal government duties, some employees had to handle over 1500 emails a day. Some departments have been able to institute decision engines that look for similar inquiries and send responses automatically. This is an example of a situation where the laws as currently written allow for automated decision making. The government also gets feedback from the citizen to see if this forwarding solved their issue. However, it must be mentioned that this process is done on a department-by-department basis and has not been implemented across the entire government.

The method of handling the chatbot inquiries in the absence of a united data pool is novel and also helps solve the issue of referring citizens to the right authorities. The design of the KrattAI chatbot POC is to have networks of many chatbots with their own knowledge which can speak to each other. They do not store the data from the transaction. This way, when a citizen contacts the chatbot and asks a question, the chatbots can refer the citizen to the chatbot with the proper knowledge base. The KrattAI chatbot POC is not yet to the point of executing government transactions but the POC has proven that a network of chatbots can allow for the proper functioning to find the proper chatbot for a transaction. This method maintains the legal boundaries put into place by Estonia while effectively handling the technical concerns from not having large data pools with which they can train the NLP engines of the chatbots.

According to the workshop attendees, in agreement with the NGDA vision paper, there are changes in the current E-governance architecture are necessary to enable the vision of virtual assistant enabled services. One change that still must be made is moving X-Road from a synchronous communication mode to an asynchronous version of communication. This could include publish, subscribe messaging patterns. The CTO has called this change introducing X-Rooms. X-Rooms would allow more than one verified entity to be party to the communication being passed and not require that both entities be connected at the exact same time. This is key for the vision to be achieved with virtual assistant driven services.

With a PPP the Estonian authorities have managed design, code, and test a system that uses AI and ML for the benefit of the citizen while attempting manage the difficulty points of these types of projects. Limitations of the research are that the number of interactions with stakeholders were few. The projects are also not that far along. The specific partnership potential with public virtual assistant providers is not able to be discussed and legally very complex. Because of these legal complexities, the options for integration to make the chatbot POC able to use virtual assistant capabilities would be conjecture.

Future work will take a specific case for which the virtual assistant capability could be used, and follow the business processes as well as specific technical processes through to the end of the transaction. If possible, an artefact will be designed to help solve a technical issue pertinent to initiatives of similar purpose.

REFERENCES

Ajzen, I and Fishbein M., 1980. Understanding Attitudes and Predicting Social Behavior Englewood Cliffs, NJ: Prentice-Hall, Inc.

Ajzen, I., 1991 The Theory of Planned Behavior, Organizational Behavior and Human Decision Processes 50, pp. 179–211.

Ajzen, I., 2010. The Theory of Planned Behavior, http://www.people.umass.edu/aizen/tpb.html.

Akkaya, C. and Krcmar, H. 2019. Potential Use of Digital Assistants by Governments for Citizen Services: The Case of Germany. In Proceedings of dg.o 2019: *20th Annual International Conference on Digital Government Research*, dg.o 2019, June 18, 2019, Dubai, United Arab Emirates. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3325112.3325241

Alharbi, N., Papadaki, M., & Dowland, P., 2016. The impact of security and its antecedents in behaviour intention of using E-Government services. *Behaviour & Information Technology,* 36(6), 620–636. doi:10.1080/0144929x.2016.1269198

Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. 2019. Transforming the communication between citizens and government through AI-guided chatbots. Gov. Inf. Q., 36, 358-367.

Carter, L., & Bélanger, F., 2005. The utilization of E-Government services: citizen trust, innovation and acceptance factors. Information Systems Journal, 15(1), 5–25. doi:10.1111/j.1365-2575.2005.00183.x

Centre for Public Impact, 2017. Destination unknown: Exploring the impact of Artificial Intelligence on Government. Retrieved from https://publicimpact.blob.core. windows.net/production/2017/09/Destination-Unknown-AI-and-government.pdf.

Colesca, S.,. 2009. Increasing e-trust: A solution to minimize risk in E-Government adoption. Journal of Applied Quantitative Methods. 4.

Economist Intelligence Unit. 2020. The inclusive Internet Index 2020. https://theinclusiveinternet.eiu.com/explore/countries/performance

Engstrom, D.F., Ho, D.E., Sharkey, C.M., Mariano-Florentino, C., 2020. Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies: Report Submitted to the Administrative Conference of the United States.

Engstrom, David Freeman and Ho, Daniel E., 2020. Artificially Intelligent Government: A Review and Agenda (March 9, 2020). Big Data Law (Roland Vogl, ed., 2020, Forthcoming), Available at SSRN: https://ssrn.com/abstract=3551549

European Parliament resolution of 12 February 2019 on a comprehensive European industrial policy on artificial intelligence and robotics (2018/2088(INI)).

Evas, T., 2020. European framework on ethical aspects of artificial intelligence, robotics and related technologies. European added value assessment. European Parliamentary Research Service, European Value Added Unit.

Gioia, D. A., Corley, K. G., & Hamilton, A. L., 2012. Seeking Qualitative Rigor in Inductive Research. Organizational Research Methods, 16(1), 15-31. doi:10.1177/1094428112452151

Grimsley, M., & Meehan, A., 2007. E-Government information systems: Evaluation-led design for public value and client trust. European Journal of Information

Systems,16(2),134–148. doi:10.1057/palgrave.ejis.3000674

Kerikmäe, T., Pärn-Lee, E., 2020. Legal dilemmas of Estonian artificial intelligence strategy: in between of e-society and global race. AI & Soc. https://doi.org/10.1007/s00146-020-01009-8

Kimmo, M.; Pappel, I; Draheim, D., 2018. E-residency as a nation branding case. Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance, ICEGOV2018 : 4 - 6 April 2018, Galway, Ireland. Ed. Kankanhalli, Atreyi; Ojo, Adegboyega; Soares, Delfina. New York, NY: ACM Press, 419–428. (ACM International Conference Proceedings Series).10.1145/3209415.3209447.

Lips, S., Pappel, I., Tsap, V., Draheim, D., 2018. Key factors in coping with large-scale security vulnerabilities in the eID field. Electronic Government and the Information Systems Perspective : *7th International Conference, EGOVIS 2018, Regensburg, Germany, September 3-5, 2018, Proceedings*. Ed. Kö, Andrea; Francesconi, Enrico. Cham: Springer Verlag, 60–70. (Lecture Notes in Computer Science; 11032).10.1007/978-3-319-98349-3_5.

Mehr, H. 2017. Artificial intelligence for citizen services and government. Cambridge, MA: Harvard Kennedy School, Ash Center for Democratic Governance And Innovation. Retrieved January 10, 2021, from https://ash.harvard.edu/files/ash/files/artificial_intelligi nte_for_citizen_services.pdf

Paide, K., Pappel, I., Vainsalu, H., & Draheim, D. (2018). On the Systematic Exploitation of the Estonian Data Exchange Layer X-Road for Strengthening Public-Private Partnerships. Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance - ICEGOV '18. doi:10.1145/3209415.3209441.

Pappel, I., Tsap, V., Draheim, D., 2020. The e-LocGov Model for Introducing e-Governance into Local Governments: an Estonian Case Study. IEEE Transactions on Emerging Topics in Computing. DOI: 10.1109/TETC.2019.2910199 [ilmumas].

Perez, C., 2002. *Technological Revolutions and Financial Capital*, Books, Edward Elgar Publishing, number 2640.

Perez, C., 2010 Technological Revolutions and Techno-economic paradigms. *Cambridge Journal of Economics,* Vol. 34, No.1, pp. 185-202

Personal Data Protection Act, (AKI).(Est). *2018*.

Scholl, R., Velsberg, O. 2020. AI FOR GOOD LIVE: How Estonia builds the next generation e-government with AI use cases. International Telecommunications Union. https://www.youtube.com/watch?v=O3jiLbYVLQU (Nov. 10, 2020)

Sikkut, S., Velsberg, O., Vaher, K. 2020. #KrattAi: The Next Stage of Digital Public Services in #Estonia, Vision and Concept.

Tepandi, J., Lauk, M., Linros, J., Raspel, P., Piho, G., Pappel, I., Draheim, D., 2017. The data quality framework for the Estonian public sector and its evaluation: establishing a systematic process-oriented viewpoint on cross-organizational data quality. In: Hameurlain, A.; et al. (Ed.). Transactions on Large-Scale Data- and Knowledge-Centered Systems XXXV (1–26). Berlin: Springer. (Lecture Notes in Computer Science; 10680).10.1007/978-3-662-56121-8_1.

Thong, James Y. L.; and Xu, X., 2016. Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead, *Journal of the Association for Information Systems*: Vol. 17 : Iss. 5 , Article 1. DOI: 10.17705/1jais.00428

Vaher, K. 2020. Next Generation Digital Government Architecture.

Varughese, B., 2020. Cross- border data transfer in the context of the GDPR and CLOUD Act. Tilburg Institute for Law, Technology, and Society (TILT) LLM. Law and Technology.

Venkatesh, V. & Bala, H. 2008. Technology Acceptance Model 3 and a Research Agenda on Interventions. Decision Sciences - DECISION SCI. 39. 273-315. 10.1111/j.1540-5915.2008.00192.x.

Venkatesh, V., Morris, M., Davis, G., & Davis. F., 2003. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly, 27(3), 425. doi:10.2307/30036540

Venkatesh, V., Thong, J. Y. L.; and Xu, X. 2016. "Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead," Journal of the Association for Information Systems: Vol. 17 : Iss. 5 , Article 1. DOI: 10.17705/1jais.00428

Yin, R., 2014. Case Study Research: Design and Methods (5th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Publication III

R. Dreyling, J. Lemmik, T. Tammet, and I. Pappel, “An artificial intelligence maturity model for the public sector: Design science approach,” *TalTech Journal of Social Sciences*, vol. 14, no. 2, p. 16, Forthcoming

An Artificial Intelligence Maturity Model for the Public Sector: Design Science Approach

Richard Dreyling
Department of Software Science
Tallinn University of Technology
Akadeemia Tee 15a
Tallinn 12618, Estonia
Email: richard.iii@taltech.ee

Juhani Lemmik
Department of Software Science
Tallinn University of Technology
Akadeemia Tee 15a
Tallinn 12618, Estonia
Email: julemm@taltech.ee

Tanel Tammet
Department of Software Science
Tallinn University of Technology
Akadeemia Tee 15a
Tallinn 12618, Estonia
Email: tanel.tammet@taltech.ee

Ingrid Pappel
Department of Software Science
Tallinn University of Technology
Akadeemia Tee 15a
Tallinn 12618, Estonia
Email: Ingrid.pappel@taltech.ee

Abstract: This paper presents the development of an Artificial Intelligence Maturity Model (AIMM) specifically tailored for public sector organizations to assess their readiness for AI adoption. Using design science methodology, the research synthesizes insights from academic literature and expert consultations to propose a comprehensive AIMM. Through iterative development and expert feedback, the study refines a model that categorizes AI maturity across eight dimensions. The model's validity is assessed through expert evaluations and questionnaires, confirming its relevance and utility in guiding public organizations toward effective AI adoption. This research contributes to the theoretical and practical understanding of AI implementation in the public sector, addressing unique challenges such as procurement models, legal compliance, and organizational capabilities.

1. Introduction

In recent years, there are many governments and public administrations attempting to introduce Artificial intelligence (AI) projects to find efficiency and effectiveness, but AI projects are notoriously difficult to implement and take a high level of various elements within the organization to be able to find success (Van Noordt and Misuraca, 2022; Dreyling, et al., 2021). Many public administrations are implementing projects but they come to varying levels of success. The difficulty is that AI is an inherently complex technology (Jöhnk, et al., 2021) that requires among other things, organizational support, leadership buy in, strong technical and data science competence (Fukas, et al., 2021). Artificial Intelligence maturity models (AIMM) are meant to give organizations a tool through which they can evaluate the organization's maturity when it comes to AI technologies (Sadiq, et al., 2021). With such an instrument, or one to assess the maturity level of a public sector organization to adopt, implement and run a particular AI project. AI expansion is relatively new, even if it cannot be said that the technology itself is new. The nature of the technology presents many potential problems in society (Wirtz, et al., 2020). Challenges in adoption of AI in the public sector include not only the success or failure of projects but broad concepts like data and privacy, and AI bias, which governments must reckon with if they choose to adopt AI technologies to make better government services (Zuiderwijk, et al., 2021; Wirtz, et al., 2019; Dreyling, et al., 2023).

This research originates from discussion with practitioners in the field of AI who have experience consulting with government entities and it identified an anecdotal problem. In the private sector there are many consultancies who go into a customer site and ensure that relevant and necessary pre-requirements for an AI project are present. In addition, private sector organizations have the ultimate way of measuring the potential success or failure of a project in an institutionally relevant way – they conduct a cost benefit analysis projection. On the other hand, in the public sector not only is this measure not a holistic or appropriate one, but many organizations procure pilots (Van Noordt and Misuraca, 2022a) without knowing whether any of the factors necessary to succeed are present in the organization. In the academic literature search for thorough AI maturity models (AIMM) reveals no thoroughly investigated and validated AIMM for use in the public sector which acknowledges the differences in goals and procurement models between the public sector and private sector.

To this end, this paper investigates the following research question: How would public sector organizations analyze the maturity of their organization to adopt AI enablement for digital public services?

To answer this question, researchers use the design science methodology to propose and validate an artifact that serves as an AI maturity model for the public sector.

Initial research is conducted by collecting and analyzing academic and gray literature on the topic of AI maturity models. Initial requirements from this are gathered from literature and interviews with and the public sector which leads to the creation of the artifact. The researchers create an initial version of the AI maturity model and then iterate three times on and validate it with experts in the field until the artifact is validated.

The AI maturity model for the public sector and this paper contribute to the academic literature and to the field by offering a maturity model for public sector organizations seeking to adopt AI by giving them a way to consider their organization and see where they can improve capabilities to have a better potential for a successful AI implementation that adheres to legal, ethical, and security guidelines.

2. Theory

2.1 AI Adoption in the Public sector

From the introduction of e-government and digitalization governments have been trying to use ICTs to improve services and digital governance (Tskhadadze, 2024). An increasing body of work discusses the use of AI in the public sector. This body has elements of looking at implementations of AI in government across the EU (Van Noordt and Misuraca, 2022a) and the

US federal government (Engstrom and Ho, 2021). The Studies across the EU have been associated with AI Watch, which surveys hundreds of public authorities who have at least begun AI related projects. Many potential uses for AI in the public sector are currently being conceived from public administrations optimizing bus routes (Hong et al., 2018), to researchers proposing machine learning systems to make legal language easier to read by citizens (Üveges, 2022).

Public authorities are seeking to adopt AI technologies to be able to claim efficiency and effectiveness gains that the technology can enable by furthering the state of capability within digital solutions to solve longstanding problems (Troitiño, 2022). Growing technological competence can have the potential to increase overall economic well-being (Stavytskyy et al., 2019). At the same time, digital public services are not something that has improved at the same rate of other measurable factors according to the Digital Economy and Society Index (DESI) (Masoura & Malefaki, 2024).

When implementing AI in the context of a public organization there are many potential challenges including social, legal, and technical challenges (Dreyling et. al., 2021; Holmstrom, 2022). The majority of the projects implemented are chatbots or other ways of communicating with the government (Van Noordt and Misuraca, 2022b). There are questions about data protection and efficacy of legislation like GDPR (Kesa & Kerikmäe, 2020). It is debatable if the technology truly contributes to efficiency and effectiveness gains or simply alter the distribution of the channels through which humans contact the government. Other types of implementations include AI uses in back-office processes (Mehr, 2017). In the US it was noted that one group in a legal group which handles disability claims, the staff created machine learning (ML) tools to be able to group like sets of cases. In this way, the judge deciding the cases would be able to study fewer relevant legal codes in a day and was also able to avoid task switching between types of cases (Engstrom and Ho, 2021). The increase in research on the topic and practical number of uses cases indicate that AI use is on the uptake in government.

2.2 Digital Transformation

Digital transformations (DT) are shifts in the organization due to the adoption of a technology that causes iterative and self-impacting changes (Mergel, 2018). This means that once a technology is adopted it may impact the structure or communication of the organization. The adoption of AI does not in and of itself mean that there will be a large-scale digital transformation in the organization.

During a digital transformation, it is important for an organization to adapt to external and internal pressures (Vial, 2021). This can also be the cause of DT. One important factor to consider in this process is the idea that the technology and organization can have mutual and reinforcing impacts. The organizational structure and competencies within the organization have an impact on how the technology is adopted within the organization. And once the technology is integrated, it further impacts the organizational operations and structure of the organization itself.

Organizations and institutions have a large amount of stakeholders and phenomena that they must reconcile in order to accomplish a digital transformation and derive productivity value from the technology adopted (Hinings, et al., 2018; Brynjolfsson and Hitt, 2000). The strategy and systems adopted should reflect the opinions of the stakeholders in the organization. AI technologies, because of the potential for an intelligent system to learn in an exponential way, make it worth considering the impacts upon the organization as well as what secondary and tertiary processes will have to be put into place to maximize the utility of the technology (Brynjolfsson and McAfee, 2014). Especially when considering modifying the channels through which humans are able to access government services, many social, legal, and technical questions must be considered.

Digital transformations are known to often be a response to external and internal forces (Vial, 2021; Mergel, 2019). It is reasonable to believe that recent developments with Large Language Models (LLMs) and the proliferation of their use in general availability may be currently be pushing the level of awareness to the populace in a way that previous iterations of technology diffusion caused firms and governments to initiate the digital transformation process.

The digital transformation process can be informative in considering the elements of what can make a successful AI adoption because it is organizationally a digital transformation with additional complexity added.

For example, the trends in how these transformations take place and have a feedback effect across the organization are valid in AI related transformations. One of the key points in the literature shows the importance of competences, strategy, and culture (Mergel, 2019; Vial, 2021). In the AI adoption process these play an outsized role because the outcome of the project depends so much on the technical talent. Even in an organization that plans to adopt the technology provided by a for profit company through public private partnership, the public personnel need to be able to understand enough from a technical perspective to ensure that the product will do what is promised by the sales team (Labanava, et al. 2022). In addition, because the public sector has laws, norms, and expectations of trust, security, privacy, and data protection the public entity would typically be expected to ensure all of these criteria are met to a reasonable level of public expectation and legal compliance. This is a key to ensuring that the AI is developed and implemented in accordance with public value perspective (PVP) detailed by Moore (1995) and expanded upon by Benington, (2011), Williams & Shearer (2011) and Pang et al., (2014) to show the need for public authorities to use their skills to explore “new opportunities in public value creation” (Lemmik, 2024).

The complexity of digital transformations combined with the difficulty of managing, implementing, and maintaining AI systems has led to a variety of maturity models based around various facets of AI (Sadiq, 2021; Fukas, et al., 2021). These maturity models are designed for sector specific applications like auditing and industry 4.0. None of them directly address the needs of the public sector as it pertains to the way unique position of stewardship that governments have with relation to their populace affects the motivations for and execution of AI adoption (Van Noordt and Misuraca, 2022b; Dreyling et al., 2021).

3. Methodology

Design science methodology in information systems research is a way of informing one of the main sides of information systems research which is the creation and evaluation of new artifacts that can be used by both business and technical practitioners in the field.

According to Hevner et al. (2004) there are 7 design science guidelines. These are design an artifact, problem relevance, design evaluation, research contributions, research rigor, design as a search process, and communication of research (pp. 83).

Becker et al. (2009) built upon the design science in information systems guidance created by Hevner et al. (2004) to create a framework that is meant for the creation of maturity models in research. This work suggests a framework for adapting existing maturity models for a new application, while following Hevner’s principles.

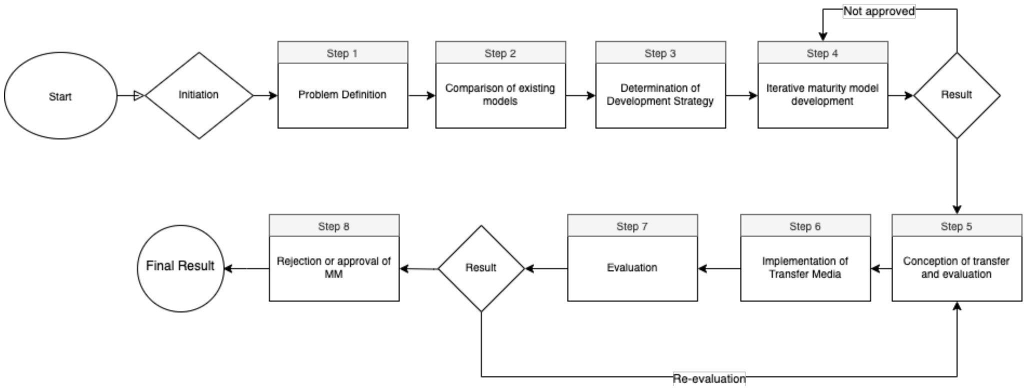

Figure 1: Design Science Steps (Hevner et al, 2004), (Becker et al 2009) and (Alsheiabni et al 2019)

To assemble the design, researchers conducted the search of academic and grey literature. In the building of the artifact for this contribution, it was important to the authors to ensure that previous work in the academic and the private sector were considered and integrated into the initial research that would build the instrument from which expert user feedback would be garnered. The team conducted a survey of the academic and grey literature. The thematic analysis of the papers informed the search process. The grey and academic literature were assembled into a CAQDAS. Researchers then coded the documents according to the elements that might be relevant to the AI Maturity model Framework that answers the research questions in this paper. The coding schema was inductive and focused on elements and the atomic parts which comprised these elements.

This aided in steps one, two, and three. Then during user requirements elicitation, the sample selection was chosen for people who have expertise participating in AI related project. During research on AI related to DT in government, the researchers asked interviewees about the ways that public sector entities evaluate AI maturity. These interviews, 10 over the course of 2023, with 7 interviewees, provide the empirical basis for understanding public administration requirements for an AIMM. The interviewees confirmed the lack of processes of engaging cross-functional teams to measure the ability of the organization to adopt the AI technology or the presence of prerequisites for the technology like data or existing tools (Dreyling et al., 2024). Understanding how to adapt the MM for a new purpose is key to Becker's work (2009). The interviews, through qualitative analysis, yielded the importance of expanding the current AIMMs to fit the public sector's application of an AIMM. The snowball sampling method was used. This follows the sampling recommendations (Cresswell, 2017) for when specific skills are important or there is a narrow target group for sampling. After the interviews and literature informed the broad requirements for categories and levels, the researchers began step four.

3.1 Build, results

The artifact is constructed in the format of maturity models that have gone before, (Fukas, et al., 2022; Sadiq 2021) following the guidelines of Becker et al., (2009) for repurposing existing maturity models for new applications. The team established the relevant categories that must be addressed for the use in question. Then the levels are determined through considering the relevant literature and iterating with expert feedback. Similar to the Fukas et al. (2022) contribution, the authors made the decision to follow the level structure followed by the Capability Maturity Model Integration (CMMI) because it is a standard practice for those in the field, and one seen in academic maturity models as well (Sadiq 2021; Warsinske, et al., 2019). In the search process, the categories that the research team, interviewees and literature agreed were relevant to the maturity model were collected and collated into a list of categories and elements. After developing the concepts that needed to be added to make the AIMM relevant for the public sector, the team commenced step five and conceived of the transfer and evaluation metrics. Step six is to implement the transfer.

3.2 Artifact meta-attributes and validation

The artifact is assembled with the categories and appropriate levels based on a combination of the literature review and expert feedback. The AIMM has eight categories that consist of "technology and tools," "data," "people and competences," "organization and processes," "leadership and strategy," "financial," "law, ethics, and trust," and "security and privacy." The levels are based on CMMI (Warsinske, et al. 2019) and go rom level one to five, initial, assessing, determined, managed, and optimized.

The validation continues until the primary metric is satisfied in Step eight. This particular maturity model's primary metric is whether it suits the purposes of an AI maturity model for the public sector. The question is posed to experts, "does this AI maturity model for the public sector adequately function for its proposed purpose of helping public sector entities evaluate their maturity for developing, implementing, and managing artificial intelligence?" Step seven is to validate the artifact. Because of the transfer concept, the team did three rounds of iteration on the AIMM and levels associated with it. The first iteration involved interviews with two experts who had experience with AI implementation and the public sector to gather feedback on the first version of the AIMM. Then the second iteration consisted of a further validation with four experts from the public sector. Ultimately, the team decided to do another iteration to improve the level definitions for the AIMM. For this a questionnaire with the final iteration of the AIMM and levels was sent to twelve experts with experience in the public sector and AI implementation.

4. Artifact

4.1 Artifact Creation

Because AI is a technology that is not as simple as a set it and forget it system, but instead operates in a system which requires that data be processed and have potential bias related effects, the adoption merits consideration of the effects across many areas that may take place. This means that public entities adopting AI have the obligation to consider the adoption in a manner that is more deliberate than private sector offerings of the same technology.

To this end the researchers conducted a search for previous academic work on “readiness assessments,” “feasibility studies,” and “maturity models” related to AI implementation in organizations. Then the AI maturity model search was expanded to understand the first principles of AI maturity models. The largest AI Maturity model systematic literature review was produced by Sadiq et al. (2021). Furthermore, because the expertise from the private sector would also be considered to be useful because much of the consultancy work is driven by private companies rather than academic works on the same topic. Due to this relevance, the search for gray literature on the topic of AI readiness assessments narrowed to the public sector clients to get an indication of what specific issues and focuses the large consultancies saw when public sector entities attempt to adopt AI.

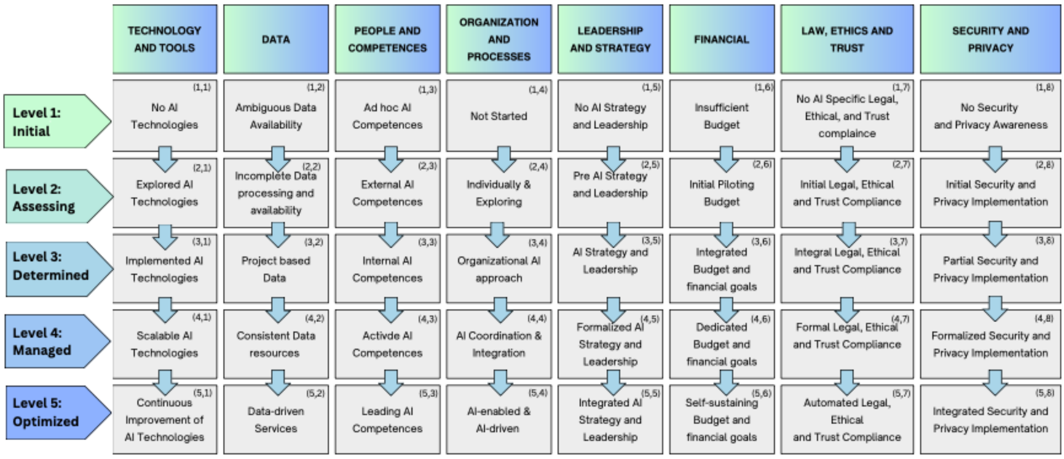

Figure 2: Artificial Intelligence Maturity Model for the Public Sector

Each category has levels associated with it. The schema for levels of maturity and the associated definitions of these levels is derived from CMMI institute's maturity levels. This schema has been used in many maturity models and more importantly, provides a practical and understandable roadmap for organizations seeking to improve their maturity (Sadiq et al, 2021; Becker; 2009, Fukas, et al., 2021). Fukas et al., (2021) is used to determine some of the categories and the use of the CMMI level structure. However, the AIMM in Fukas et al. (2021) was originally intended for use with private sector auditing organizations. Because of this, the team decided to take inspiration, but change many of the metrics and determine which additional categories would be beneficial. In addition, the levels in Fukas et al., (2021) are very brief and it was deemed that it would useful to have more practically descriptive items for the use of officials in the public sector. For example, the budget category and concept of return on investment was removed and replaced with concepts that are closer to the public sector. In addition, products and services was removed in favor of an added security and privacy category to bring in some of the concepts in Das et al. (2023).

Below, tables indicate the categories and definitions of their levels. Each level was determined from a combination of the interviewees and the academic literature applied to the guidelines provided by Warsinske et al., (2019) as they pertain to the levels of organizational maturity in the CMMI framework.

Table 1: Maturity Levels - Technology

Technology	
Initial (1,1)	Organisation is exploring specific AI opportunities and related technologies; it has no access to specialized AI technological solutions (even if it has invested in related technologies, such as DevOps, robotic process automation, or advanced analytics), nor has it partnered with anyone outside the organisation to obtain such solutions.
Assessing (1,2)	There is a good understanding – through Proof of Concept or piloting – of technological needs of an AI initiative; experiments are conducted on personal computers or ad-hoc private, hybrid, or secure public cloud-based environments; there is an agreement on collaboration with a technology partner to support the implementation of AI initiative; AI model training still occurs manually, without automated resource management.
Determined (1,3)	Organisation moves from the piloting phase to the implementation phase based on a chosen AI technology; infrastructure is used to harness GPU power beyond personal computers and ad-hoc deployments; technical controls are established to facilitate human oversight and incorporate explainability features outlined by AI governance practices for production deployment.
Managed (1,4)	There is a technology architecture that allows for cost-effective scaling of existing AI models and streamlined deployment of new models; deployment architecture is formalized; AI solutions are deployed from one business process area to others (e.g. through code reuse); there is the capability to manage AI technology within the organisation through understanding of requirements, need for flexibility and scaling, which are also covered by technology policies (if technology is managed in-house; otherwise this capability is obtained through a partner organisation).
Optimized (1,5)	There is the capability to provide continuous technical support for continuous monitoring (implementing techniques for monitoring AI models in operational environments) and retraining of AI solutions; AI deployment architecture becomes standardized and efficient, aligning with the organisation's strategic objectives; new use cases drive technological advancements; AI is leveraged to manage technology infrastructure automatically (possibly through a partner organisation), optimizing resource provisioning and enabling innovative use cases.

Table 2: Maturity Levels - Data

Data	
Initial (2,1)	Raw data beginning to prepared for AI modelling; some analysis of what data is needed to train the AI model.
Assessing (2,2)	Data for AI modelling is collected, but there may be biases in the information, or it may be incomplete; data is not cleaned and validated and/or properly structured for AI modelling.
Determined (2,3)	Data is compatible with analytics techniques and AI models; data inconsistencies are eliminated and data is standardized and prepared; data is used for AI modelling but only for particular applications. External data sources are tested and able to be used with preprocessing.
Managed (2,4)	Standardised data resources are made available across the organisation; data quality is consistently monitored and improved through data audits with the view to improving existing or deploying new AI models; a significant number of employees have participated in a formal data literacy or proficiency programme.
Optimized (2,5)	Data needs for creating new AI applications are proactively considered and data drives new services; AI applications are constantly improved by providing more or improved data and responding to changing needs and technology.

Table 3: Maturity Levels - People and competencies

People and Competencies	
Initial (3,1)	Some employees have an understanding of the AI possibilities, but these skill sets are ad hoc in random parts of the organisation; there is no formal expertise in data science or AI.
Assessing (3,2)	External resources have been hired or otherwise regularly consulted; some employees in the organisation are familiar with AI possibilities and have tried to build real use cases on that; there is an understanding in the organisation that AI skills are needed, but no systematic approach to their development has been introduced.
Determined (3,3)	Some forms of training on AI have been provided to a sizeable number of employees, either through internal or external training or as on-the-job support; data scientist(s) have been employed or a designated resource for data analytics has been assigned to deploy AI use cases, but on a project basis; external sources may be consulted, but the leading role for working on AI use cases is within the organisation.
Managed (3,4)	AI competencies are defined and consistently developed throughout the organisation to prepare a sizeable number of employees for AI deployment and employees comprehend and contribute to the organisation's AI vision; administrative positions are dedicated to supporting AI solution development efforts and building AI competence.
Optimized (3,5)	The organisation employs leading AI competencies that are also consulted by external parties; AI expertise is fully integrated into every major business unit; constant innovation and improvement take place with organisation's own resources responding to needs and technological changes.

Table 4: Maturity Levels - Organisation and Processes

Organisation and Processes	
Initial (4,1)	There is a basic approach to AI solution development and AI information collection, neither are there defined processes behind first attempts to explore AI use cases; momentum is building but no formally assigned AI sponsor or champion is designated to mature AI capability and tools.
Assessing (4,2)	Initiated AI projects have project-level processes and procedures; solutions are decentralised; refining and testing hypotheses regarding potential AI solutions for business problems takes place through trials and Proof of Concepts; AI champions and teams have been established; evaluation of tools and standards is being considered; cultural change has been addressed.
Determined (4,3)	Organisation-wide approaches are documented for AI projects; there is a common approach for introducing AI initiatives and standardised process of AI delivery; AI project teams are capturing metrics and documenting best practices and tools to achieve outcomes, although it may not yet be systematic across the organisation; AI governance and promotion of AI culture are in place, distinguishing between AI development and maintenance; senior leadership takes a more active role in sponsoring change.
Managed (4,4)	AI initiatives follow policy, governance, and technical standards; AI project teams are capturing metrics and documenting best practices and tools to achieve outcomes across the organisation; leadership makes decisions by analysing data against defined and captured metrics.
Optimized (4,5)	Organisation has broken down silos to integrate data and resources more effectively; best practices and lessons learned from all AI projects are captured to share with each other; leadership is updating policies and procedures by analysing data against defined and captured metrics to optimise organisation-wide impact and create more public value.

Table 5: Maturity Levels - Leadership and strategy

Leadership and Strategy	
Initial (5,1)	Organisational leadership has perceived the initial need to explore the possibilities of AI; there may have been an attempt to develop an AI use case in the organisation; there is no AI strategy and the direction for AI is unclear.
Assessing (5,2)	Leadership supports experimenting with AI and is interested in the results without a commitment to act; there are no goals for AI although the AI project-level objectives are in place. AI strategy takes shape through the analysis of various AI project-level experiences in the organisation.
Determined (5,3)	Leadership recognises and supports the opportunities of AI orally, with resources and actions; there is a collaboration between public organisations driven by the need to find common solutions to common problems through AI; there is an AI strategy with objectives in place signifying a commitment to its deployment in the entire organisation.
Managed (5,4)	AI strategy has become more comprehensive with an emphasis on safety, mandates and accountability, performance metrics, and maintainability of AI models over time; AI strategy aligns with broader organisational strategies.
Optimized (5,5)	AI is a core part of the organisation's vision and AI permeates the entire organisation and its business operations; there are regular discussions at the management level on the effectiveness of the AI strategy, its value and future, and its capacity to innovate constantly.

Table 6: Maturity Levels - Financing

Financing	
Initial (6,1)	Sufficient funding to cover AI pilot, or the funding is external to the organisation.
Assessing (6,2)	There is project-based funding for AI in the phase of experimentation, but not for its deployment.
Determined (6,3)	There is an understanding what is the financial effect of deploying AI projects (financial goals, costs, and benefits); AI projects are managed with clear project-level financial controls; funding of AI projects is through organisation's development budget competing with other innovations.
Managed (6,4)	There is an understanding of the baseline cost of the business area where the AI project will be introduced; a portfolio management plan guides investments into AI; a dedicated AI budget exists.
Optimized (6,5)	Unified budgeting schemes and indicators encompass both business and AI technology domains; there is a long-term perspective on the financing needs of the AI tools in use and on their return on investment and the financing is stable.

Table 7: Maturity Levels - Law, ethics, and trust

Law, ethics, and trust	
Initial (7,1)	Legal, ethical, and trust considerations for AI are not addressed beyond the regular data privacy and protection aspects.
Assessing (7,2)	Awareness about the legal, ethical, and trust considerations is developing, but approaches are incomplete and taken at the project level.
Determined (7,3)	Legal, ethical, and trust considerations are an integral part of every AI project; AI risk assessment is carried out for each AI project before its launch; there is a formal procedure to collect feedback from stakeholders on any issues related to legality, ethics, or trust in the AI system; there is no formal standard yet with what AI projects should comply with.
Managed (7,4)	The organisation checks compliance with formalised AI-related standards through an organisation-wide formalised mechanism before the launch of every AI project; compliance is also monitored throughout the AI implementation phase.
Optimized (7,5)	There is an automated and continuously improving legal, ethical, and trust compliance process, which is optimised to the organisation's needs, with a flexible system that continuously improves and adapts to legal and ethical changes in the AI landscape and the situation on the ground.

Table 8: Maturity Levels - Security and privacy

Security and privacy	
Initial (8,1)	Organisation and/or AI-related data and privacy security policies are rudimentary in the organisation causing unregulated security and information storage and exchange.
Assessing (8,2)	Begin to integrate AI-related specificities into the organisation data and privacy security policies for controlled security and information exchange; checking of AI systems to ensure that private data does not get presented to users.
Determined (8,3)	Defined and implemented organisation data and privacy security policies and processes (including specific AI-related aspects) for digital data exchange are in place ensuring that no private data is released on the part of AI systems; qualitative security risk management implemented in AI projects and beginning research into quantitative risk management.
Managed (8,4)	Implementation and monitoring of organisation-wide AI-specific privacy and data security policies and processes are in place based on standards and testing, ensuring safety and secure digital data exchange, no private data leakage from AI systems; quantitative cyber risk management processes implemented at AI project level.
Optimized (8,5)	Continuous improvement and optimization of organisation-wide organisational and privacy and data security policies and processes are in place based on standards and measurement of trends; regular auditing of AI systems for private data leaks; regular quantitative cyber risk management with mitigating actions is widely enforced across the organisation.

4.2 Artifact Evaluation

For the initial evaluation, as defined in the methodology feedback was garnered from four experts with experience in the public sector, with two of them also having academic and consulting experience. The final artifact was presented via a questionnaire to confirm the levels and maturity model. The answer was in the affirmative to the step eight heuristic of whether the experts believe the maturity model functions for its intended purpose with four respondents. However, two of the experts thought that it did not suit the use. One believed that it was too complicated and weighty for an organization to be able to implement. Another believed that it had language in the levels that led them to believe that it was geared toward the private sector. More clearly defined levels are a benefit in that not many of the maturity models have them. Fukas et al. (2021) created a contribution that actually defined the levels. Many in Sadiq et al. (2021) which is a systematic literature review and represents the body of work for academic maturity models until that year, did not define levels in any explicit way. However, it also creates a challenge; the comments from the two experts who did not think the AIMM was ideal revolved around the specific language and the cumbersome nature of evaluating the levels. The AIMM does what it purports to do. But this does not mean that it will work for every situation. The necessity to have multiple versions, possibly one that is a lightweight checklist for public authorities and another that is the expanded version contained in this paper would be an option to relieve this challenge. This would be a way to popularize the results according to step six in Hevner (2004) and would go beyond the primary methods through which the AIMMs discussed in Sadiq (2021) and found in the literature have historically done.

5. Summary and outlook

The AIMM in this paper is holistic in that it includes many categories that are relevant to the AI maturity of entities of the public sector. However, this becomes a challenge in that many of the potential respondents the questionnaire to validate the final artifact had a response rate of around 50%. In addition, although the team garnered feedback from personnel from seven different countries across the validation process, it is still feasible to get a different result given a different group of experts participating in the validation process. Future work would be conducting an evaluation of an organization and then getting feedback on both the maturity model and the process. In Lemmik (2024) conducted design science research of a self-evaluation method to operationalize this maturity model with a government authority. However, one challenge is that it was not possible to conduct research that provided value to the organization through a live process that engages cross-functional teams to ensure that all the relevant functions and personnel were able to contribute to the knowledge base to give a proper evaluation and be educated on the areas in which the administration can improved. Other future work would be the potential for the AIMM and accompanying documentation to be published in a web application that is available to aid other public administrations in conducting these sessions and give feedback to the researchers.

The concepts and levels of these categories can be discussed among the entire cross-functional team and ensure that maturity levels are satisfactory prior to the procurement or beginning of the project and more importantly that everyone in the organization has cohesion of goals and expectations of the project. Ultimately the results of this maturity model and its evaluation show that it is a difficult task to please all stakeholders when it comes to the idea of implementing AI in the public sector. From discussions with the interviewees in the validation of the maturity model it became apparent that the majority of public institutions are on level one or two across the categories of the maturity model. This paper contributes to the knowledge in the field by suggesting an academic maturity model for use in the public sector adoption of AI. This will hopefully raise the awareness of public authorities as to what organizational maturity in relation to AI adoption can be and give them a method to build a road map for their public administration and thereby improve the potential for successful AI implementation that adds to public value.

References

Alsheibani, S., Cheung, Y. and Messom, C.H., 2019, July. Towards An Artificial Intelligence Maturity Model: From Science Fiction To Business Facts. In PACIS, pp. 46.

Becker, J., Knackstedt, R. and Pöppelbuß, J., 2009. Developing maturity models for IT management: A procedure model and its application. *Business and Information Systems Engineering*, *1*, pp.213-222.

Benington, J., 2011. From private choice to public value. Public value: Theory and practice, pp.31-51.

Brynjolfsson, E. and Hitt, L.M., 2000. Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic perspectives*, *14*(4), pp.23-48.

Brynjolfsson, E. and McAfee, A., 2014. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton and Company.

Creswell, J.W. and Poth, C.N., 2016. *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Das, S.D., Bala, P.K. and Mishra, A.N., 2023. Towards Defining a Trustworthy Artificial Intelligence System Development Maturity Model. Journal of Computer Information Systems, pp.1-22.

Dreyling, R., Jackson, E.B., Tammet, T., Labanava, A. and Pappel, I., 2021, April. Social, Legal, and Technical Considerations for Machine Learning and Artificial Intelligence Systems in Government. In *ICEIS (1)* (pp. 701-708).

Dreyling R., Erlenheim, R., Tammet, T. and Pappel, I., 2023. AI Readiness Assessment for Data-driven Public Service Projects: Change Management and Human Elements of Procurement. *Human Factors, Business Management and Society*, *97*(97).

Dreyling R., Tammet, T., Pappel, I., 2024. Technology Push in AI-Enabled Services: How to master Technology Integration in Case of Bürokratt, SN Computer Science, Future Data Science Engineering 2024.

Engstrom, D.F. and Ho, D.E., 2021. Artificially intelligent government: a review and agenda. *Research Handbook on Big Data Law*, pp.57-86.

Fukas, P., Rebstadt, J., Remark, F. and Thomas, O., 2021. Developing an Artificial Intelligence Maturity Model for Auditing. In *ECIS*.

Hevner, A.R., March, S.T., Park, J. and Ram, S., 2004. Design science in information systems research. *MIS quarterly*, pp.75-105.

Hinings, B., Gegenhuber, T. and Greenwood, R., 2018. Digital innovation and transformation: An institutional perspective. *Information and organization*, *28*(1), pp.52-61.

Holmström, J., 2022. From AI to digital transformation: The AI readiness framework. *Business Horizons*, *65*(3), pp.329-339.

Hong, S., Kim, Y., Park, J: Big data and smart city planning: The caseof Owl Bus in Seoul. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 4492–4500. IEEE.

Jöhnk, J., Weißert, M. and Wyrtki, K., 2021. Ready or not, AI comes—an interview study of organizational AI readiness factors. *Business and Information Systems Engineering*, *63*(1), pp.5-20

Kesa A. and Kerikmäe T. 2020. Artificial Intelligence and the GDPR: Inevitable Nemeses?. TalTech Journal of European Studies, 10 (3), pp. 68-90. https://doi.org/10.1515/bjes-2020-0022

Labanava, A., Dreyling III, R.M., Mortati, M., Liiv, I. and Pappel, I., 2022, November. Capacity Building in Government: Towards Developing a Standard for a Functional Specialist in AI for Public Services. In International Conference on Future Data and Security Engineering (pp. 503-516). Singapore: Springer Nature Singapore.

Masoura M. and Malefaki S., 2023. Evolution of the Digital Economy and Society Index in the European Union: A Socioeconomic Perspective. TalTech Journal of European Studies, Vol.13 (Issue 2), pp. 177-203. https://doi.org/10.2478/bjes-2023-0020.

Mehr, H., Ash, H. and Fellow, D., 2017. Artificial intelligence for citizen services and government. *Ash Cent. Democr. Gov. Innov. Harvard Kennedy Sch., no. August*, pp.1-12.

Mergel, I., Edelmann, N. and Haug, N., 2019. Defining digital transformation: Results from expert interviews. *Government information quarterly*, *36*(4), p.101385.

Moore, M.H., 1997. Creating public value: Strategic management in government. Harvard university press.

Pang, M.S., Lee, G. and DeLone, W.H., 2014. IT resources, organizational capabilities, and value creation in public-sector organizations: a public-value management perspective. Journal of Information Technology, 29, pp.187-205.

Sadiq, R.B., Safie, N., Abd Rahman, A.H. and Goudarzi, S., 2021. Artificial intelligence maturity model: a systematic literature review. *PeerJ Computer Science*, *7*, p.e661.

Stavytskyy A.,Kharlamova G. and Stoica E., 2019. The Analysis of the Digital Economy and Society Index in the EU. TalTech Journal of European Studies, Vol.9 (Issue 3), pp. 245-261. https://doi.org/10.1515/bjes-2019-0032

Troitiño D. (2022) The European Union Facing the 21st Century: The Digital Revolution. TalTech Journal of European Studies, Vol.12 (Issue 1), pp. 60-78. https://doi.org/10.2478/bjes-2022-0003

Tskhadadze K. (2024) E-Government Implementation on the Example of Georgia. TalTech Journal of European Studies, Vol.14 (Issue 1), pp. 253-270. https://doi.org/10.2478/bjes-2024-0012.

Üveges I. (2022) Comprehensibility and Automation: Plain Language in the Era of Digitalization. TalTech Journal of European Studies, Vol.12 (Issue 2), pp. 64-86. https://doi.org/10.2478/bjes-2022-0012.

Van Noordt, C. and Misuraca, G., 2022. Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union. *Government Information Quarterly*, *39*(3), p.101714.

Van Noordt, C. and Misuraca, G., 2022. Exploratory insights on artificial intelligence for government in Europe. *Social Science Computer Review*, *40*(2), pp.426-444.

Vial, G., 2021. Understanding digital transformation: A review and a research agenda. *Managing digital transformation*, pp.13-66.

Warsinske, J., Henry, K., Graff, M., Hoover, C., Malisow, B., Murphy, S., Oakes, C.P., Pajari, G., Parker, J.T., Seidl, D. and Vasquez, M., 2019. *The Official (ISC) 2 Guide to the CISSP CBK Reference*. John Wiley and Sons.

Williams, I. and Shearer, H., 2011. Appraising public value: Past, present and futures. Public administration, 89(4), pp.1367-1384.

Wirtz, B.W., Weyerer, J.C. and Geyer, C., 2019. Artificial intelligence and the public sector—applications and challenges. *International Journal of Public Administration*, *42*(7), pp.596-615.

Wirtz, B.W., Weyerer, J.C. and Sturm, B.J., 2020. The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, *43*(9), pp.818-829.

Zuiderwijk, A., Chen, Y.C. and Salem, F., 2021. Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government information quarterly*, *38*(3), p.101577.

Publication IV

R. M. Dreyling, T. Tammet, and I. Pappel, "Artificial intelligence use in e-government services: A systematic interdisciplinary literature review," in *International Conference on Future Data and Security Engineering*, pp. 547–559, Springer, 2022

Artificial Intelligence Use in e-Government Services: A Systematic Interdisciplinary Literature Review

Richard Michael Dreyling III(✉), Tanel Tammet, and Ingrid Pappel

Tallinn University of Technology, Ehitajate Tee 5, 19086 Tallinn, Estonia
richard.iii@taltech.ee

Abstract. The objective of this paper is to conduct an interdisciplinary systematic literature review of the current state of the art related to the use of Artificial Intelligence in the field of e-Government services that includes technical applications. The study uses the systematic literature review methodology prescribed for software science. Of over 500 resulting articles, the final relevant number of articles is 29. The results include a large cross-section of disciplinary approaches. One surprise result is that even technical articles considered the ramifications of the use of AI in government services on underserved populations. The field of use of AI in government services for service provision is still a new area of investigation and more literature is being published constantly. Because of this, a recommendation for potential areas of future research include readiness assessment frameworks and security.

Keywords: Artificial intelligence · Government services · Literature review

1 Introduction

Governments have begun to use AI in various ways to increase effectiveness and efficiency [1] Different governments and organizations have found AI useful in the provision of government services, or in providing governance.

Artificial intelligence has the potential to be a transformative technology but it also comes with potential drawbacks. As governments insist on learning how to use AI, the potential for it to have a negative effect on citizens' lives must be considered. Issues like bias that can be the result of historical data or hard coded into the algorithms can have a larger impact on those sections of societies least able to defend themselves. This can include bias in policing, or even credit decisions [2].

The systematic literature reviews previously conducted research to answer some of the important questions about AI and its uses in government, public governance, sustainable development and business models [1, 3, 4]. However, the unique situation in which the social effects of AI can depend on less obvious technological aspects it is necessary to conduct a holistic literature review that combines all of the interdisciplinary parts of the research to understand how the use of AI in government services can be

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
T. K. Dang et al. (Eds.): FDSE 2022, CCIS 1688, pp. 547–559, 2022.
https://doi.org/10.1007/978-981-19-8069-5_37

achieved without causing harm to the populace through bias, legal issues, or breaches of privacy [2].

An understanding of technical topics as they apply to the problem of the application of AI to the provision of government services is largely absent in the literature. The literature review of AI use in public governance conducted by Zuiderwijk et al. [2019] gives a thorough format to approach the research questions [1]. However, this literature review disqualified all technical literature and the inclusion of the technological element has the potential to provide interesting insights in the area of using AI for government services.

The objective of this literature review is to conduct an interdisciplinary systematic literature review to understand the field as it concerns technologies involved as well as the uses of AI in public governance as it applies to e-government services. As presented in Burgers et al. [2019] an interdisciplinary approach will be followed to bring together disparate perspectives on the same problem to help better understand the application at hand [5].

The methodology of this research modeled Kitchenham's [2009] systematic literature methodology [6]. The primary purpose of this research is to answer the research questions to understand the literature as it stands currently in the field across many disciplines.

Research questions:

1. How has the use of AI in public services been researched?
What disciplines and approaches are prevalent in the literature?

Main contribution:

This paper gives results as to research items specifically relevant to the topic of the use of AI in public services.

2 Methodology

The methodology of this research modeled Kitchenham's [2009] systematic literature methodology [6]. The primary purpose of this research is to answer the research questions to understand the literature as it stands currently in the field across many disciplines. As such, quite broad research questions were selected to give an indication of the existing literature.

2.1 Keywords and Search Process

The search process included the databases Web of Science and Scopus. These were chosen because they have a large number of disciplines and are known as general research databases that give an accurate picture of the literature that is currently available. The general databases were in this case chosen to limit the amount of bias and conduct the review in as systematic a manner as possible. Initially, a search which included all of the potential keywords separated by "OR" returned results of over 10,000 between the two databases. After this initial search adjustments were made using the top results as feedback to try to hone the results to be more relevant to the research questions. The final selection of keywords was placed for the Scopus query as follows: "(TITLE-ABS-KEY ("Artificial Intelligence" OR "AI") AND TITLE-ABS-KEY ("Public Sector" OR "Government") AND TITLE-ABS-KEY ("Public Service" OR "Government Service" OR "EGovernance" OR "E-Governance" OR "E-Government" OR "Egovernment")).” A search for a similar keyword string and Boolean values was placed in Web of Science. However, the search criteria in Web of Science were “All Fields” as the “TITLE-ABS-KEY” which search titles, abstracts and keywords were not available. These keywords were chosen because they would accurately narrow the specific items that would apply to the research questions through the boolean requirements. The “ORs” selected in the keyword search were designed to cover all potential analogous terms. With the terms, “AI” and “Artificial Intelligence” the idea was that the keywords would cue on the general terminology even if the article discussed a more specific component technology of AI, such as machine learning (ML) or natural language processing (NLP). The Scopus search yielded 566 results. The Web of Science search returned 321 records.

2.2 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were chosen to be able to get an idea of multiple disciplines as they apply to the research questions. Because of the interdisciplinary nature of the topic, the discipline of the article was not used as a refining technique, to avoid unintentionally disqualifying articles pertinent to the topic. All languages were considered, as well as unpublished materials to attempt to ameliorate English language and publication bias. However, the majority of the articles returned were in the English language.

Inclusion Criteria

- All disciplines.
- Topic relevant to research questions.
- AI used for e-government or public services.
- AI use or application in Public Services as a part of the purpose of the research.
- Consideration of the application of AI in public services discussed.

Exclusion Criteria

- Editorials, letters, book chapters, presentations are not included.
- AI is only part of research methodology not the subject being researched.
- Pure technical article without explicit application to a government or e-service.
- AI is not directly part of providing an e-service or government service.
- AI mention is incidental.
- Early non-AI chatbots.

Duplicates were removed. 384 records were present after duplicate removal. Then the researcher conducted an initial review of titles and abstracts to determine which of these were relevant to the topic. All of the search results were given values of yes, no, and maybe.

2.3 Quality Evaluation

After the initial review of abstracts and titles, all relevant and possibly relevant articles were read to determine a final judgment of relevance and quality. For each of the articles, a quality evaluation was conducted based on the table format given in [7]. Table two gives metrics for rating empirical studies based on the type of data collection. And Table one takes into account table two's score as well as the relevant categories for the quality assessment. These tables give an indication of how complete a piece of research could be based upon metrics like the amount of data collected depending upon the method of collection. For example, different numbers are required when doing face to face qualitative interviews than when considering questionnaires. From these table based metrics, the author derived a percentage quality score by evaluating each point. Due to not all studies having scores in all of the categories, only relevant ones, the quality scores varied to a large degree. The table used for statistical collation and analysis of the quality numbers primarily used the percentage to try to judge articles on as level a basis as possible. This percentage was derived by the following equation: (Score / Potential Relevant Score * 100). The number of articles for consideration in the data extraction portion of the review-was 29, specific numbers at each stage will be discussed in section three. The author read the literature that was accepted for the study. Coding took place within the documents looking for connections and congruities from the broad research.

The search conducted yielded 29 articles which were found from a total of 386 records after the initial application of the deduplication. After the title and abstract review n = 66. Upon further review and quality check, the number of articles selected for study was 29.

2.4 Data Collection and Analysis

The author read the literature that was accepted for the study and coded the documents looking for connections and congruities from the research. Once coding was complete, the author analyzed the codes and data for similarities and through lines in the research while paying special attention to the methodology and discipline of the research.

3 Results

The preliminary answers are stated in this chapter.

3.1 Search Results

As stated above, the search conducted yielded 29 articles which were found from a total of 386 records after the initial application of the timeliness criteria and removal of duplication. After the title and abstract review n = 66. Upon further review and quality check, the number was 29.

These items can be seen in alphabetical order in Table 1. The Systematic Identifier "S" with a number is used for clarity so the reader knows which items were selected for the study and which were referenced in other areas. Where applicable, the overall ACM citation number will be used at the end of the sentences as prescribed by citation rules.

3.2 Answers to Research Questions

As previously stated, the research questions are as follows:

1. How has the use of AI in public services been researched?
2. What disciplines and approaches are prevalent in the literature?

Contrary to expectation, the contributions made by the items included in the study were primarily dealing with practical applications of AI in public services. Because the research area is so new, the hypothesized trend was to see a variety of theories which would explain how the use of AI in public services would work. However, even the overviews of the state of the art had the goal primarily of stating the ways in which AI was used in public services currently and could be in the future. This could be because of the particular angle through which the author is finding the gap in research. Several articles [13, 14, 23, 31] analyzed the phenomenon from theoretical perspectives like the Technology adoption model [23], public value theory, problematization, and digital inclusion. One also suggested a maturity model for e-invoicing to build toward. AI capabilities in automated public services [29].

Due to the nature of the inquiry as well as the field itself, it is expected that the relevant items would be multidisciplinary in many ways. The fields which investigate the use of AI in public services are varied. Below, the discipline through which the items were approached are shown. Information research and public administration comprised the largest portion of the accepted papers. Research focusing on the data science and computational intelligence of AI in public services were the other approach that was apparent. The other disciplines had a single item each, including one article that was pure data science (Fig. 1).

This literature review resulted in an interesting categorical distribution of the types of contributions that were made in the field. Because of the interdisciplinary basis of the inquiry many different types of contributions were made, but many of them may be organized into categories based on the ending contribution. For example, there are multiple overviews of how AI can be implemented in government services including [8–10, 17, 20, 36]. In addition, there are multiple different overviews of AI implementations that have specific geographical regions as a component of the research. These are [22, 23, 25, 28]. These two categories, comprise the majority of the articles in related categories in the study. Not all of these articles are from the same discipline. In addition, some other articles gave an overview of an implementation of a technology or technologies that use AI in public services in the context of a specific region or geography [11, 22, 30, 31, 33]. Only one article [16] specifically measured citizen opinion about using AI systems in public services, which means that only one directly related to user experience or user

Table 1. Texts selected for study

Identifier	Authors	Title
[8]	Ahn M.J., Chen Y.-C	Artificial intelligence in government: Potentials, challenges, and the future
[9]	Akkaya, C; Krcmar, H	Potential Use of Digital Assistants by Governments for Citizen Services: The Case of Germany
[10]	Al-Mushayt O.S	Automating E-Government Services with Artificial Intelligence
[11]	Anwer, MA; Shareef, SA; Ali, AM	Smart Traffic Incident Reporting System in e-Government
[12]	Balta D., Kuhn P., Sellami M., Kulus D., Lieven C., Krcmar H	How to Streamline AI Application in Government? A Case Study on Citizen Participation in Germany
[13]	Chatterjee S., Khorana S., Kizgin H	Harnessing the Potential of Artificial Intelligence to Foster Citizens' Satisfaction: An empirical study on India
[14]	Chen, T; Ran, LY; Gao, X	AI innovation for advancing public service: The case of China's first Administrative Approval Bureau
[15]	Dreyling R., Iii, Jackson E., Pappel I	Cyber security risk analysis for a virtual assistant G2C digital service using FAIR model
[16]	Drobotowicz K., Kauppinen M., Kujala S	Trustworthy AI Services in the Public Sector: What Are Citizens Saying About It?
[17]	Engin Z., Treleaven P	Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies
[18]	Fatima S., Desouza K.C., Buck C., Fielt E	Business model canvas to create and capture AI-enabled public value
[19]	Gunaratne H., Pappel I	Enhancement of the e-Invoicing Systems by Increasing the Efficiency of Workflows via Disruptive Technologies
[20]	Henman P	Improving public services using artificial intelligence: possibilities, pitfalls, governance
[21]	Hong, S; Kim, Y; Park, J	Big data and smat (sic.) city planning: The case of Owl Bus in Seoul

(*continued*)

Table 1. (*continued*)

Identifier	Authors	Title
[22]	Kuziemski M., Misuraca G	AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings
[23]	Marri A.A., Albloosh F., Moussa S., Elmessiry H	Study on the Impact of Artificial Intelligence on Government E-service in Dubai
[24]	Medhane D.V., Sangaiah A.K	PCCA: Position Confidentiality Conserving Algorithm for Content-Protection in e-Governance Services and Applications
[25]	Misuraca G., Van Noordt C., Boukli A	The use of AI in public services: Results from a preliminary mapping across the EU
[26]	Mittal P	A multi-criterion decision analysis based on PCA for analyzing the digital technology skills in the effectiveness of government services
[27]	Montoya L., Rivas P	Government AI Readiness Meta-Analysis for Latin America and the Caribbean
[28]	Nam T	How did Korea use technologies to manage the COVID-19 crisis? A country report
[29]	Pappel I., Gelashvili T., Pappel I	Maturity Model for Automatization of Service Provision and Decision-Making Processes in Municipalities
[30]	Park S., Humphry J	Exclusion by design: intersections of social, digital and data exclusion
[31]	Petersen A.C.M., Cohn M.L., Hildebrandt T.T., Møller N.H	'Thinking problematically' as a resource for AI design in politicised contexts
[32]	Rafail P., Efthimios T	Knowledge Graphs for Public Service Description: The Case of Getting a Passport in Greece
[33]	Snowdon J.L., Robinson B., Staats C., Wolsey K., Sands-Lincoln M., Strasheim T., Brotman D., Keating K., Schnitter E., Jackson G., Kassler W	Empowering Caseworkers to Better Serve the Most Vulnerable with a Cloud-Based Care Management Solution

(*continued*)

Table 1. (*continued*)

Identifier	Authors	Title
[34]	van Noordt C., Misuraca G	New Wine in Old Bottles: Chatbots in Government: Exploring the Transformative Impact of Chatbots in Public Service Delivery
[35]	Van Noordt C., Misuraca G.,	Evaluating the impact of artificial intelligence technologies in public services: Towards an assessment framework
[36]	Wirtz, BW; Weyerer, JC; Geyer, C	Artificial Intelligence and the Public Sector-Applications and Challenges

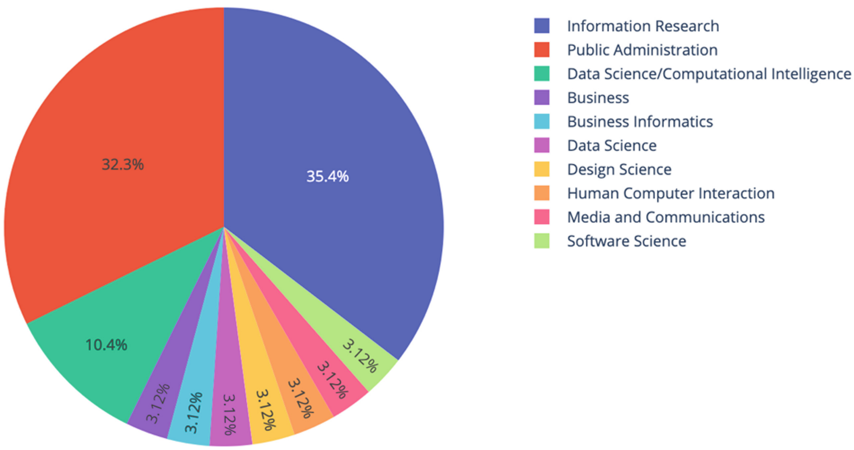

Fig. 1. Discipline of items accepted for the study

interactions and how the public felt. Design of AI techniques and Systems which would implement AI in public services were another large category, with four articles [11, 19, 24, 32]. Theoretical analyses comprised a smaller number of articles in this study, only three articles analyzed AI in public services specifically from a theoretical perspective, or proposed a new theoretical way to look at the phenomenon. One article [18] designed an approach which was essentially a method of completing a business model canvas as a way to aid in the creation of AI enabled government services. [26, 27, 29] proposed discussed readiness levels. [27] specifically explained readiness in a geography. And [29] proposed a maturity model in the context of readiness as a specific geography builds toward the ability to implement a higher level of AI enabled government services (Fig. 2).

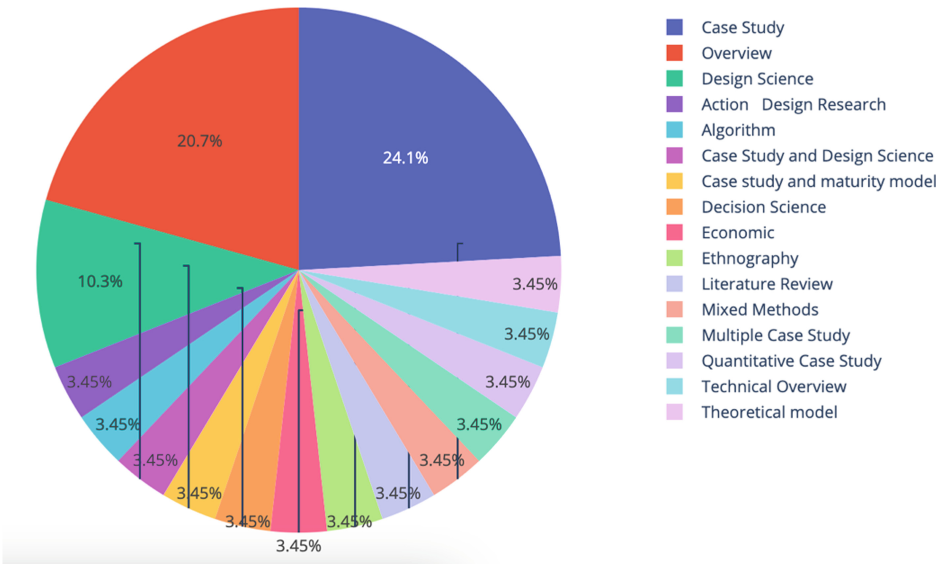

Fig. 2. Methodologies of items accepted for study

The methodologies used in the articles were disparate. The largest percentage of them came in the form of case studies and overviews of the topic. These categories together represented 44.8% of the articles. Only one paper [15] specifically focused on security at all in the area of the use of AI in public services.

4 Discussion and Limitations

This study takes into account the best of the researcher's knowledge at the time of this writing but the process is iterative. The keywords are indicative of the current state of the art as it pertains to the use of artificial intelligence in the provisioning of government services, but with such a new area of study new research will constantly be published. The focus of this paper is a novel area of research even though it is well known that researchers have been publishing research in artificial intelligence for a very long time. This can be considered a limitation of this study. Because it is such a new area of research, the total number of studies that meet the criteria of this piece of research is likely drastically lower because the focus on the use of AI in public services is low. Even though many governments have vision documents referring to how they would like to deploy AI in their governments, in reality very few have gone to the production phase of operation [37]. This logically means that of the case studies that are included in this research many of them are, even if not explicitly stated, primarily derived from cases which have not had real implementation and citizen interaction to gain user feedback or understanding how the technical solution can affect the social environment.. As of the writing of this paper, few projects have been implemented in a production environment that has run long enough to see differences in the before and after state or what the projects have achieved [38].

One theme that came through in the analysis of all the documents is the concept of "can" versus "should." Many of the overviews [2, 35] and previous literature reviews [1] approach the topic from a primarily non-technical disciplinary perspective. In these subjects, one would expect discussions of ethereal topics such as the conceptual, "Given the ability to use a technology to increase efficiency and effectiveness of a public service, should society do this?" This is especially considering the potential and documented exponential downsides for those populations in societies who have historically been underserved. This makes sense to question the ethical ramifications and was in accordance with the hypotheses of the author prior to beginning the study. However, what was not expected was that some of the more technical articles also considered these issues. For example, in the article discussing optimizing the bus lines in South Korea for the late night hours, the authors state:

> "The most profitable late night bus routes generally connect downtown and high income neighborhoods. From an equity point of view, however, the government could not implement the service only in those areas. If the government did implement such a service, low income neighborhoods would be further marginalized and the already unequal regional distribution of public infrastructure would have been exacerbated. This is why governments sometimes make decisions that conflict with evidence guided by the analysis of big data" [21]

This quotation shows that the researchers involved believed that they could optimize the use of the public service of late-night buses. However, they also had the understanding of how this technological improvement would affect the social environment in this area.

In addition, the author did not expect that particular case studies would compare the programs of multiple parts of a specific country, similar at least in the national law, if not the state, and give both the benefits and drawbacks of these programs considering the intent of the programs when they were implemented [30]. These case studies gave insight into the lives of citizens affected by these programs. This understanding is important if politicians and policy makers are going to implement these programs, they should understand in depth the way these programs will affect citizens.

Several areas that lacked a strong breadth of work became apparent during this study. Given that data protection and privacy is an often-mentioned topic in the introduction section of many papers on AI use in public services, it is unexpected that only one of the papers in the study dealt with security [15]. This paper discussed cybersecurity risk. However, there would seem to be a large gap for researchers wanting to research security in this area because an evaluation of risk is only an introductory amount of work that could be done.

Some studied proposed maturity models related to the topic [14, 29]. Only one study [29] suggested the necessity for preambular requirements to build to a successful implementation of AI, in the form of a maturity model for EDRMS systems that would be, at the highest level, able to integrate into an AI system in a public service. Because of the lack of production implementations of AI in public services from which researchers can derive insights, and the pilot-first approach to the use of AI in the public sector, the next step for research would be to establish a framework for AI readiness assessments in the public sector including feasibility studies for pilots. This is common practice in the private sector and could have the effect of more successful AI related public service projects and less waste and loss from projects that currently go to the pilot phase but do not have prerequisite elements for a successful pilot. It would be useful to have more projects that have been completed which use AI in order to discuss with the departments and creators of the projects the outcomes of them and the effect on public services.

5 Conclusion

The many overviews even in the past five years suggest that the science in the area of artificial intelligence implementation in the specific area of public services is not yet established. The future direction of research in this area should do a better job of melding the technical with the social in an interdisciplinary manner and also include the way these technologies will affect the lives of citizens. It is one thing to give specific economic indicators of time saved in measures of efficiency and effectiveness, but the impact on citizens should also be measured. Only one of the studies accepted into this systematic literature review looked at the ramifications of an example of an AI enabled public service and that study [15] did this from an almost theoretically vanilla IT implementation. Perhaps another person can write a maturity model of the literature research into the use of AI in public services. Ideally, public policy makers and politicians should have academically cogent research that goes as far as quantitatively discussing the impact of the issues that arise when AI is implemented in the public sector. This could be accomplished through analyzing the ramifications of things like black box decision-making in AI determination of credit worthiness and other items that should be but are not often regulated. If it can be shown that an AI system will cause a percentage of the population to lose a certain amount of economic potential in currency, then politicians will have a better understanding of how regulation or implementation will impact the citizenry. If systems currently do not have the ability to explain the methods through which they decide who should and should not have access to benefits or services, the considerations of accountability of decision makers and systematic decision making are very important.

Acknowledgements. This work in the project "ICT programme" was supported by the European Union through European Social Fund.

References

1. Zuiderwijk, A., Chen, Y., Salem, F.: Implications of the use of artificial intelligence in public governance: a systematic literature review and a research agenda. Government Information Quarterly **38**(3), 101577 (2021)
2. Wirtz, B., Weyerer, J.C., Sturm, B.J.: The dark sides of artificial intelligence: An integrated AI governance framework for public administration. Int. J. Public Adm. **43**(9), 818–829 (2020)
3. Di Vaio, A., Palladino, R., Hassan, R., Escobar, O.: Artificial intelligence and business models in the sustainable development goals perspective: a systematic literature review. J. Bus. Res. **121**, 283–314 (2020)

4. de Sousa, W.G., de Melo, E.R.P., Bermejo, P.H.D.S., Farias, R.A.S., Gomes, A.O.: How and where is artificial intelligence in the public sector going? A literature review and research agenda. Government Information Quarterly **36**(4), 101392 (2019)
5. Burgers, C., Brugman, B.C., Boeynaems, A.: Systematic literature reviews: four applications for interdisciplinary research. J. Pragmat. **145**, 102–109 (2019)
6. Kitchenham, B.O., Brereton, P., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering–a systematic literature review. Inf. Softw. Technol. **51**(1), 7–15 (2009)
7. Beecham, S., Baddoo, N., Hall, T., Robinson, H., Sharp, H.: Motivation in software engineering: a systematic literature review. Inf. Softw. Technol. **50**(9–10), 860–878 (2008)
8. Ahn, M.J., Chen, Y.C.: Artificial intelligence in government: potentials, challenges, and the future. In: The 21st Annual International Conference on Digital Government Research, pp. 243–252 (2020)
9. Akkaya, C., Krcmar, H.: Potential use of digital assistants by governments for citizen services: The case of Germany. In: Proceedings of the 20th Annual International Conference on Digital Government Research pp. 81–90 (2019)
10. Al-Mushayt, O.S.: Automating E-government services with artificial intelligence. IEEE Access **7**, 146821–146829 (2019)
11. Anwer, M.A., Shareef, S.A., Ali, A.M.: October. Smart traffic incident reporting system in e-government. In: Proceedings of the ECIAIR 2019 European Conference on the Impact of Artificial Intelligence and Robotic (2019)
12. Balta, D., Kuhn, P., Sellami, M., Kulus, D., Lieven, C., Krcmar, H.: How to streamline AI application in government? A case study on citizen participation in Germany. In: International Conference on Electronic Government, pp. 233–247. Springer, Cham (2019)
13. Chatterjee, S., Khorana, S., Kizgin, H.: Harnessing the Potential of Artificial Intelligence to Foster Citizens' Satisfaction: An empirical study on India. Government information quarterly 101621 (2021)
14. Chen, T., Ran, L., Gao, X.: AI innovation for advancing public service: The case of China's first Administrative Approval Bureau. In: Proceedings of the 20th Annual International Conference on Digital Government Research, pp. 100–108 (2019)
15. Dreyling, R., Jackson, E., Pappel, I.: Cyber Security Risk Analysis for a Virtual Assistant G2C Digital Service Using FAIR Model. In: 2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG), pp. 33–40. IEEE (2021)
16. Drobotowicz, K., Kauppinen, M., Kujala, S.: Trustworthy AI Services in the Public Sector: What Are Citizens Saying About It?. In: International Working Conference on Requirements Engineering: Foundation for Software Quality, pp. 99–115. Springer, Cham (2021)
17. Engin, Z., Philip Treleaven, P.: Algorithmic government: automating public services and supporting civil servants in using data science technologies. The Computer Journal **62**(3), 448–460 (2019)
18. Fatima, S., Desouza, K., Buck, C., Fielt, E.: Business model canvas to create and capture AI-enabled public value. In: Proceedings of the 54th Hawaii International Conference on System Sciences, p. 2317 (2021)
19. Gunaratne, H., Pappel, I.: Enhancement of the e-invoicing systems by increasing the efficiency of workflows via disruptive technologies. In: International Conference on Electronic Governance and Open Society: Challenges in Eurasia, pp. 60–74. Springer, Cham (2020)
20. Henman, P.: Improving public services using artificial intelligence: possibilities, pitfalls, governance. Asia Pacific Journal of Public Administration **42**(4), 209–221 (2020)
21. Hong, S., Kim, Y., Park, J.: Big data and smat city planning: The case of Owl Bus in Seoul. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 4492–4500. IEEE (2018)

22. Kuziemski, M., Misuraca, G.: AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. Telecommunications policy **44**(6), 101976 (2020)
23. Al Marri, A., Albloosh, F., Moussa, S., Elmessiry, H.: Study on the impact of artificial intelligence on government E-service in Dubai. In: 2019 International Conference on Digitization (ICD), pp. 153–159. IEEE (2019)
24. Medhane, D.V., Sangaiah, A.K.: PCCA: Position confidentiality conserving algorithm for content-protection in e-governance services and applications. IEEE Transactions on Emerging Topics in Computational Intelligence **2**(3), 194–203 (2018)
25. Misuraca, G., van Noordt, C., Boukli, A.: The use of AI in public services: results from a preliminary mapping across the EU. In: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, pp. 90–99 (2020)
26. Mittal, P.: A multi-criterion decision analysis based on PCA for analyzing the digital technology skills in the effectiveness of government services. In: 2020 International Conference on Decision Aid Sciences and Application (DASA), pp. 490–494. IEEE (2020)
27. Montoya, L., Rivas, P.: Government AI readiness meta-analysis for Latin America and The Caribbean. In: 2019 IEEE International Symposium on Technology and Society (ISTAS), pp. 1–8. IEEE (2019)
28. Nam, T.: How did Korea use technologies to manage the COVID-19 crisis? A country report.". International Review of Public Administration **25**(4), 225–242 (2020)
29. Pappel, I., Gelashvili, T., and Pappel, I.: Maturity Model for Automatization of Service Provision and Decision-making Processes in Municipalities. In: Proceedings of Sixth International Congress on Information and Communication Technology, pp. 399–409. Springer, Singapore (2022)
30. Park, S., Humphry, J.: Exclusion by design: intersections of social, digital and data exclusion. Inf. Commun. Soc. **22**(7), 934–953 (2019)
31. Petersen, A.C., Cohn, M.L., Hildebrandt, T., Møller, N.H.: 'Thinking Problematically' as a Resource for AI Design in Politicised Contexts. In: CHItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter, pp. 1–8 (2021)
32. Petersen, A.C., Cohn, M.L., Hildebrandt, T., Møller, N.H.: 'Thinking Problematically' as a Resource for AI Design in Politicised Contexts. In: CHItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter, pp. 1–8 (2021)
33. Snowdon, J.L., Robinson, B., Staats, C., Wolsey, K., Sands-Lincoln, M., Strasheim, T., Kassler, W.: Empowering caseworkers to better serve the most vulnerable with a cloud-based care management solution. Applied Clinical Informatics **11**(04), 617–621 (2020)
34. Noordt, C.V., Misuraca, G.: New wine in old bottles: Chatbots in government. In: International Conference on Electronic Participation, pp. 49–59. Springer, Cham (2019)
35. van Noordt, C., Misuraca, G.: Evaluating the impact of artificial intelligence technologies in public services: towards an assessment framework. In: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, pp. 8–16 (2020)
36. Wirtz, B.W., Weyerer, J.C., Geyer, C.: Artificial intelligence and the public sector—applications and challenges. Int. J. Public Adm. **42**(7), 596–615 (2019)
37. Misuraca, G., Van Noordt, C.: AI Watch-Artificial Intelligence in public services: Overview of the use and impact of AI in public services in the EU. JRC Working Papers, (JRC120399) (2020)
38. van Noordt, C., Misuraca, G.: Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union. Government Information Quarterly, 101714 (2022)

Publication V

R. Dreyling, K. McBride, T. Tammet, and I. Pappel, "Navigating the AI maze: Lessons from Estonia's Bürokratt on public sector AI digital transformation," *SSRN*, 2024

Navigating the AI Maze: Lessons from Estonia's Bürokratt on Public Sector AI Digital Transformation

Richard Dreyling[a], Keegan McBride[b], Tanel Tammet[c], Ingrid Pappel[a]

[a]*NextGen Digital State Research Group, Tallinn University of Technology, Akadeemia tee 15A, Tallinn, 12169, Estonia*
[b]*Oxford Internet Institute, 1 St Giles', Oxford, OX1 3JS, United Kingdom*
[c]*Applied Artificial Intelligence Group, Tallinn University of Technology, Akadeemia tee 15A, Tallinn, 12169, Estonia*

Abstract

This study delves into the complexities of adopting Artificial Intelligence (AI) in public services, with a focused case analysis of Estonia's pioneering Bürokratt initiative. As governments worldwide grapple with the integration of AI into public sector service delivery, Estonia's experience offers critical insights. This qualitative case study, grounded in the Technology-Organizational-Environmental (TOE) Framework and AI governance principles, investigates two primary research questions: What factors influence the technical and organizational architecture for implementing AI-enabled e-services in public sector organizations? And how can these organizations leverage digital transformation to facilitate the adoption of AI-enabled digital public services and enhance service delivery? This explanatory case study uses semi-structured interviews, document analysis, and thematic analysis. A significant finding is the pivotal role of organizational flexibility, open collaboration, and citizen-centric approaches in navigating the AI implementation landscape. The study contributes to the field of digital transformation by highlighting the recursive relationship between technology adoption and organizational structure, especially in the context of AI-enabled public services. Our findings underscore the potential of small nations like Estonia in setting global benchmarks for AI in digital public services.

Keywords: keyword one, keyword two
PACS: 0000, 1111
2000 MSC: 0000, 1111

1. Introduction

Artificial intelligence is everywhere, from behind the scenes of the most used search engines in the world to the recommendation engines used by social media companies that claim over half the world as users. Even with recent headlines being grabbed by the quick adoption of generative AI tools, the direction of the use of AI in the area of public services does not have an answer (van Noordt and Misuraca, 2022). Implementing AI is by its nature more complex than the types of digital transformations that comprise the majority of projects (Jöhnk et al., 2021). Due to the difficulty of digital transformations, public administrations should strive to understand the best practices of the DT process as it relates to the organizational and technical transformations required for adopting AI for external services.

Digital Transformations are a relatively new field of study, and although there are conceptual frameworks in place (Vial, 2019; Zaoui and Souissi, 2020; Mergel et al., 2019) the literature continues to expand as researchers add individual contributions. A specific type of Digital Transformation literature tends to focus on the factors that can lead to success or failure in the process of adopting a new technology within organizations (Breaugh et al., 2023). One reason for organizations to move toward digital transformations in the context of the public sector is to react to the contextual change in their environment(Mergel et al., 2019; Vial, 2019).

Adopting artificial intelligence in the context of any organization adds a level of complexity due to the nature of the technology (Jöhnk et al., 2021). When adopting a new technology for an enterprise a team would need to consider an already complex array of factors including but not limited to: organizational structure, and the internal and external capabilities of the people for the technology, processes, IT infrastructure, current technological implementations, development plans, implementation, support, maintenance. Now teams and management must consider all of the above items along with the social, legal, and technical challenges posed by AI (Dreyling et al., 2021). AI maturity models attempt to measure how prepared an organization may be able to reckon with these challenges based on certain factors (Sadiq et al., 2021).

Regardless of these challenges, many public administrations are investigating, testing, and beginning the adoption of AI because of its vast potential for increasing effectiveness and efficiency both in internal business processes as well external facing services (van Noordt and Misuraca, 2022). Competence in what could be called traditional digital transformations or a high ranking on AI maturity models do not predict success in AI transformations but it could potentially aid in certain aspects like the existence of data.

This qualitative case study research seeks to find the answer to following research questions:

RQ1: What factors influence the technical and organizational architecture of public sector organizations implementing AI enabled e-services?"

RQ2: How can public sector organizations targeting a digital transformation facilitate the adoption of AI-enabled services?

This paper seeks to add to the academic literature on the topic of AI use in public services by analyzing the development and implementation of the Estonian Bürokratt initiative as well as the organizational structure and dynamics of the teams involved through the lenses of Digital Transformation, Artificial Intelligence Maturity Models, and AI governance.

Section 2 describes the state of the art and theoretical background. Section 3 states the research methodology. Section 4 decribes the case. Section 5 presents and the results. Section 6 .... Section 8 .... The paper finishes with a conclusion in Section 9.

2. Background

Across time organizations have struggled to introduce the latest technologies to their organizations. This has remained true across the technoeconomic paradigms that comprise the recent history of the world. Digital transformation is a research topic that tries to reckon with this difficult topic and is particularly important to consider when looking at the public service adoption of AI technologies.

2.1. Digital transformation

One concept of Digital Transformation that leads to its name is an acknowledgement that implementing a new technology tends to have an effect on not only the technological systems of the organization but also on the structure of the organization itself. This can be thought of as a feedback loop.

The literature on digital transformation seems to be trying to understand the many ways in which the organization is affected by the initiatives that require wholesales changes in the organization as well as what factors can lead to success in these endeavors. Although there are differences in the type of the organization including sizes, motives, and sectors a failed digital transformation project means the disappointment of stakeholders and clients (Randall & Alter, 2019). If one takes on assumption that is external disruptions like macroeconomic changes lor expectations of consumers (Vial, 2019) that cause the need for a digital transformation project then the failure of a project like this could mean being left behind in the marketplace or angry citizens.

Digital transformation goes further than digitalization and digitization of services (Mergel et al., 2019) which is a narrower focused shift in an organization. The digitization process implicitly means there will be changes in the organization's business processes, not taking an analog process with analog procedures and putting it to a digital format. Digital transformation goes further and includes changes to the way the organization's structure and operations. This distinction may seem simplistic but is important to consider. Having to print, sign and email a form back is much different than being able to as a part of the workflow, digitally sign in a secure way to complete a process. If one were to consider the back end processes even as a thought process, it is clear that the nature of the interaction between the person receiving the service and the entity providing the service would greatly change based on whether the organization digitalized or truly digitized a service.

According to (Vial, 2019) once there is a disruption in the environment, it triggers strategic responses which include digital business and digital transformation strategies. These rely on the use of digital technologies. This further fuels disruptions and enables changes in value creation paths. Organizational barriers and structural changes affect the value creation path changes. And this then creates negative impacts to things like security and privacy and potential positive impacts in a variety of areas as well. Understanding the continuous feedback loop model of digital transformation of (Mergel et al., 2019) is key to understanding how digital transformations can have a recursive effect on the organization. Although it has not been studied, one hypotheses is that AI has the opportunity to have an exponential effect on organizations and society because of the self learning nature of most machine learning algorithms.

2.2. Technology adoption and acceptance within organizations and institutions

An entire area of information systems research studies how technologies spread and how to predict the odds that a potential user of a technology will choose to actively use it. This field of study has its roots in the social psychological sphere of research beginning with the theory of planned behavior expanding on the theory of reasoned action (Ajzen, 1991). An entire series of theories that follow attempt to explain whether and why someone will use a technology. These include the technology acceptance model (TAM), its multiple extensions, as well as the unified theory of acceptance and use of technology (UTAUT). Even the more advanced versions of these are not particularly excellent compared to a coin flip at predicting the use of technology by individuals (Venkatesh et al., 2003) but they can help understand what factors may influence the reasons that people may decide to adopt a technology.

Because AI is still a relatively new technology to be offered to individual users without the superuser abilities necessary until this point to use it the average user is still reckoning with the understanding and trust of AI systems. Recent releases of large language models and other generative AIs are having the effect of spreading awareness. However, outside what one might consider entertainment or experimental purposes of these models and software systems, it is not clear the comfort level of average consumers to adopt a system that uses AI.

There are multiple reasons that people could choose to use a new technology. Often they include things like the perceived usefulness, and ease of use (Davis, 1989). One other key point is trust. This level of trust, a key antecedent in the literature regarding technology adoption in e-government services (Carter and Bélanger, 2005) may inhibit AI enabled e-government services from making the leap from the early adopters to the early majority. This is commonly considered a key decision point in the success of disruptive technologies (Moore, 2014). Carter and Bélanger (2005) divided trust into a two separate factors including trust of the internet and trust of government.

Although many factors contribute to the adoption of technology, as can be seen in the variety of frameworks which have attempted to study this phenomenon because this paper discusses the specific use of AI for government services, it is necessary to note that trust is a key factor. Trust is a key factor for e-government use and adoption of these services by ministries and people. And AI adds a layer of complexity to this when compared to other technologies. By association with the self-learning definition of AI implementations, many people feel slightly differently about trusting AI as they do about technologies in general. This can be seen in the application of attachment style specifically related to AI and trust by humans (Gillath et al., 2021). Based on their attachment styles and priming toward a specific attachment style, people will have varying trust in AI. The problem of lack of transparency within AI algorithms and how they make decisions can cause people to not trust AI (von Eschenbach, 2021). In addition, it should be considered that when it comes to AI there are many factors in which an AI revolution has the potential to affect modern societies (Wirtz et al., 2022). It may complicate people's willingness to trust machines to conduct services or other transactions when the same technology may have a negative impact in another sphere in one's life.

The technology adoption and acceptance theories are meant to predict the adoption of a technology, and for a digital transformation there is a necessity for feedback loops, i.e. adoption. Then, it is reasonable as a first step to attempt to understand adoption antecedents through this lens. The organizational and institutional architectures within governments are set in many cases by law (Draheim et al., 2021) and have complicating factors outside of those covered in technology adoption and acceptance models when it comes to digital transformations and large scale changes (Mintzberg, 1989) (Hinings et al., 2018).

2.3. AI use in government and public services

As adoption of AI technologies across the citizenry continues in the spheres of life outside of their governmental relations and obligations, there seems to be a top-down pressure to adopt AI in anticipation of the external demand that will come. This can be seen in the many national AI strategies being adopted, and in Europe in particular, in the striving to pass regulatory measures to set the environment (Schiff, 2022). Although different countries write their strategies based on different assumptions and views of the future (Bareis and Katzenbach, 2022) much of the effects of AI adoption in the area government services is at this point comes from a similar ethereal theoretical place (van Noordt and Misuraca, 2022) rather than measurements that show tangible metrics in most cases.

The many different narrow AIs that comprise the current types of popular algorithms have a huge number of potential uses. Governments see an opportunity to make internal processes and in some cases external citizen facing functions more efficient and effective. The intent seems to be to save money on the cost of provisioning these functions. An example of making a process more efficient is that in the United States, some government departments have used text analysis techniques to organize judicial dockets to increase efficiency by saving judges from context switching (Engstrom et al., 2020). This is a case that should bring about less ethical concerns than some uses.

Across Europe governments have created over 250 use cases of AI (van Noordt and Misuraca, 2022). The typology of these uses cases ranges from AI-Empowered knowledge management to Security analytics and threat intelligence (van Noordt and Misuraca, 2022). Of these 115 of the 250 applications the authors categorized as "Public Service Delivery" related governance functions. Many uses for AI that help streamline what would be called "back-office" processes in the business world indicate the potential for AI to help build efficiency and effectiveness in the public sector without directly being involved in the customer-facing element of services. The customer facing element is also germane to the discussion of trying to automate or create proactive virtual assistant enabled services which the Estonian state has outlined in the vision papers. The typology with the greatest number of public service delivery cases and the highest ratio within the category was "Chatbots, Intelligent Digital Assistants, Virtual Agents and recommendation systems." This means that these citizen-facing systems, which represent roughly one-fifth of the AI uses cases of the total cases studied across Europe are a popular approach compared to the other classification categories; "policy making" and "internal management functions" (van Noordt and Misuraca, 2022).

The future of use of AI in the provisioning of public services is not yet fully explored. Predictions include proactive services that are fully automated (Kuziemski and Misuraca, 2020). The canvas conducted by van Noordt and Misuraca (2022) did not find specific instances of public services being fully automated or proactive.

2.4. Readiness, prerequisites, and capacities

When considering a digital transformation, or any type of change in an organization, one must ensure that the right prerequisites are present to give a better chance of success. One of these prerequisites is leadership. Previous research by Breaugh et al. (2023), shows the need for leadership. Not only must there be support in the organization from leaders at a high level, but there must be specific types of leadership. Having a leader who can provide a vision at the beginning of project – what would fall under the "strategic responses" portion of the Vial model (Vial, 2019) – as well as the ability to change gears and provide more flexible leadership that was easily able to navigate across functional groups and teams later in the project – would give an increased chance of success.

In many ways digital transformation processes and projects by their nature cross functional groups and organizational siloes. In many cases the projects may even include stakeholders from completely different organizations or even sectors. The more people added to a project, the more complexity comes with it. This is due to the way that groups and human beings function. When given different incentives it can be difficult for members of different groups to understand each other and function in a cohesive manner (Solheim-Kile and Wald, 2019). To illustrate this consider the different incentive structures of a private sector company implementing an IT project for a public sector organization. The private sector company's primary objective is to maximize profit. Rather than being a static process from which leaders can set a strategy and allow it to be followed, a digital transformation has inherent feedback loops. For example in the previous example, the "institutional arrangements" including "motivation and pressure" and have an effect on the organizational forms (Gong et al., 2020). Such a complex and fast moving situation demands flexibility in the overall structure of the enterprise in question (Gong et al., 2020).

Some of the issues faced by organizations attempting a digital transformation have not only to do with organizational structure and composition but other factors such as having the right people with the required competencies in the organization (Blanka et al., 2022). A digital transformation, as it is the logical extension of the process of digitalization, and digitization brought to the point of having massive changes across the organization and modifications in the "boundaries" of the organization to take advantage of new technologies (Vial, 2019; Holmström, 2022). As with such a large change in any organization it is important to have competence in the relevant areas to give a higher probability of success.

In the case of an AI related technology, the digital transformation made even more difficult because of the nature of AI systems. Adopting a new technology that requires the adjustment of one thing like an ERP system has its own complexities in changing the way those in the organization use the technology, ensure that they know how to use the technology, and the communication patters around the use of that technology. The ERP system has to be implemented and to have integrations made with the existing technologies in the organization. In some cases, databases and other existing components necessary to the function must be migrated. But when one decides to implement an AI system in a public sector organization, a whole host of complexities come along with it. Instead of having to focus only on the the factors above and how the technology works inside the organization, the implementation team must consider factors relating to the citizen perception of use of algorithms and data for the conduct of government services.

2.5. Bias and ethics

There are plenty of challenges that come from the need to maintain effective yet non-biased use of AI in the public sector. It goes without saying that governments use of AI should not be biased. But bias is of growing concern among researchers (Engstrom and Ho, 2021; Wirtz et al., 2022). Bias can be a result of the algorithm (Engstrom and Ho, 2020) or potentially result from the use of historical data (Engstrom et al., 2020). This means that in the implementation phase of an AI project, both of these should be considered and controlled for. Some in the EU even advocate for continual auditing of adopted systems to ensure that the system is not biased and behaving in the way that it was created to (van Noordt and Misuraca, 2022).

One way to accomplish this, recommended by Das et al. (2023) is to ensure a robust auditing process through the development, implementation, and production phases of AI use. Ethics in the case of the use of AI in public services is important. However, the methods in which AI is typically used currently in the access of public services is simply another avenue through which a person can complete or gain information about the service. Given that the nature of the service is subject to the ability of humans living in a country, state, or municipality to successfully complete a given task in their interactions with the government the ideal is that the all eligible citizens or residents should be able to complete a task. In the situation in which a country, state, or municipality has implemented artificial intelligence to make a service more efficient or effective, the ideal is that the service will be accessible and useable by the same eligible population through a different vector of engagement. This has the large given factor of the ability of the person in question to be able to speak or write in the language in question. This is a form of bias but cannot be reasonably attributed to the Ai system but to the mechanisms within a society.

Although bias in AI and the marginally related topics of ethics and governance are very popular in the literature related to AI use in the public sector (Ballester, 2021), when focusing on the use of AI to enable public services purely through weak AI means like natural language processing (NLP) to interpret intents, the bias that would be more appropriate to investigate would be that of non-native speaker NLP interpretive issues. Automatic speech detection systems in the Netherlands have been shown to recognize the speech of non-native speakers of Dutch at a much less accurate rate than native speakers across all gender and age groups (Feng et al., 2021). Similar other challenges arise in the case of written language in various areas, such as the pretrained neural language models Anoop et al., sentiment analysis of non-native speakers (Blodgett et al., 2020) as well as GPT detection against non-native speakers (Liang et al., 2023).

3. Methodology

The methodology in this piece of research is a qualitative explanatory case study. The case study is the recommended methodology for analysis of a real-world phenomenon with the ability to investigate in-depth in the environment in which the subject of the case is occurring (Yin, 2017). This allows the researchers to engage with the organization conducting the project and learn without affecting the organization or project implementation.

The research questions start with "what" and "how" which support the use of an explanatory case study (Yin, 2017). Both questions seek to understand the mechanisms and causal relationships in a specific context. This does not describe a situation or explore a new phenomenon, it seeks to understand how and why certain outcomes are happening in the context of the case. The nature of this inquire examines the feedback loops between technology and

	Interviewee	Participation
#1	Former CIO of Estonia, Managing partner of consulting firm	Interview
#2	Project Manager for AI	Two Interviews, Workshop
#3	Product Manager for Bürokratt	Two Interviews, Workshop
#4	Bürokratt Architect	Two Interviews
#5	Project Manager for Bürokratt Services	Interview, Workshop
#6	AI Researcher and Consultant	Interview
#7	Head of Machine Learning and NLP	Interview
#8	Team Lead – Client Organization	Interview

Table 1: List of Interviewees

organization as well as the influence of historical and local factors and streategies employed in the development and implementation phases of a digital transformation, which fits more appropriately in the explanatory cse study methodology. Explanatory case studies are well suited for situations where the research is examining existing interventions or real-life events to understand the underlying principles or causes. This tstudy aligns with this as it exampines the real-life implementation of AI in Estonian public services, aiming to explain the factors that contribute to its success and strategies employed to navidgate challenges.

Interviews allow for the researcher to gather information when direct observation is not a feasible option (Creswell and Poth, 2016). The semi-structured format provides the ability to have a preplanned interview schedule with the flexibility to delve deeper into topics where necessary for clarity and understanding.

The data collection methods include document review, semi-structured interviews and a workshop. The case study focuses on a particular technology development and adoption process within the scope of a larger digital transformation within the Estonian government. The overall concept of the research design is to attempt to ensure internal and external validity while gaining as many insights as possible regarding the implementation of an AI enabling public service project. Internal validity measures include sharing the interview schedule with the team and gaining insight as to how the questions may be perceived, and to focus on making sure that the proper questions are being asked. External validity measures range from ensuring that the questions of investigation are relevant to be used in other international contexts to documenting all the research steps to ensure repeatability.

The sample selection method was snowball convenience sambpling because the purpose of the research gave a limited number of people who understood the context of the case. Thus the sample is limited to team members of the core team, former members of the team, client organizations, and outside researchers and consultants who had contact with the project in order to gain another perspective. Each interviewee provided potential interviewees with whom the research team could have an interview to gain additional insight into the program. The sampling stopped when no further interviewees were forthcoming and willing to be interviewed. Overall, seventeen queries to potential interviewees resulted in nine interviews and a workshop focused on the needs of the clients and customers of the Bürokratt program.

In addition to the data collected from the interviews publicly available documents, the GitHub archive in which the procurement for the program, the backlog, and other development related documents and code repositories are publicly available, and public comments, were used to triangulate the results to provide validity.

The research team transcribed the interviews using automated software and then edited the transcripts for accuracy. Once this process was completed, the team followed Qualitative coding procedures to label and find themes in the data. Then the team completed a thematic analysis to extract insights from the data.

4. Case background and description

Estonia has the reputation of a leader in the field of digital governance and e-government. The Estonian government adopted digital technologies ahead of some of the older countries for a variety of reasons. In recent times, Estonia was one of the earlier countries to adopt an AI national strategy. As a part of the national strategy, Estonia determined the necessity to ensure development of AI capabilities across the public sector, private sector, academia, and the populace (Sikkut et al., 2020). During this period, the CIO, CDO, and CTO of Estonia penned a vision paper describing what they believed would be the future state of e-Governance as well as a path toward this vision. One of the concepts to which they decided they would build was government services accessible via virtual assistant.

As a path to the development of this vision of a virtual assistant enabled proactive relationship between civilians and government, the Ministry of Economic Affairs and Communication (MKM) proposed a proof of concept for a chatbot that would be a one stop shop for those who wanted to ask questions of different government ministries. After a successful chatbot proof of concept, this ministry began to build a team to develop a production version of the chatbot that could be deployed for wide use.

Estonia's focus on digitization and position in the world gives some factors that can make a digital transformation which uses data some advantages and also challenges. Because Estonia has a long operating digital state, there is a history of innovation. This aids in the ability for ministries to be willing to be comfortable with technology in general. However, this also may create a sense of comfort and unwillingness to move to new options. The data exchange platform, called X-tee in Estonian, or X-Road allows for government ministries to exchange data in a manner that is secure and route. However, the principle that people give their data only to one government body and the data protection law passed by Estonia, mean that data has to be stored where it is captured. For example, any government entity querying a resident's residence would request the information from the population registry. This would leave a log file with a hashed time stamp from which the resident in question can see what organization queried the information and when it was queried. This protection law though, means that it is not legally feasible for the ministries in question to create a data lake through which mass training or queries can be conducted. Because of this, a network of interoperable localized chatbots associated with individual government entities and databases needed to be built rather than a single government chatbot.

This could provide some challenges because each implementation of the chatbot part of Bürokratt will have to be localized with NLP data to be able to adequately answer questions related to the queries citizens may ask of that particular government entity.

5. Results

5.1. Technical and organizational considerations

5.1.1. Blending the old with the new

One thematic throughline with reference to the technical and organizational architecture is a combination of the ideas presented by Christensen (2013) and Mintzberg (1989). Governments, with the exception of newly formed departments (Lunenburg, 2012) typically are what Mintzberg would consider machine organizations. This means that they have an inherently difficult task when faced with change. At the same time, Christensen's concept of the innovator's dilemma outlines the challenge faced by large organizations that are forced to confront new technologies. Christensen's concept is meant to apply to for-profit companies. However, the idea that an organization has to upkeep development and maintenance on their existing services while simultaneously developing a new technology and the challenges this presents is also applicable to governments. Contrary to the suggestions Christensen developed for corporations, it is not possible for government entities to create spin offs to enable the building of AI enabled public services. Thus the only option is to develop "ambidexterity" as "dynamic" organizational capacity (O'Reilly and Tushman, 2008). Principally, in this case it illustrates the tension between existing technologies and new ones.

The organizational and technical infrastructures in this case inherently tied. Both the technical organizational architectures as they relate to an AI enabled public service must be configured in a manner which allows for them to integrate with existing organizational structures and technical architectures.

For example, in the Estonian case, the Bürokratt program has been designed from the beginning to comply with existing technical infrastructures. This is illustrated by interviewee 4 as they state, "In Estonia, we have about 3000, e- services. We are working on a solution that anybody can create e-services by combining these three 3000 e-services plus any number of public REST endpoints "The e-services in Estonia use a date exchange platform called "X-Tee" the Estonian version of X-road as an enabling technology. This has the follow on effect that any application or program attempting to aid in the transaction of e-services is required to be interoperable with X-road, in this case Bürokratt. The history of e-service transactions completed through X-Road provides a set of data from existing architectures that can be used moving forward.

Additionally, the Bürokratt technical architecture has to solve for challenges with X-Road's original design and it's integration with an AI-enabled tool as laid out in the *next generation digital state* vision paper by Vaher (2020) which they solved in a different way than in the pilot of the technology, "we have this the DMR in in our architecture distributed message rooms it's we have a proof of concept for its from June last year we haven't implemented in production yet."

This is further confirmed in discussion with interviewee 5 when describing the approach that Estonia has taken toward digital transformation as it relates to the data exchange platform X-road, "Many countries probably have some similar problems, just because we work in a similar field. But again, because we have solved things, previously very differently, we have X-road. Other countries don't have X-road. This is not to say that because of X-road. If you don't have x-road, you will never have digital transformation, that I don't just, I don't really believe in that. Because you have lived as a country up to this day without x-road. So probably very likely, you have other processes going on." This statement acknowledges the interaction between technology and business processes in oganizations.

5.1.2. Development tied to organizational processes

The development process itself is tied intimately to the business processes in the government organizations it is meant to enable. Interviewee 4 for examples states, "We don't have this kind of discussion with the business side said that we want to provide this kind of service around them if we can do it or we should we have to skip it. Like a product's business requirements. And technically, we provide everything that the business side wants and for our clients, it's just a matter of using our deploy scripts to get everything up and running." Before making any technological decisions, the architect asks for a business case. This can be considered a standard practice. However, in this case it provides the important role of tying the business processes to the technological developments and ideally keeps the focus on the requirements of adopting organizations.

5.1.3. Silo-breaking activities

The aforementioned combination of the need for an organization to address faster moving technologies and the difficulty changing the organizational architecture of a machine like organization necessitates addressing these needs through novel means. The speed at which organizations process information in the information age architecture (Mendelson, 2000) is a determinant factor for success. Given constraints on the structure of the organization, silo breaking activities, including overlays and a knowledge sharing practices are a potential solution. The project manager for AI projects in Estonia explains the nature of their overlay position: "As project manager here, but I'm not directly working together with [the Bürokratt team], because my job is basically to be making sure all the goals that were set in the AI strategy we have for Estonia, that those goals will be met. So my job basically, is to know all the different private public sector organizations, what they're working on what kind of AI projects they have started. And so I should have a database of all the AI projects done in Estonia in the public sector." (Interviewee 2) In this manner the interviewee's role is what can be called an overlay position. This allows for communication across organizational and developmental siloes. Because this individual reports directly to the Chief Data Officer of Estonia, there is a cross-organizational communication path and ability for groups to not duplicate efforts. They state, "Because often when [ministries] start developing something they kind of reinvent the wheel sometimes, because actually, it happens often. So they, they think that this is something completely new that nobody has ever done it before. But then it turns out that somebody else has already done it. So maybe they should talk to each other. So that's that's also one of my responsibilities; to make sure I bring the right people together and when somebody is struggling I can I should be able to tell them where to go and who to interview or who can help them." (Interviewee 2)

In addition the group conducts knowledge building activities among government employees which also allow a venue to communicate across siloes. Interviewee 5 states, "At least once a month, we have webinars where agencies that have implemented something will come and talk about it. And then you have the connections already made. Whereas for Bürokratt we have webinars where you have agencies that have already implemented Bürokratt, but it's more so we are sharing information."

The technical architecture ties into the organizational architecture, including legacy structures of both, when it comes to how the Bürokratt tool is developed. The Bürokratt tool is developed centrally and with instances managed by adopting agencies, like X-Road before it (Blake Jackson et al., 2021). From a legal perspective, the agency or ministry must have new contracts signed between the relevant service providers and service consumers as well as the state authentication system "TARA". Interviewee 5 states, "Tara means fence in Estonia. You need to use TARA [authentication], because Bürokratt's back office uses data for you to log in. But for that you need another contract.

5.1.4. Necessary activities

As can be seen from the example of the state authentication system integration in a legal and technical sense, the technical and organizational architectures are shaped by factors outside these two phenomena. Rather, they are both influenced by the legal framework and regulatory environment. The above example, is a clear situation in which the law affects both the technology and the organization. In addition, the approach to building Bürokratt mirrors in some ways the development X-Road by the Estonian state nearly 20 years ago (Blake Jackson et al., 2021). There is a single responsible team for the development inside of the government, and they use public private partnerships. In the case of Bürokratt, the RIA (State Information Authority) team responsible for building the program conducts the public private partnership procurements in the open on GitHub. Interviewee 4 states, "We produce a software, we use X-road services, we don't provide any central services. We observe our Bürokratt instances." The program is there to provide a chat interface and service module to open another digital channel through which residents are able to instantiate digital public services, operated by other government ministries through existing infrastructure.

As can be expected, when technology is in use in the government, at least in an EU country, regulations apply. The ministry gives only certain portions of the development to outside providers and takes it upon itself to maintain the proper security readiness, Interviewee 4 states: "[Development partners] have to use our base components. Then when it comes to pentesting. When something is red, [they] are not affected. These are our core components, you won't have any access to them. And you just wrote these [simple text files in YAML (yet another markup language)] and that's it. And this is also like a major argument. Everybody says yeah, fine. It's great now, because penetration testing in public sector is not something that's taken lightly."

Data protection and privacy are also important essential activities that tie into the legal and regulatory framework. The development of the distributed messaging rooms (DMR) mentioned above is an example of how the technical architecture is affected by the regulatory framework, in this case the personal data protection law in Estonia. Another example of this is the approach of the MKM team as the approach the continued development of the Bürokratt system in the post ChatGPT general availability world. Interviewee 7 states, "If we want to have citizens specific request inside there the we shouldn't use that way. It's a big risk and uh, we would have a have it risking so that open AI in some point will use retrains their model and uses the data that will given them. And inside there is private information about some citizens." So rather than using the most widely adopted solution, in order to protect the people using the system, it is necessary to consider Gen AI options that will be able to prevent private personal data from being used to retrain the model later. The potential hallucinations and other challenges of GenAI or LLM use in government services is beyond the scope of this paper.

Many factors influence the technical and organizational architectures. These include the historical architectures used by the government and the extent to which they can be changed, as well as the legal and regulatory framework and planned improvements.

5.2. Digital transformation towards AI-enabled public services

Estonia has a large role in consulting with other nations internationally on how to increase the maturity of digital public services. It could then of interest to understand from an Estonian use case how another country or government would go about managing a digital transformation in related to adopting AI-enabled digital public services.

5.2.1. Have a purpose!

"I think they should, I think countries should really actually focus on themselves, have think tanks, really understand. What are our issues, let's really get the root issues down... At this way, it's way more important to understand the problem, and then start looking at the solutions that maybe somenoe esle did." (Interviewee 5)

Having an understanding of what a person wants to build or to write, whether through a prototype or an outline, can ease the difficulty of the building or writing. Similarly, with a governmental digital transformation, rather than looking at something that is already in existence, starting with the use case and then considering the technical and organizational legacy in the region can ensure that whatever is built suits the context in which it is operates.

5.2.2. Execute on the vision!

Many countries have written AI strategies and road maps. Many countries are currently planning, piloting, and executing AI projects in government. However, the journey from the vision to the implementation and operation of the projects can be a fraught one.

Interviewee 1 details how there can be challenges before development starts: So we assumed when we were writing the paper and stuff, for example, 'I don't know linking up to Siri is gonna be like quite an easy thing, right?' What came out in the work is that well, look actually the way that the series the Google Assistant, the others are built up right, they were nowhere near ready to what bureaucrat was meant to offer... So obviously that's took from the original plan a few steps back compared to where we hope to be. The initial vision may have some challenges when it comes to the capabilities of available technology for the use case. And in the context of these challenges, the interviewee explains the importance of experimentation and agility in the approach. "That's not that's exactly like agility of that, that's the that's what happens if you set out with an experimentation program like Bürokratt, I really see that experimentation program like we had this hunch that hurts how it could work, but would it work how it would work?"

This focus on agility for government entities extends to the way that the current development operates. There is debate about how effective agile development is in the public sector Kupi and McBride (2021). Nevertheless, the groups developing the Bürokratt program use several practices typically associated with agile methodologies and Silicon Valley. They use Jira, and have product managers similar to many technology firms. They at one point even instituted mini-procurements at the beginning of the project which allowed for them to have individual partners develop specific back log items in one month long sprints. Each mini procurement would then have a development cycle associated with a one-month agile sprint. However, in keeping with organizational agility (Vial, 2019) when they made a shift when necessary. Interviewee 4 states, "We changed it. There were two reasons. One of it was that we got so many developments so far that we couldn't test it and it was really a problem. And there are some solutions to automate this part. The other thing ... was that if you have so many small tiny procurements, it's totally overwhelming for the legal department and that's why we had to adopt [a] different kind of approach."

The organization also uses techniques associated with technology firms for maintaining alignment and driving achievement. The team uses Obectives and Key Requirements (OKRs) embedded in the Jira platform. Interviewee 2 states, "We use the same system, we have our OKRs. And then, and of course, often, with innovation, your goals might change a bit like at first, they might say, Okay, we'll do this thing, but then you realize that actually, that's not the thing you can do...And let's agree on this other thing we're going to do, and it's okay... I always have my my objectives there. And we know what I need to do." Because OKRs are agreed upon by all relevant parties, it aids in maintaining cohesion and alignment across the organizational. In addition, as stated by the interviewee, they have built in flexibility.

5.3. Adoption challenges

5.3.1. Lack of citizen engagement

Many of the AI-realted initiatives from the public sector are a result of funding availability and top-down strategies. One interviewee states, "So citizens very often are not included in any AI related projects or somewhere at the back you know, verify this for us...They're not a key stakeholder in already the beginning, so in few cases I've seen, they come as an afterthought(Interviewee 6)." This can result in a product or service that is difficult to get citizens to use. Additionally, citizens may not have the same excitement for a government developed system as they do for the latest customer focused private sector application. Interviewee 5 states, "For us, we are selling a national service. So it's not that exciting to many people. And even with the uptake of apps, like Tiktok comes out, and everyone's like, 'Oh, this is so great. You don't really need it, but it's just so great.' [Then] the Covid [infection location tracking] App comes around, [and[no one downloads." This is a challenge that many governments face as they develop citizen facing Government to citizen (G2C) technologies.

5.3.2. Organizational resistance

Resistance is another factor in the challenge to go from the vision to implementation. In the government as well as the private sector, the challenge is that individuals working in the organization may have resistance due to risk aversion. Interviewee 5 illustrates, "The agency wants to not make any mistakes... They have the understanding of how does this affect me? And how do I sell this to my higher ups and not get fired? Obviously not get fired. But you need to have a strong use case that you can back everything up.“ One way to ease this transition is to understand this risk aversion.

And contrary to the prevailing opinion, it is not necessarily the leading technological organizations in the government that will be most willing to adopt an AI related technology. Interviewee 2 states, “they feel like they are already way ahead, and that they are the digitally transformed. And they don’t see the need for AI.”

5.3.3. Overcoming resistance

Looking at the process as an educational one and showing the value to adopting organizations seems to be most successful. “You have to somehow make them realize that there’s so much value in data, and they can actually use it... and that they are missing out on so much. But I guess it’s also like, it’s about educating them and teaching them (Interviewee 2).”

Another important factor that has been identified in the literature is to get the leadership involved (Vial, 2019). Interviewee 5 sees this dynamic and states, “If the agency has an understanding, its Director General, and it’s people that actually are going to implement it. Both have a it’s very important for the Director General also to understand because they will need to give their okay, because it’s, it is gonna take money, to understand what it takes, then they are there for us to actually call them clients.” When leadership is involved and tracking the implementation and adoption, it gives cover for the individuals in the organization and incents them to make the adoption successful.

Even with all these methods to overcome resistance and accomplish the mission, this does not mean that the digital transformation will be successful. It gives a slightly better probability of success. Even then, some find that opening a new digital channel doesn’t necessarily reduce the level of work. Interviewee 8 states, “I think the bot can answer the easier questions and ... now all the ‘I want‘ requests that we get for from for our email. The questions aren’t simple.” It may even increase some of the human work.

6. Discussion

The Bürokratt initiative is being developed and implemented in a unique context. The findings of this research support much of the digital transformation literature in that there is a feedback mechanism between the technology an organization adopts and the organizational structure. In the case of the visionary authority, the Ministry of Economic Affairs and Communication (MKM) and the integral development authority, the State Information Systems Authority (RIA), although largescale wholesale changes have not occurred organizational shifts in the form of cross-functional teams and overlay positions facilitate increased communication and alignment. Interviewee 2 is a project manager who has a view into nearly all AI development projects in Estonia, of which Bürokratt is only one, and can facilitate communication and knowledge sharing between these organizations. This person reports directly to the Chief Data Officer (CDO) of Estonia even though they work in RIA. In addition, interviewee 5, the project manager for Bürokratt services, works at MKM and reports to the CDO as well, but as a part of their job, works very closely with the Bürokratt Product manager, who is interviewee 3. Human resource constraints in the Estonian government cause this to be individual contributors, but in a larger governmental context it would make sense for this to be cross-functional teams and multiple overlay personnel who facilitate moving information between siloes.

Much of the digital transformation theory focuses on digital transformations across private and public entities. The level of flexibility in structuring a private enterprise would be higher than that of a government, which necessarily has many environmental and legal constraints as to how one can change the structure in response to feedback due to a technological change. This would from an institutional perspective pdiscussed by Hinings et al. (2018) mean that in this case the government adopdted digital institutional building blocks, but not digital organizational forms like AirBnB or Uber, leading to digital innovation in the adopting organizations within the Estonian government.

Trust is a factor that is key to the adoption of e-governance and e-government related technologies as discussed above. One way to aid trust is to enhance the user experience so that the people using the service perceive that those developing it are similar (Colesca). The Estonian authorities understand this. They commissioned a consultancy to understand the readiness of the population to adopt AI enabled digital government services and lead with the needs of the citizen as their “why” in their discussions on the topic. This came across in the interviews and the workshops.

The dynamic interaction of the technical and organizational architectures in the case of MKM and RIA could be seein in the technical innovation and flexibility required to be able to integrate AI technologies like natural language processing and the a decision to develop a network of interoperable chatbots. The organizational adaptability and strategy could be seen in the use of agile and collaborative approaches like agile development and the use of objectives and key results to maintain alignment across the organizational structures that typically are heavily bureaucratic. The ongoing alignment of strategic government objectives with the technological advancement and those involved in that development demonstrates this as well. The environmental challenges posed to the government like the challenge of figuring out a way to make the national language of Estonia available in data and operable in the chatbot and regulatory environment provide the teams at RIA and MKM with an opportunity to introduce unique technical solutions to make a vision into reality.

With limited human capital, the teams realized the importance of reducing the complexity from the adopting organization perspective and even the developer perspective. With multiple iterations of the final product, this should aid when the service module reaches general availability and Estonians are able to use Bürokratt to access digital public services.

7. Recommendations for strategy and policymakers

The results of this case study indicate that it is not possible for leaders from foreign nations to be able to copy and paste the Bürokratt program into their context. To quote one of the interviewees, "You look at your problems, and understand your core root issues and find a solution to it. Maybe I should rephrase it as what should the outcome be like? This you should do before you start looking at other countries, or just in my opinion. Then you don't ask a question, how can x-land be like Estonia? But then you look at how can I fix my problems? And I think people would be you know people in public sector or any place would find it easier to start working on this very huge task."

Another clear relevant factor to decisionmakers in the literature and the results of the case study is the importance of upper management support or leadership. Without funding from leadership nothing will get done. But in addition to this it is important that leadership is on board from the ideation phase of the initiative. In the case of the Estonian Bürokratt initiative, the vision originated from the leadership and that driving "why" is responsible for not only the funding and the organizationally agile context, but also for hiring the right people in key positions to be able to make the vision meet reality. This forward looking leadership chose to adopt agile practices adapted for the public sector and emplace processes and practices to maintain strategic and tactical alignment like using objectives and key results rather than the typical key performance indicator way of measuring performance.

8. Limitations and future research

The size and scope of this case study is limited by design. This comes with limitations in the potential applicability of the results. In addition the time scope is limited. This case study represents a small period of time in the development of the Bürokratt initiative so it is unable to determine of the program is a success in achieving the vision of virtual assistant enabled digital public services.

Future work could give deep analysis of similar initiatives in other technological, organizational, and environmental contexts. Relevant factors for AI maturity models and readiness assessments in the public sector would be an interesting topic of research. In addition, with the recent rapid expansion of large language models (LLMs) future research could investigate how similar projects might implement LLMs or foundation models as well as the potential ethical ramifcations of such an implementation and governmental adaptation to their digital transformation strategies.

9. Conclusion

The Bürokratt initiative as a case study demonstrates many of the relevant points of the academic literature related to digital transformation as they pertain to the adoption of AI enabled digital public services. It illustrates the adaptations in technological innovation and organization strategy that the Estonian government team has used to address contextual challenges and maintain a strong human centered focus in service delivery. The importance of leadership and effective public private partnerships in response to limited human capital provides valuable lessons for similar digital transformations in other contexts. Even though this case study captures a point in time early in the development of the project, it shows the potential of strategic AI implementation in providing efficient and effective digital public services.

References

Ajzen, I., 1991. The theory of planned behavior. Organizational Behavior and Human Decision Processes 50, 179–211.

Anoop, K., Manjary P. Gangan, Deepak, P., Lajish, V., . Towards an enhanced understanding of bias in pre-trained neural language models: A survey with special emphasis on affective bias, in: Mathew, J., Santhosh Kumar, G., P., D., Jose, J. (Eds.), Responsible Data Science: Select Proceedings of ICDSE 2021, Springer, Singapore. pp. 13–45. doi:10.1007/978-981-19-4453-6_2.

Ballester, O., 2021. An artificial intelligence definition and classification framework for public sector applications, in: Proceeding of DG.O'2021 – the 22nd Annual International Conference on Digital Government Research, ACM. pp. 67–75.

Bareis, J., Katzenbach, C., 2022. Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. Science, Technology, & Human Values 47, 855–881.

Blake Jackson, E., Dreyling, R., Pappel, I., 2021. A historical analysis on interoperability in estonian data exchange architecture: Perspectives from the past and for the future, in: Proceedings of ICEGOV'21 – the 14th International Conference on Theory and Practice of Electronic Governance, ACM. pp. 111–116.

Blanka, C., Krumay, B., Rueckel, D., 2022. The interplay of digital transformation and employee competency: A design science approach. Technological Forecasting and Social Change 178, 121575. doi:10.1016/j.techfore.2022.121575.

Blodgett, S.L., Barocas, S., Daumé III, H., Wallach, H., 2020. Language (technology) is power: A critical survey of" bias" in nlp. arXiv preprint arXiv:2005.14050 .

Breaugh, J., Rackwitz, M., Hammerschmid, G., 2023. Leadership and institutional design in collaborative government digitalisation: Evidence from Belgium, Denmark, Estonia, Germany, and the UK. Government Information Quarterly 40, 101788. doi:10.1016/j.giq.2022.101788.

Carter, L., Bélanger, F., 2005. The utilization of e-government services: citizen trust, innovation and acceptance factors. Information Systems Journal 15, 5–25.

Christensen, C.M., 2013. The innovator's dilemma: when new technologies cause great firms to fail. Harvard Business Review Press.

Creswell, J.W., Poth, C.N., 2016. Qualitative inquiry and research design: Choosing among five approaches. Sage Publications.

Das, S.D., Bala, P.K., Mishra, A.N., 2023. Towards defining a trustworthy artificial intelligence system development maturity model. Journal of Computer Information Systems , 1–22.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly , 319–340.

Draheim, D., Krimmer, R., Tammet, T., 2021. On state-level architecture of digital government ecosystems: From ict-driven to data-centric, in: Transactions on Large-Scale Data-and Knowledge-Centered Systems XLVIII: Special Issue In Memory of Univ. Prof. Dr. Roland Wagner. Springer, pp. 165–195.

Dreyling, R., Jackson, E.B., Tammet, T., Labanava, A., Pappel, I., 2021. Social, legal, and technical considerations for machine learning and artificial intelligence systems in government, in: Proceedings of ICEIS'2023 – the 25th International Conference on Enterprise Information Systems, SCITEPRESS. pp. 701–708.

Engstrom, D.F., Ho, D.E., 2020. Algorithmic accountability in the administrative state. Yale Journal on Regulation 37, 800.

Engstrom, D.F., Ho, D.E., 2021. Artificially intelligent government: a review and agenda. Research Handbook on Big Data Law , 57–86.

Engstrom, D.F., Ho, D.E., Sharkey, C.M., Cuéllar, M.F., 2020. Government by algorithm: Artificial intelligence in federal administrative agencies. Technical Report Public Law Research Paper No. 20-54. NYU School of Law. New York.

von Eschenbach, W.J., 2021. Transparency and the black box problem: Why we do not trust AI. Philosophy & Technology 34, 1607–1622.

Feng, S., Kudina, O., Halpern, B.M., Scharenborg, O., 2021. Quantifying bias in automatic speech recognition. arXiv preprint arXiv:2103.15122 .

Gillath, O., Ai, T., Branicky, M.S., Keshmiri, S., Davison, R.B., Spaulding, R., 2021. Attachment and trust in artificial intelligence. Computers in Human Behavior 115, 106607.

Gong, Y., Yang, J., Shi, X., 2020. Towards a comprehensive understanding of digital transformation in government: Analysis of flexibility and enterprise architecture. Government Information Quarterly 37, 101487.

Hinings, B., Gegenhuber, T., Greenwood, R., 2018. Digital innovation and transformation: An institutional perspective. Information and Organization 28, 52–61.

Holmström, J., 2022. From AI to digital transformation: The AI readiness framework. Business Horizons 65, 329–339.

Jöhnk, J., Weißert, M., Wyrtki, K., 2021. Ready or not, AI comes – an interview study of organizational AI readiness factors. Business & Information Systems Engineering 63, 5–20.

Kupi, M., McBride, K., 2021. Agile development for digital government services: Challenges and success factors, in: Edelmann, N., Csáki, C., Hofmann, S., Lampoltshammer, T.J., Alcaide Muñoz, L., Parycek, P., Schwabe, G., Tambouris, E. (Eds.), Proceedings of ePart'2021 – the 13th IFIP WG 8.5 International Conference on Electronic Participation, Springer International Publishing, Cham. pp. 139–150.

Kuziemski, M., Misuraca, G., 2020. AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. Telecommunications Policy 44, 101976.

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., Zou, J., 2023. GPT detectors are biased against non-native English writers. Patterns 4, 100779. doi:10.1016/j.patter.2023.100779.

Lunenburg, F.C., 2012. Organizational structure: Mintzberg's framework. International Journal of Scholarly Academic Intellectual Diversity 14, 1–8.

Mendelson, H., 2000. Organizational architecture and success in the information technology industry. Management Science 46, 513–529. doi:10.1287/mnsc.46.4.513.12060.

Mergel, I., Edelmann, N., Haug, N., 2019. Defining digital transformation: Results from expert interviews. Government Information Quarterly 36, 101385. doi:10.1016/j.giq.2019.06.002.

Mintzberg, H., 1989. Mintzberg on management: Inside our strange world of organizations. Free Press.

Moore, G.A., 2014. Crossing the Chasm: Marketing and Selling Disruptive Products to Mainstream Customers. Harper Business, New York. 3rd edition.

O'Reilly, C.A., Tushman, M.L., 2008. Ambidexterity as a dynamic capability: Resolving the innovator's dilemma. Research in Organizational Behavior 28, 185–206. doi:10.1016/j.riob.2008.06.002.

Sadiq, R., Safie, N., Abd Rahman, A., Goudarzi, S., 2021. Artificial intelligence maturity model: a systematic literature review. PeerJ Computer Science 7. doi:10.7717/peerj-cs.661.

Schiff, D., 2022. Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. International Journal of Artificial Intelligence in Education 32, 527–563.

Sikkut, S., Velsberg, O., K., V., 2020. #KrattAI: the Next Stage of Digital Services in #eEstonia. Republic of Estonia GCIO Office. URL: `https://tinyurl.com/2vn9xupm`.

Solheim-Kile, E., Wald, A., 2019. Extending the transactional view on public–private partnership projects: Role of relational and motivational aspects in goal alignment. Journal of Construction Engineering and Management 145, 04019030. doi:10.1061/(ASCE)CO.1943-7862.0001643.

Vaher, K., 2020. Next Generation Digital Government Architecture. Republic of Estonia GCIO Office.

van Noordt, C., Misuraca, G., 2022. Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union. Government Information Quarterly 39, 101714. doi:10.1016/j.giq.2022.101714.

Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D., 2003. User acceptance of information technology: Toward a unified view. MIS Quarterly 27, 425–478.

Vial, G., 2019. Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems 28, 118–144. doi:10.1016/j.jsis.2019.01.003.

Wirtz, B.W., Weyerer, J.C., Kehl, I., 2022. Governance of artificial intelligence: A risk and guideline-based integrative framework. Government Information Quarterly 39, 101685.

Yin, R.K., 2017. Case Study Research and Applications: Design and Methods. SAGE. 6th edition.

Zaoui, F., Souissi, N., 2020. Roadmap for digital transformation: A literature review. Procedia Computer Science 175, 621–628.

VI

R. Dreyling, T. Koppel, T. Tammet, and I. Pappel, "Challenges of genai chatbots in public services: An integrative review," *SSRN*, 2024

Challenges of Generative AI Chatbots in Public Services – An Integrative Review

Challenges of Generative AI Chatbots in Public Services

An Integrative Review

Richard M. Dreyling

Tallinn University of Technology, Department of Software Science, richard.iii@taltech.ee

Tarmo Koppel

Tallinn University of Technology, Department of Business Administration, tarmo.koppel@taltech.ee

Tanel Tammet

Tallinn University of Technology, Department of Software Science, tanel.tammet@taltech.ee

Ingrid Pappel

Tallinn University of Technology, Department of Software Science, ingrid.pappel@taltech.ee

As governments attempt to improve public service efficiency and effectiveness through the implementation of artificial intelligence applications like chatbots, the widespread adoption of generative AI technologies like ChatGPT have potentially changed the expectations of users. Public sector adoption of chatbots can lead to the opening of additional digital channels for public services, like the Bürokratt initiative in Estonia. This paper uses the case of the Bürokratt program to ground discussion of the challenges and opportunities of using GenAI chatbots in public service applications. Although LLM-based chatbots have the potential to greatly improve the customer experience for the people who would use them to access services many challenges exist. The researchers conducted an integrative review which identified challenges including but not limited to technical challenges in development, handling complex queries, security, privacy and bias, organizational barriers, as well as the legal and regulatory challenges and need for transparency and trustworthiness as well as suggest some solutions based on the literature in the field.

CCS CONCEPTS•Applied computing~ Computers in other domains~ Computing in government~ E-government•Computing methodologies~Artificial intelligence~ Natural language processing~ Natural language generation•Social and professional topics~Computing / technology policy~Government technology policy~Governmental regulations•Security and privacy~ Human and societal aspects of security and privacy~ Privacy protections

Additional Keywords and Phrases: Artificial Intelligence, Chatbot, GenAI, LLM

ACM Reference Format:

Richard M Dreyling, Tarmo Koppel, Tanel Tammet, Ingrid Pappel. 2024. Challenges of Generative AI Chatbots in Public Services – A Commentary: ACM Manuscript Submission Template.

1 INTRODUCTION: DEVELOPMENTS IN GENERATIVE AI

Governments around the world have varying levels of maturity when it comes to e-government and e-governance [1]. When it comes to offering AI enabled public services, a prerequisite is to have public services that are already offered via digital channels. Some maturity models have automation at the higher levels of their measures of maturity. Although AI and automation are not synonymous, in some sense, if a public service were offered through an AI enabled digital channel, it would be analogous to an automated business process between the citizen and government. Chatbots can potentially aid in management of relationships between government and citizens if executed well [2]. Citizens seem to believe that digital channels of communication are easier than traditional methods of communication [3].

One potential role for a generative AI chatbot based on a foundation model would be to create better responses for consumers of public services. This way, the functionality and level of response would be similar to the LLM-based tools the consumers are using in their daily life. This creates a better user experience which has the possibility to increase trust. However, this type of implementation comes with challenges.

The Bürokratt project [4] serves as an example of a project for building a conversational AI in public service: a virtual assistant using natural language to simplify interaction with the state agencies. The project was initiated in 2020 and the first version was published in 2022. The system is built around an existing chatbot toolkit from rasa.com and is primarily rule-based, while employing a limited amount of machine learning, mainly for intent discovery. However, up to 2024 it has seen very limited deployment and use, mostly due to the complexities of achieving sufficient conversational quality for practical uptake. In order to make the system give relevant answers to queries, the government agency deploying the system would have to create and maintain a large number of complex rules specific to the agency, plus perform training. This has clearly been too hard a task for the agencies. It seems likely that instead of attempting to continue with the deployment of Bürokratt, a new version will be built, centered around large language models (LLMs) and retrieval augmented generation (RAG).

1.1 Integrative Review

This paper consists of the results of an integrative review which was conducted by the researchers based on the methodological protocol detailed in [5]. This consists of six steps which are. Step one is to select a concept. The research team selected the concept of the challenges relating to adopting Generative AI chatbots in public services due to the appropriateness of the expansion of Generative AI chatbot use at the same time that the Estonian government has been implementing Bürokratt which is developed using traditionally developed chatbot technology. Step two is to determine the aim of the analysis. In this case the team chose to search of all of the technical, social, and legal literature relating to GenAI chatbots that would be relevant for their use in public services. Step three is to conduct the review, and step four is to organize and evaluate the data. Step five is analysis and synthesis. Step six is discussions and conclusions. And step seven is to disseminate the findings. Steps three through six were conducted among the team and resulted in the output for step seven, which comprises the rest of the paper.

1.2 Generative AI Chatbots

Following Conversational agents, traditionally developed modern chatbots, use NLP and ML capabilities, but the most advanced chatbots today use generative knowledge-based techniques. These techniques involve training the bot on large amounts of data, allowing it to generate responses based on patterns and information it has learned. This approach allows chatbots to have more natural and varied conversations, as they can generate responses in real-time rather than relying on pre-defined answers. Generative knowledge-based chatbots are becoming increasingly popular as they can provide more engaging and personalized interactions with users.

Chatbots using machine learning approaches use NLP to extract content from user input and can learn from conversations. These chatbots consider the entire dialog context, not just the current turn, and do not rely on predefined responses. Training these chatbots requires extensive datasets, which can be difficult to find as available datasets may be insufficient. These chatbots offer more advanced capabilities compared to pattern matching-based chatbots [6].

1.3 Large Language Models (LLMs)

Large Language Models (LLMs) are pre-trained models with the capability for fine-tuning for various language-related tasks [7]. Recent advancements in large language model (LLM) development encompass a diverse array of topics. These include innovations in architecture, strategies for training, enhancements in context length, fine-tuning methodologies, the emergence of multi-modal LLMs, applications in robotics, the creation of specialized datasets, benchmarking efforts, and pursuits focused on optimizing efficiency [8]. An AI system utilizing a Large Language Model (LLM) expands its scope beyond mere text generation, facilitating capabilities such as engaging in conversational interactions, completing tasks, logical reasoning, and demonstrating a degree of independent behavior [9]. However, integrating LLMs with conventional software systems like databases, planners, explicit rules, up-to-date texts, etc. is a complex question without sufficiently good answers yet [10]. The most promising and widely used approach right now is Retrieval Augmented Generation (RAG) which uses vector databases to search for related snippets of texts to be automatically inserted into the prompt [11].

1.4 Conversational AI in Public Services

In the last years the relevance of generative artificial intelligence (GenAI) and large language models (LLM) have become prevalent in the public discourse, and governments are still reckoning how to deal with these technologies and the drivers and challenges which they pose. One of the obvious challenges that these technologies are providing is the change in the environment external to governments. To put it plainly, when consumers of government services have used ChatGPT or other GenAI technologies based on LLMs it can affect the expectations of those citizens when they attempt to use government chatbots or other digital technologies to help people use another digital channel to access information or services which creates pressure to recreate the same experience [12]. At the same time, traditionally developed chatbots including current phase of development the Estonian government chatbot called Bürokratt, pose significant challenges on their own potentially creating an opening for the guarded implementation of LLM-based chatbots in governments.

2 CHALLENGES RELATED TO GENERATIVE AI CHATBOTS

2.1 Technical Challenges

Developing chatbots is not a simple task. Governments are reckoning with how to handle the many requirements and competences that are necessary to build a chatbot capable of answering citizens in a realistic manner with appropriate responses. Some governments are currently attempting to build chatbots on the road map of a greater vision to build personal virtual assistants with the capacity to conduct digital public services, like Estonia and the Bürokratt program [13].

2.1.1 Complex Queries

Chatbot users may enter long and complex queries where they present information in a structured and layered way. LLMs may struggle in parsing and understanding such inputs and have lower accuracy compared to simple queries. Deep semantic and syntactic understanding is required to process such texts. Latest LLMs, such as OpenAI's GPT-4 have demonstrated unprecedented cognitive capabilities, but still may fail in processing complex queries. A straight forward conversation with the chatbot may turn into a complex one even when the user switches to new topics within the same conversation. Chatbots may face challenges in distinguishing each of the cases and transition to the one that the user considers relevant.

2.1.2 Context Handling

Natural language if often vague. Words and phrases can have different meanings and understanding the message is dependent on the context of the conversation. Some conversations are more context dependent than others. Generative AI chatbots struggle especially in the context dependent conversation, where their understanding of user's intent may suffer. Understanding the conversation context over the length of the conversation remains a challenge for chatbots.

2.1.3 Cultural and Linguistic Insensitivity

Users' context is also due to the cultural and linguistic background. When generative AI chatbots are based on foundational models, these may be insufficient to account for cultural and linguistical nuances. When the chatbots misses its target user's cultural-linguistic background, it may fail in to compose appropriate and respectful responses. It may even misinterpret user's intent or be insensitive in other ways. Misinterpretations will disallow the organization to provide the same quality and consideration of service to all population groups. This will lead to dissatisfaction and even other consequences of ethical nature.

2.1.4 Variability in Users' Self-expression

Chatbot users are different in the way they speak and express themselves. This includes different language styles, spelling and grammar. Chatbots need to be up to the task when users express themselves in diverse ways. Chatbot's failure to identify user's intent may be due to the variability of user input and the shortcomings in the training data. Developers face questions, how to train the LLM based chatbot, so that it would be consistent and reasonable in its responses. Understanding on users' diverse expressions is crucial for the next era chatbots.

2.1.5 Lack of Domain Specific Understanding

As LLMs are trained on public domain data, these are good in general understanding and common ground knowledge. However, domain specific areas, such as many government services, may have not been included in LLMs training data sets. This may result in failure to understand the specialty context, terminology, industry-specific standards or other domain specific knowledge. Many domains have specialty jargon and terminology that were not included in LLMs training data. Understanding the proper context within a discrete domain in crucial for the chatbot service. When the users input specialized language, this may lead to misunderstandings and consequent incorrect responses by the chatbot.

2.1.6 Lack of up-to-date Information

Some domains, like government services, require constant updates. In some fields, which require regulatory compliance, regulations may change frequently. Chatbots should be aware of the latest developments and regulatory requirements. LLMs are trained on datasets that are historical in nature. This also means that generative AI chatbots lack the ability to access real-time information. Users often seek information about currently available services or service updates, that they've ordered. When LLM-based chatbots are disconnected from the presence, this likely results in outdated responses to user queries. Regularly re-training or fine-tuning LLMs with new specialized information, and forgetting outdated information, may be complex and prohibitively expensive. Developers need to find ways to link LLMs to the organization's internal and external databases, to make sure, the service provided is up to date with latest requirements, procedures, events, resource availability, service status and other dynamic aspects of the service.

2.1.7 Integration to Databases and Responsiveness

Database integration is inevitable for any chatbot that does more than just providing static textual knowledge or seeking up latest status. Linking a chatbot to the organization's or governmental databases rises many challenges. Integrations need to be robust, secure and seamless. These databases may use different technologies and formats, which presents technical challenges for compatibility and interoperability. Additionally, the system needs to be scalable, as when the chatbot's user base grows, the demands for traffic to the database also increases. Growing traffic may hence diminish or even compromise the performance of the system. Real-time integrations to various modules and databases increase the latency and response time of requests. Users expect prompt responses, where delays in response generation and information retrieval will likely impact overall user experience.

Challenges for responsiveness are tightly connected to how much contextual information chatbot stores in its memory. Efficient memory management is crucial for optimal chatbot performance. This is a balancing challenge, where on the one hand the developers need to minimize memory usage for the sake of responsiveness, but from another hand, need to allow the relevant information be stored in the memory [10] [11].

2.1.8 Memory

Generative AI chatbots face several challenges related to remembering the previous exchange with the user and long-term memory. These issues impact chatbot's ability to maintain context and to provide accurate responses. As LLM context window is limited, generative AI chatbots struggle to retain context over extended or repeating conversations that takes place over multiple sessions. As a result, users may be asked the same information over again, or in worse case, chatbot would presumes such information non-existent nor asks the user to (re)state such information. When conversations take place over multiple sessions, chatbots are challenged also to account references to previous parts of the conversation, that may refer to pronouns, specific terms or other context or case dependent information. Some LLMs prioritize recent inputs over what was said over long time ago, even if that historical information fits into the context window. These shortcomings lead to a narrow focus on the most recent interactions and potentially leave out important details form the start, when the user bought their case to chatbot's attention.

Storing and managing conversational history to improve the performance of the chatbot raises additional issues about data security and privacy. Chatbot designers cannot account for all types of information that the users would enter into the system. Other users may be uncomfortable about entering sensitive data via chatbot channel. On the other hand, the legislations, such as the GDPR, obligate the operators of such technology to follow strict data protection rules.

2.1.9 Data Security and Privacy

Chatbots often deal with personal or sensitive data that requires specific methods and procedures to handle. Guaranteeing user's privacy and ensuring that their data is stored securely is a major challenge. This data is exchanged between the user's device and the chatbot service. Additionally, internal or externals services, such as databases may be connected to the pipeline. The physical computing devices, together with the data links between them, and the data transmission protocols may all be vulnerable for breaches.

Integrating chatbots with databases poses serious challenges in ensuring data security and privacy. Databases may contain sensitive information, such as the user's health data. Access to sensitive data must be restricted, calling for encryption mechanisms to protect the data from unauthorized access. The security concerns for the database data are not only about data leakage, but also about data manipulation by bad actors, e.g., for sabotage purposes. Therefore, a chatbot connected to a database, should include robust authentication mechanisms.

The development of chatbots within the government sphere provides more challenges related to data privacy than many for profit implementations of chatbots. The latter do not have to deal with as much, or organizations like Open AI consider less of a priority. Governments, by the nature of their existence as the entity which creates and enforces laws have to consider topics like security and privacy of citizens as well as the veracity of responses [13]. Similarly, these concerns pose a challenge when governments are developing chatbots of their own.

2.1.10 Bias and Fairness

Chatbots will inevitably perpetuate biases present in the training data of LLM which it was based on. Unintentional discrimination or misinformation may happen due to the unbalanced training data. Training of a large language model requires vast amounts of textual data. Such foundational models are never comprehensive, meaning these do not cover all the business domains, hence are limited when faced with unforeseen scenarios. Importantly, the bias may introduce serious errors in the answers produced. In 2021 the Dutch government fell as a consequence of the errors made by the machine learning algorithm used for checking childcare benefits [14].

2.2 Legal and Regulatory Challenges

Given certain conditions in the level of digitization in a country, the legal framework combined with the sharing of knowledge regarding technologies across national borders can work as a barrier to the adoption or as an opportunity [15]. The necessity for a regulatory framework to govern the use of AI can be seen in the EU developing a preliminary EU Act as far back as 2021 and the Biden administration's release of an executive order focused on the safety, and security of AI in the United States [16]. Canada, China, and Australia are investigating how to approach the regulation of generative AI technologies as well [17].

2.2.1 Compliance with Data Protection Laws

The EU has typically functioned as the leader in the area of technology regulation due to the passing of the EU General Data Protection Regulation (GDPR) [18]. GDPR and similar national laws can have an effect on the ways in which an AI system is designed or adopted. For example, Estonia has the personal data protection act (PDPA) which closely mirrors GDPR but also stipulates that the government must store data where it is collected [19]. This means that when designing the original Bürokratt chatbot, rather than using a data pool and having the chatbot connect to it, they instead had to design a network of chatbots which would refer the user to the relevant government ministry and have that chatbot answer the question of the citizen.

Under GDPR the data and privacy of citizens using web technologies should be anonymized or deleted upon request. One application of this would be to ensure the deletion of private information of citizens from conversations with AI based chatbots. Some researchers also provide general guidelines for the regulation of AI and chatbots. External audits for privacy and data protection in the use of chatbots is one way to ensure that the key concerns of regulators and researchers are appropriately addressed [20]. However, one challenge is that not everyone who uses AI enabled chatbots and services fall under the jurisdiction of the EU and therefore this may lack the type of enforceability necessary to have a large effect in many countries [20].

2.2.2 The EU AI Act

The EU AI act recently made headlines because decisionmakers reached an informal agreement on what should be present in the law. Researchers have called for the regulation of AI in recent years, highlighting the need for guidance in the US [21] and attempting to limit monopolistic behaviors among the large AI providers [22]. Others have called for approaches to regulating AI through loping and iterative frameworks [23]. However, the EU AI Act takes its final shape will have effects on the AI legislation across the world whether informally or formally in what is called the "Brussels effect" [20]. This effect refers to an informal effect, not to be confused with enforceability outside EU citizen context. The EU AI Act divides types of systems into risk categories and has suggestions and requirements for systems that may have an impact on people's lives which could potentially transgress their rights [24]. In this concept of the systems the chatbot falls below the risk classification of high-risk [25]. However, there has been some discussion of regulating large-language models and foundation models in a similar way to the high-risk systems [26]. Some researchers believe that the way the law was drafted may not actually help the EU achieve "trustworthy AI" but instead has a misalignment problem by assigning risk categories instead of focusing on elements that may actually increase the trustworthiness of the AI systems [27].

2.2.3 Potential Liabilities

Contemporary concerns emerge due to technological advancements, which have allowed a significant reduction in the expenses associated with gathering, storing, analyzing, and utilizing data on a large scale. This expands information imbalance well beyond individual transactions. These advancements are commonly encapsulated by the phrases "big data" and "AI." [28]. These risks and liabilities can be associated with the use of chatbots in public services. One major concern is data protection and privacy issues. Compliance with data protection laws such as the GDPR is crucial when implementing AI systems like chatbots. Failure to comply with these regulations can lead to legal consequences, fines, and damage to the reputation of the organization. The risk of misinformation or bias mat take precedence in the responses provided by chatbots. If the algorithms powering the chatbots are not properly trained or monitored, these may provide inaccurate information ort bias, which leads to legal implications.

Privacy breaches might be considered a critical concern with any AI system. If chatbots collect and store personal data without consent or fail to secure this data properly, it can result in violations of data protection laws such as the GDPR. Individuals have the right to control their personal information and organizations that fail to uphold these rights can face legal action. Privacy in AI systems requires assessing each component against privacy frameworks. Current AI poses risks in knowledge representations, NLP, reasoning, and ML model creation and use [29].

Another liability risk is the potential for security breaches and cyber-attacks. Chatbots, like any other digital technology, are vulnerable to hacking and data breaches. If sensitive information is compromised, it can result in financial losses, legal action, and harm to individuals whose data has been exposed. Accountability and transparency of chatbots cannot always be guaranteed. If something goes wrong or a mistake is made, it may be difficult to determine who is responsible – the developers, the organization using the chatbot, or the AI system itself. The primary legal instruments essential for mitigating the dangers of discrimination propelled by AI are anti-discrimination and data protection laws. If effectively implemented, both these legal mechanisms hold the potential to combat unlawful discrimination [30].

Ethical conflicts may bring additional legal problems that may arise if technical challenges are not adequately addressed. One significant legal issue that may emerge is discrimination. If chatbots exhibit bias, whether intentional or unintentional, in the responses or decisions they make, it can lead to discriminatory outcomes, which is illegal in many jurisdictions. For example, if a chatbot provides different information or services based on factors like race, gender or disability, it can violate anti-discrimination laws and lead to legal repercussions.

When chatbots are not transparent in how they operate or make decisions, it can raise legal issues around accountability and compliance. Regulatory bodies require organizations to provide explanations for how their AI systems work, especially in sensitive areas like healthcare or finance. Failure to do so can lead to legal challenges and regulatory scrutiny.

The primary concern so far has been that companies are not held fully responsible for the risks they introduce to consumer privacy and data security [28]. Regulation and oversight are struggling to keep up with rapidly advancing technologies that can self-improve, leading to a risk of falling far behind [31].

2.3 Organizational Challenges

Governments seek to implement AI tools like chatbots to increase efficiency and effectiveness. These are an example of an AI tool that can be adopted in order to open a new digital channel for communication and potentially services with consumers of government services, usually residents or citizens of that country. With digital transformations associated with new technologies, typically there are changes in organizations [32]. However, according to Van Noordt and Misuraca [33] the feedback mechanisms causing large scale shifts within the organizational mechanism of the public administration are limited when it comes to the adoption of traditionally developed chatbots like conversational agents [33]. It is not clear yet if this is the case with LLM based chatbots.

2.3.1 Drivers of organizational technology adoption

The reasons for organizations choosing to adopt new technologies can stem from many factors. However, one of the most important factors in this decision is the environment in which the organization exists [32]. Changes in the external environment can cause pressure in a variety of ways on the organization itself. This can be external pressure for change when the potential user base for public sector services sees the efficiency and effectiveness of solutions that they use in a non-public sector context, the demand increases to have a similar user experience from the government [12]. There can also be mimetic pressure, the imitation of other organizations [34]. Because the work force that comprises the public sector does not exist in a vacuum, large changes in technology can also affect the work force and cause internal pressure for digital transformation [9]. Demand for better service seems to be the primary driver in the adoption of GenAI [34]

According to Brynjolfsson [35] there is reason to believe that GenAI could be similar to historical general-purpose technologies and provide a lot of value creation and reduction of costs. However, it has also been recognized that these technologies take a lot of time to come to fruition in terms of measurable worker productivity [36]. This can be attributed to the need for investments related to the technological innovation as well as changes to business processes, structures, and skills [36]. However, with Gen AI in particular, many of the accompanying required components for widespread adoption are already in place [37]. Initial research shows the ability for GenAI to significantly boost productivity in customer service applications when paired with human workers [38].

2.3.2 Skills and competences

For a public entity seeking to adopt AI in hopes of increasing productivity it is important to have AI skills in the organization. In public administrations it is important even when hiring teams of outside developers or reusing existing software to ensure that there is enough AI skill in the organization to understand what is being delivered and that it follows governance rules and the legal and regulatory environment [39]. This can be difficult because these are in demand skills and public administrations may face a competitive labor market for these services. An organization which has internal AI talent would be considered more mature in this metric [40].

2.3.3 Governance

Governance is a challenge for organizations seeking to adopt Gen AI governance can be a potential area of contention because governance is now more challenging [41]. With the research that has been conducted specific to GenAI governance for organizations it is clear that the governance challenge is still being worked out in regulations and organizations [42]. The topic is complex and has many overlapping pieces. Schneider, Abraham, and Meske recommend a framework that has structural, procedural, and relational mechanisms [41]. For governmental organizations seeking to implement GenAI in public services governance should include the procedural mechanism of compliance monitoring, possibly even with additional outside models.

3 CONCLUSION

Though the amount of speed in the development of the technology itself makes it hard for researchers, businesses, and governments to have a clear view of the potential of GenAI in ten years, it appears that the risks are somewhat clearer [16] [42]. Risk management methods have been suggested by papers in both the United States and [17] [43]. The AI risk management standards for GPAIs and foundational models is grounded in the documents put forth by the National Institute of Standards and Technology (NIST) [30]. The Australian risk considerations come from a paper by the office of the Chief scientist of Australia [17]. One clear through line in the research on the topic is the necessity for transparency [44] as well as data protection [20] [45].

User trust in chatbots is dependent of most of the aforementioned challenges. For example, when the chatbot would provide outdated information, this would undermine users' trust not only toward the chatbot, but toward the government service in general. Users rely on the government on reliable information, which needs to be comprehensive, meaning nothing is left out, and timely, meaning it's latest.

Some research has been conducted on users and how they may feel about privacy and data protection. Atkins et al concluded that the responsible AI guidelines that were current in 2021 were not a good tool for users to be able to arbitrate whether the AI chatbot is using responsible AI. One potential solution which was tested on users of a chatbot for this issue was to have embedded conversational privacy prompts that the chatbot would give to users. This research showed that users were not opposed to such privacy prompts to remind them [46].

In addition, legal researchers have proposed some potential solutions to the challenges faced by governments reckoning with the rapid spread of LLM based applications like ChatGPT. One suggestion is to use the data which is the lifeline and value proposition of the GenAI applications to regulate the corporations producing the applications (Aaronson). Other researchers have suggested that even though the European Union is passing an AI Act, the operative regulation under which they will most likely regulate many of the GenAI companies will actually be GDPR [47]. In the United States, one piece of research underlined the need for transparency and suggested the creation of an AI agency for regulation and enforcement [31]. Some believe that applying the high-risk category to LLMs and foundation models under the AI act is a way to handle LLMs under the existing EU AI Act proposal [26]. For the purpose of regulatory compliance, it is understood that there will need to be some sort of collaboration across the AI value chain in which it is necessary to differentiate between the developers, implementors, and the consumers of AI products including the use of foundation models [25].

The most likely solution to the challenge of generative AI chatbots is through some manner of implementing an open-source AI foundation model and fine-tuning it for specific use cases with RAG. This could potentially have the benefit of having not only the specific answers to questions someone would ask a particular government authority, but the model could be adjusted to use authorization methods or an API to complete government services, be implemented on premises to comply with data protection regulations, ensure no personal data is used as training data for the next version of the model, and potentially to not speculate or give wrong answers to ameliorate the hallucination problem. Such a fine-tuned GenAI enhanced chatbot would have to be continually audited for safety and trustworthiness sometimes using other models to ensure compliance, but could have a better chance at giving the citizen or resident the user experience they now expect in a post ChatGPT world.

REFERENCES

[1] Dirk Draheim, Robert Krimmer, and Tanel Tammet. 2021. On state-level architecture of digital government ecosystems: From ICT-driven to data-centric. In Transactions on Large-Scale Data- and Knowledge-Centered Systems XLVIII: Special Issue In Memory of Univ. Prof. Dr. Roland Wagner, 165–195. Springer Berlin Heidelberg.

[2] Ana Marta Goucha de Carvalho. 2023. Government Augmented Intelligence-The Use of AI to Improve Citizen Relationship Management. PhD Dissertation.

[3] Aggeliki Androutsopoulou, Nikos Karacapilidis, Euripidis Loukis, and Yannis Charalabidis. 2019. Transforming the communication between citizens and government through AI-guided chatbots. Gov. Inf. Quart. 36, 2 (2019), 358–367.

[4] Ministry of Economic Affairs and Communication, Estonia. 2023. Bürokratt vision. https://www.kratid.ee/en/burokratt-visioon.

[5] Anna M. Kutcher, and Virginia T. LeBaron. "A simple guide for completing an integrative review using an example article." Journal of Professional Nursing 40 (2022): 13-19.

[6] Eleni Adamopoulou and Lefteris Moussiades. 2020. Chatbots: History, technology, and applications. Machine Learning with Applications 2 (2020), 100006.

[7] Terrence J. Sejnowski. 2023. Large language models and the reverse Turing test. Neural Comput. 35, 3 (2023), 309–342.

[8] Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Nick Barnes, and Ajmal Mian. 2023. A comprehensive overview of large language models. arXiv preprint arXiv:2307.06435 (2023).

[9] Andrew Zhao, Daniel Huang, Quentin Xu, Matthieu Lin, Yong-Jin Liu, and Gao Huang. 2023. Expel: LLM agents are experiential learners. arXiv preprint arXiv:2308.10144 (2023).

[10] Grégoire Mialon, Roberto Dessì, Maria Lomeli, Christoforos Nalmpantis, Ram Pasunuru, Roberta Raileanu, Baptiste Rozière, et al. 2023. Augmented language models: a survey. arXiv preprint arXiv:2302.07842 (2023).

[11] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, et al. 2020. Retrieval-augmented generation for knowledge-intensive NLP tasks. In Advances in Neural Information Processing Systems 33 (NeurIPS 2020), 9459–9474.

[12] Gregory Vial. 2021. Understanding digital transformation: A review and a research agenda. Managing Digital Transformation (2021), 13–66.

[13] Omar Ballester. 2021. An artificial intelligence definition and classification framework for public sector applications. In Proceedings of DG.O2021: The 22nd Annual International Conference on Digital Government Research, 67-75.

[14] Rahul Rao. 2023. Artificial Intelligence in Government. IEEE Spectrum. https://spectrum.ieee.org/artificial-intelligence-in-government.

[15] Yulia Petriv, Regina Erlenheim, Valentyna Tsap, Ingrid Pappel, and Dirk Draheim. 2020. Designing effective chatbot solutions for the public sector: A case study from Ukraine. In Proceedings of the 6th International Conference on Electronic Governance and Open Society: Challenges in Eurasia (EGOSE 2019), St. Petersburg, Russia, November 13–14, 2019, 320-335. Springer International Publishing.

[16] Joseph R. Biden. 2023. Executive order on the safe, secure, and trustworthy development and use of artificial intelligence. (2023).

[17] Genevieve Bell, Jean Burgess, Julian Thomas, and Shazia Shadiq. 2023. Rapid Response Information Report: Generative AI-language models (LLMs) and multimodal foundation models (MFMs). (2023).

[18] Charlotte Siegmann and Markus Anderljung. August 2022. How, E.U.: The Brussels Effect and Artificial Intelligence. August 2022.

[19] Richard Dreyling, Eric Blake Jackson, Tanel Tammet, Alena Labanava, and Ingrid Pappel. 2021. Social, Legal, and Technical Considerations for Machine Learning and Artificial Intelligence Systems in Government. In Proceedings of ICEIS (1), 701-708. 2021.

[20] Martin Hasal, Jana Nowaková, Khalifa Ahmed Saghair, Hussam Abdulla, Václav Snášel, and Lidia Ogiela. 2021. Chatbots: Security, privacy, data protection, and social aspects. Concurrency and Computation: Practice and Experience 33, 19 (2021), e6426.

[21] Chris Lewis. 2022. The Need for a Legal Framework to Regulate the Use of Artificial Intelligence. U. Dayton L. Rev. 47 (2022), 285.

[22] Ganesh Sitaraman and Tejas N. Narechania. 2023. An Antimonopoly Approach to Governing Artificial Intelligence. SSRN (2023).

[23] Sarah Kreps and Adi Rao. 2023. AI and the Regulatory Challenge: A New Framework Using the Seto Loop. SSRN (2023).

[24] Lilian Edwards. 2021. The EU AI Act: a summary of its significance and scope. Artificial Intelligence (the EU AI Act) 1 (2021).

[25] Philipp Hacker, Andreas Engel, and Marco Mauer. 2023. Regulating ChatGPT and other large generative AI models. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, 1112-1123.

[26] Philipp Hacker. 2023. The European AI Liability Directives–critique of a half-hearted approach and lessons for the future. Comput. Law & Secur. Rev. 51 (2023), 105871.

[27] Johann Laux, Sandra Wachter, and Brent Mittelstadt. 2024. Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk. Regul. & Gov. 18, 1 (2024), 3–32.

[28] Ginger Zhe Jin. 2018. Artificial intelligence and consumer privacy. In *The economics of artificial intelligence: An agenda*, pp. 439-462. University of Chicago Press, (2018).

[29] Curzon, James, Tracy Ann Kosa, Rajen Akalu, and Khalil El-Khatib. 2021. "Privacy and artificial intelligence." *IEEE Transactions on Artificial Intelligence* 2, no. 2 (2021): 96-108.

[30] Frederik Zuiderveen Borgesius. 2018. Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy (2018): 42.

[31] Blake Murdoch. Privacy and artificial intelligence: challenges for protecting health information in a new era. BMC Medical Ethics 22 (2021): 1-5.

[32] Ines Mergel, Noella Edelmann, and Nathalie Haug. 2019. Defining digital transformation: Results from expert interviews. Government Information Quarterly. 36, 4 (2019), 101385.8

[33] Colin Van Noordt and Gianluca Misuraca. 2019. New Wine in Old Bottles: Chatbots in Government: Exploring the Transformative Impact of Chatbots in Public Service Delivery. In Proceedings of the 11th IFIP WG 8.5 International Conference on Electronic Participation (ePart 2019), San Benedetto Del Tronto, Italy, September 2–4, 2019, 49-59. Springer International Publishing.

[34] Rohit Rohit, and Mona Ashok. 2024. Making sense of AI benefits: a mixed-method study in Canadian public administration." *Information Systems Frontiers* (2024): 1-35.

[35] Erik Brynjolfsson, Daniel Rock, and Chad Syverson. 2019 Artificial intelligence and the modern productivity paradox. *The economics of artificial intelligence: An agenda* 23 (2019): 23-57.

[36] Erik Brynjolfsson, and Lorin M. Hitt. 2000. Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic perspectives* 14, no. 4 (2000): 23-48.

[37] Andrew McAfee, Daniel Rock, and Erik Brynjolfsson. 2023. How to Capitalize on Generative AI. Harvard Business Review 101, no. 6 (2023): 42-48

[38] Erik Brynjolfsson, Danielle Li, and Lindsey R. Raymond. Generative AI at work. No. w31161. National Bureau of Economic Research, (2023).

[39] Labanava, Alena, Richard Michael Dreyling III, Marzia Mortati, Innar Liiv, and Ingrid Pappel. 2022. "Capacity Building in Government: Towards Developing a Standard for a Functional Specialist in AI for Public Services." In *International Conference on Future Data and Security Engineering*, pp. 503-516. Singapore: Springer Nature Singapore, (2022).

[40] Philipp Fukas, Jonas Rebstadt, Florian Remark, and Oliver Thomas. 2021. Developing an Artificial Intelligence Maturity Model for Auditing. In *ECIS*. (2021).

[41] Johannes Schneider, Rene Abraham, Christian Meske, and Jan Vom Brocke. Artificial intelligence governance for businesses. *Information Systems Management* 40, no. 3 (2023): 229-249.

[42] Sadiq, Shazia. (2023) Generative AI: Language models and multimodal foundation models. (2023).

[43] Anthony Barrett, Jessica Newman, Brandie Nonnecke, D. Hendrycks, EVAN R. Murphy, and Krystal Jackson. 2023. AI risk-management standards profile for general-purpose AI systems (GPAIS) and foundation models. Center for Long-Term Cybersecurity, UC Berkeley. https://perma.cc/8W6P-2UUK (2023).

[44] Gordon Unzen. 2023. Artificial Intelligence and the Administrative State: Regulating the Government Use of Decision-Making Technology. Minn. J. Law, Sci. & Technol. 25, 1 (2023), 209.

[45] Fouad Leboukh, Emmanuel Baba Aduku, and Omar Ali. 2023. Balancing ChatGPT and data protection in Germany: challenges and opportunities for policy makers. J. Pol. and Ethics in New Tech. and AI 2, 1 (2023), e35166.

[46] Birgit Brüggemeier, and Philip Lalone. 2022. Perceptions and reactions to conversational privacy initiated by a conversational user interface. Comput. Speech & Lang. 71 (2022), 101269.

[47] Josephine Wolff, William Lehr, and Christopher S. Yoo. 2023. Lessons from GDPR for AI Policymaking. Available at SSRN 4528698 (2023).

Received January 2024; Resubmitted April 2024

Curriculum Vitae

Personal data

Name	Richard Dreyling III
Date and place of birth	10 June 1984 St. Paul, Minnesota, USA
Nationality	American

Contact information

Email richard.iii@taltech.ee
rmdreyling@gmail.com

Education

2020–2024	Tallinn University of Technology, School of Information Technologies, Information Technology, PhD studies
2018–2020	Tallinn University of Technology, School of Information Technologies, E-Governance Technologies and Services , MSc
2007–2011	University of California - Berkeley, United States of America Political Economy, BA

Language competence

English	native
Russian	B1

Professional employment

2014–2018	World Wide Technology, Inc., Consulting Solutions Architect
2013–2014	Complete Solar Solution, Business Development Representative
2002–2013	United States Marine Corps, Signal Intelligence Operator and Analyst

Supervision (Defended)

- Alissa Tšerešneva, Feasibility of Artificial Intelligence Technologies in Estonian State Budget Planning, MSc, supervisor **Richard Dreyling**, Tallinn University of Technology, 2024.

- Juhani Lemmik, Operationalising an Artificial Intelligence Maturity Model in the Public Sector, MSc, supervisor **Richard Dreyling**, Tallinn University of Technology, 2024.

- Maria Bušujeva, Measuring Technical Debt in Public Sector: Case Study of Estonia, MSc, supervisor **Richard Dreyling**, Tallinn University of Technology, 2024.

- Tii Tammik-Võsu, Data Sharing Practices in Estonian Agricultural Sector: Exploring Trust as a Facilitating Factor, MSc, supervisor **Richard Dreyling**, co-supervisor Alena Labanava, Tallinn University of Technology, 2024.

- Nikita Vlassenkov, How the establishment of the new e-Residency issue points affecting the number of e-Residents from these countries and area around with Causal-Impact method, MSc, supervisor **Richard Dreyling**, Tallinn University of Technology, 2023.

- Aghahuseyn Aghabayli, Internet voting in Azerbaijan: Possible implementation, challenges, and potential, MSc, supervisor **Richard Dreyling**, Tallinn University of Technology, 2023.

- Ebru Shentyurk, Data Protection Regulations Effect on Startups: GovTech Ecosystem in Lithuania, MSc, supervisor **Richard Dreyling**, Tallinn University of Technology, 2023.

- Martín Paul Peñaherrera Maldonado, Tallinn as a City Powered by Collaboration: Co-designing a Platform to Improve Citizen-Municipality Interactions, MSc, supervisor Ingrid Pappel, co-supervisor Eric Jackson, **Richard Dreyling**, Tallinn University of Technology, 2022.

- Simone Dalledone, AI-based services to help Italian high school seniors in their post high school transition, MSc, supervisor Luigi Buzzachi, co-supervisor Ingrid Pappel, **Richard Dreyling**, Politecnico di Torini, 2022.

- Katariina Muru, Pilot for smart vaccination certificates, MSc, supervisor Ingrid Pappel, co-supervisor Eric Jackson, **Richard Dreyling**, Tallinn University of Technology, 2021.

- Olena Roraff, Cybersecurity policy for the satellite industry: Governance challenges and solutions, MSc, supervisor Paul Liias, co-supervisor Eric Jackson, **Richard Dreyling**, Tallinn University of Technology, 2021.

- Snowbar Akbar, Chatbot for assisting parents of children with chronic long-term illness, MSc, supervisor Ingrid Pappel, co-supervisor **Richard Dreyling**, Tallinn University of Technology, 2021.

- Karl Andreas Sprenk, Becoming smart with and for young people - Case study of Estonian municipalities, MSc, supervisor Ingrid Pappel, co-supervisor **Richard Dreyling**, Tallinn University of Technology, 2021.

- Ganenthra Ravindran, Market efficiency of Bitcoin: Evidence from the Efficient Market Hypothesis (EMH), BSc, supervisor Simona Ferraro, co-supervisor **Richard Dreyling**, Tallinn University of Technology, 2022.

Publications

1. R. Dreyling, , T. Tammet, and I. Pappel, "Technology push in ai-enabled services: How to master technology integration in case of bürokratt," *SN Computer Science*, vol. Future Data Science Engineering, no. 5, p. 738, 2024

2. R. Dreyling, E. B. Jackson, T. Tammet, A. Labanava, and I. Pappel, "Social, legal, and technical considerations for machine learning and artificial intelligence systems in government.," in *ICEIS (1)*, pp. 701–708, 2021

3. R. Dreyling, J. Lemmik, T. Tammet, and I. Pappel, "An artificial intelligence maturity model for the public sector: Design science approach," *TalTech Journal of Social Sciences*, vol. 14, no. 2, p. 16, Forthcoming

4. R. M. Dreyling, T. Tammet, and I. Pappel, "Artificial intelligence use in e-government services: A systematic interdisciplinary literature review," in *International Conference on Future Data and Security Engineering*, pp. 547–559, Springer, 2022

5. R. Dreyling, K. McBride, T. Tammet, and I. Pappel, "Navigating the AI maze: Lessons from Estonia's Bürokratt on public sector AI digital transformation," *SSRN*, 2024
6. R. Dreyling, T. Koppel, T. Tammet, and I. Pappel, "Challenges of genai chatbots in public services: An integrative review," *SSRN*, 2024
7. R. Dreyling, E. Jackson, and I. Pappel, "Cyber security risk analysis for a virtual assistant g2c digital service using fair model," in *2021 Eighth International Conference on eDemocracy and eGovernment (ICEDEG)*, pp. 33–40, 2021
8. R. Dreyling, R. Erlenheim, T. Tammet, and I. Pappel, "Ai readiness assessment for data-driven public service projects: Change management and human elements of procurement," *Human Factors, Business Management and Society*, vol. 97, no. 97, 2023
9. R. M. Dreyling III, T. Tammet, and I. Pappel, "Digital transformation insights from an ai solution in search of a problem," in *International Conference on Future Data and Security Engineering*, pp. 341–351, Springer, 2023
10. E. Blake Jackson, R. Dreyling, and I. Pappel, "A historical analysis on interoperability in estonian data exchange architecture: Perspectives from the past and for the future," in *Proceedings of ICEGOV'21 – the 14th International Conference on Theory and Practice of Electronic Governance*, pp. 111–116, ACM, 2021
11. E. B. Jackson, R. Dreyling, and I. Pappel, "Challenges and implications of the who's digital cross-border covid-19 vaccine passport recognition pilot," in *2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG)*, pp. 88–94, IEEE, 2021
12. A. Labanava, R. M. Dreyling III, M. Mortati, I. Liiv, and I. Pappel, "Capacity building in government: Towards developing a standard for a functional specialist in ai for public services," in *International Conference on Future Data and Security Engineering*, pp. 503–516, Springer, 2022
13. A. Labanava, R. M. Dreyling, and A. Norta, "Potential of smart contracts in the pharmaceutical supply chain of belarus," in *2022 IEEE 1st Global Emerging Technology Blockchain Forum: Blockchain & Beyond (iGETblockchain)*, pp. 1–6, IEEE, 2022

Elulookirjeldus

Isikuandmed

Nimi	Richard Dreyling III
Sünniaeg ja koht	10. juuni 1984, St. Paul, Minnesota, USA
Kodakondsus	Ameeriklane

Kontaktandmed

E-post	richard.iii@taltech.ee rmdreyling@gmail.com

Haridus

2020–2024	Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Infotehnoloogia, doktoriõpe
2018–2020	Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, E-valitsemise tehnoloogiad ja teenused, magistriõpe
2007–2011	California Ülikool - Berkeley, Ameerika Ühendriigid Poliitökonoomia, bakalaureus

Keelteoskus

Inglise keel	emakeel
Vene keel	B1

Töökogemus

2014–2018	World Wide Technology, Inc., konsultatsioonilahenduste arhitekt
2013–2014	Complete Solar Solution, ärilahenduste esindaja
2002–2013	Ameerika Ühendriikide Mereväekorpus, signaalluure operaator ja analüütik

Juhendamine (Kaitstud)

- Alissa Tšerešneva, Tehisintellekti tehnoloogiate teostatavus Eesti riigieelarve planeerimisel, MSc, juhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2024.
- Juhani Lemmik, Tehisintellekti küpsusmudeli rakendamine avalikus sektoris, MSc, juhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2024.
- Maria Bušujeva, Tehnilise võla mõõtmine avalikus sektoris: Eesti juhtumianalüüs, MSc, juhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2024.
- Tii Tammik-Võsu, Andmete jagamise tavad Eesti põllumajandussektoris: Uurides usalduse rolli hõlbustava tegurina, MSc, juhendaja **Richard Dreyling**, kaasjuhendaja Alena Labanava, Tallinna Tehnikaülikool, 2024.
- Nikita Vlassenkov, Uute e-residentsuse väljastuspunktide loomise mõju e-residentide arvule nendest riikidest ja piirkonnast CausalImpact meetodi abil, MSc, juhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2023.

- Aghahuseyn Aghabayli, Internetihääletus Aserbaidžaanis: võimalikud rakendused, väljakutsed ja potentsiaal, MSc, juhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2023.

- Ebru Shentyurk, Andmekaitsereeglite mõju idufirmadele: GovTech ökosüsteem Leedus, MSc, juhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2023.

- Martín Paul Peñaherrera Maldonado, Tallinn kui koostööpõhine linn: Platvormi kooskujundamine kodanike ja omavalitsuse suhtluse parandamiseks, MSc, juhendaja Ingrid Pappel, kaasjuhendaja Eric Jackson, **Richard Dreyling**, Tallinna Tehnikaülikool, 2022.

- Simone Dalledone, Tehisintellekti põhised teenused, mis aitavad Itaalia keskkooliõpilastel üleminekul pärast keskkooli, MSc, juhendaja Luigi Buzzachi, kaasjuhendaja Ingrid Pappel, **Richard Dreyling**, Politecnico di Torini, 2022.

- Katariina Muru, Tarkade vaktsineerimissertifikaatide pilootprojekt, MSc, juhendaja Ingrid Pappel, kaasjuhendaja Eric Jackson, **Richard Dreyling**, Tallinna Tehnikaülikool, 2021.

- Olena Roraff, Küberjulgeoleku poliitika satelliiditööstuses: juhtimisväljakutsed ja lahendused, MSc, juhendaja Paul Liias, kaasjuhendaja Eric Jackson, **Richard Dreyling**, Tallinna Tehnikaülikool, 2021.

- Snowbar Akbar, Vestlusrobot, mis abistab krooniliste pikaajaliste haigustega laste vanemaid, MSc, juhendaja Ingrid Pappel, kaasjuhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2021.

- Karl Andreas Sprenk, Tark koos ja noorte jaoks – Eesti omavalitsuste juhtumianalüüs, MSc, juhendaja Ingrid Pappel, kaasjuhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2021.

- Ganenthra Ravindran, Bitcoini turu efektiivsus: tõendid tõhusa turu hüpoteesist (EMH), BSc, juhendaja Simona Ferraro, kaasjuhendaja **Richard Dreyling**, Tallinna Tehnikaülikool, 2022.

Publikatsioonid

1. R. Dreyling, , T. Tammet, and I. Pappel, “Technology push in ai-enabled services: How to master technology integration in case of bürokratt,” *SN Computer Science*, vol. Future Data Science Engineering, no. 5, p. 738, 2024

2. R. Dreyling, E. B. Jackson, T. Tammet, A. Labanava, and I. Pappel, “Social, legal, and technical considerations for machine learning and artificial intelligence systems in government.,” in *ICEIS (1)*, pp. 701–708, 2021

3. R. Dreyling, J. Lemmik, T. Tammet, and I. Pappel, “An artificial intelligence maturity model for the public sector: Design science approach,” *TalTech Journal of Social Sciences*, vol. 14, no. 2, p. 16, Forthcoming

4. R. M. Dreyling, T. Tammet, and I. Pappel, “Artificial intelligence use in e-government services: A systematic interdisciplinary literature review,” in *International Conference on Future Data and Security Engineering*, pp. 547–559, Springer, 2022

5. R. Dreyling, K. McBride, T. Tammet, and I. Pappel, “Navigating the AI maze: Lessons from Estonia’s Bürokratt on public sector AI digital transformation,” *SSRN*, 2024

6. R. Dreyling, T. Koppel, T. Tammet, and I. Pappel, “Challenges of genai chatbots in public services: An integrative review,” *SSRN*, 2024

7. R. Dreyling, E. Jackson, and I. Pappel, “Cyber security risk analysis for a virtual assistant g2c digital service using fair model,” in *2021 Eighth International Conference on eDemocracy and eGovernment (ICEDEG)*, pp. 33–40, 2021

8. R. Dreyling, R. Erlenheim, T. Tammet, and I. Pappel, “Ai readiness assessment for data-driven public service projects: Change management and human elements of procurement,” *Human Factors, Business Management and Society*, vol. 97, no. 97, 2023

9. R. M. Dreyling III, T. Tammet, and I. Pappel, “Digital transformation insights from an ai solution in search of a problem,” in *International Conference on Future Data and Security Engineering*, pp. 341–351, Springer, 2023

10. E. Blake Jackson, R. Dreyling, and I. Pappel, “A historical analysis on interoperability in estonian data exchange architecture: Perspectives from the past and for the future,” in *Proceedings of ICEGOV’21 – the 14th International Conference on Theory and Practice of Electronic Governance*, pp. 111–116, ACM, 2021

11. E. B. Jackson, R. Dreyling, and I. Pappel, “Challenges and implications of the who’s digital cross-border covid-19 vaccine passport recognition pilot,” in *2021 Eighth International Conference on eDemocracy & eGovernment (ICEDEG)*, pp. 88–94, IEEE, 2021

12. A. Labanava, R. M. Dreyling III, M. Mortati, I. Liiv, and I. Pappel, “Capacity building in government: Towards developing a standard for a functional specialist in ai for public services,” in *International Conference on Future Data and Security Engineering*, pp. 503–516, Springer, 2022

13. A. Labanava, R. M. Dreyling, and A. Norta, “Potential of smart contracts in the pharmaceutical supply chain of belarus,” in *2022 IEEE 1st Global Emerging Technology Blockchain Forum: Blockchain & Beyond (iGETblockchain)*, pp. 1–6, IEEE, 2022

ISSN 2585-6901 (PDF)
ISBN 978-9916-80-199-4 (PDF)