Democratizing Open Data Through AI: The Case of Sweden

Proscovia Svärd
Proscovia.Svard@Sorbonne.ae
https://orcid.org/0000-0001-9959-4336
Associate Professor at Sorbonne University Abu Dhabi

Abstract

The chapter explores the democratization of open data through the lens of artificial intelligence (AI) and its implications for citizen participation and engagement. "Open government" initiatives aim to promote increased transparency and accountability in governance issues, boost innovation and create inclusive and efficient government institutions. Since Sweden has the oldest Freedom of Press Act and the Right to Access Public Records Act, its citizens are used to free access to government information. However current developments pushing the opening of government data, are different from it because they focus on the commercialization of government information/data to create an information market. Using the secondary research method, the article explores the intersection of artificial intelligence (AI) and open data within the public sector in Sweden, highlighting both the opportunities and challenges presented by these technologies. It discusses the barriers to open data utilization, such as privacy concerns, information quality, and technical challenges. AI can enhance the accessibility and reusability of government data while addressing ethical and social implications. The potential of AI to democratize open data is also examined. The article advocates an inclusive approach to open data to avoid a data-divide. By leveraging AI, the article posits that the public sector can improve service delivery and foster innovation, while also emphasizing the importance of transparency and citizen engagement in the open data movement.

Keywords

Open Data Democratization, Public Sector, Accessibility, Reusability and Artificial Intelligence.

1. Introduction

This chapter examines the democratization of open data in Sweden through the lens of Artificial Intelligence (AI). AI tools could promote the democratization of open data through effective records management regimes of the public authorities. Governments can harness AI tools to facilitate access and use of open data by the citizens. The technological advancements of the 21st century therefore present a significant opportunity for open data democratization (Awasthi and George, 2020). Fraunholz and Bauer (2023) contend that the development of machine learning (ML) models utilizing automatic data descriptions has facilitated the emergence of a data-first paradigm in artificial intelligence (AI), wherein the emphasis is placed on data engineering rather than the construction of models themselves. This shift is likely to promote a human-centric approach to AI, thereby simplifying the model development process and rendering AI technology more accessible to a wider array of users. This assertion is grounded in the increasing volumes of available data, the declining costs associated with information systems, and the enhancement of data-related skills among prospective users.

The global imperative for governments to enhance access to open government data has gained significant attention. Movements advocating for the Right to Information and Open Government Data are compelling governments to make their data accessible to citizens (Ubaldi, 2013). There is an increasing expectation for governments worldwide to operate with transparency, foster participatory governance, and engage in collaborative practices. Consequently, governments are implementing initiatives aimed at establishing what is commonly referred to as "open government." Such initiatives may encompass improvements to governmental websites for the dissemination of information, proactive publication of open data, and strategies designed to augment citizen participation in decision-making processes (Bowles et al., 2014).

For instance, the G8 governments have collectively agreed to adhere to a set of principles that facilitate the access, release, and reuse of data. These principles include the notion of open data as a default standard, the emphasis on quality and quantity, usability for all stakeholders, the release of data to enhance governance, and the provision of data to stimulate innovation (G8, 2013). Furthermore, the evolution of e-Government is recognized as a mechanism that can enhance governmental transparency, accountability, inclusivity, and the promotion of sustainable development. The advancement of e-Government involves implementing organizational changes and dismantling hierarchical structures in favor of process-oriented frameworks. This transition includes the deployment of information systems that facilitate the management of information and processes, the execution of change management strategies, and the provision of training for government officials to equip them with the necessary skills to navigate a dynamic organizational environment. Most data generated within governmental administration today arises from the creation and delivery of electronic services. Therefore, the extent of e-Government development directly influences the quality of government data that is made accessible to citizens (Layne and Lee, 2001).

Open government initiatives are designed to enhance transparency and accountability in governance, foster innovation, and establish inclusive and efficient governmental institutions. This evolution involves the dissemination of government information to the public. Public Sector Information (PSI) encompasses data derived from various sources, including geographic information systems (GIS), land registries, public meteorological services, and

other forms of information produced by governmental entities. The term "open data" refers to the data released by public sector organizations in a manner that is both legally and technically reusable. This initiative aims to address a range of social, economic, and political challenges by facilitating innovation and economic development, promoting political accountability and democratic engagement, and improving the efficiency of the public sector (Zuiderveen Borgesius et al., 2016). Furthermore, Section 4 of the European Public Sector Information Directive 2013/37/EU, which amends the original PSI Directive of 2003, articulates the importance of these principles:

“Open data policies which encourage the wide availability and re-use of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints, and which promote the circulation of information not only for economic operators but also for the public, can play an important role in kick-starting the development of new services based on novel ways to combine and make use of such information, stimulate economic growth and promote social engagement.”

Sweden is notable for having established the oldest Press Act in 1776, which enshrines the right of citizens to access public records. Historically, Swedish citizens have enjoyed a robust framework for accessing governmental information (Kassen, 2017). Despite this long-standing commitment to transparency, in 2010, Sweden was required to implement the European Public Sector Information (PSI) Directive, a mandate from the European Union. The original Public Sector Information Directive, 2003/98/EC, which focused on the re-use of government information, was enacted in 2003; however, it has since been superseded by the Open Data Directive (Directive (EU) 2019/1024), which was enacted on July 16, 2019, which serves as a legislative framework aimed at promoting the accessibility and re-use of public sector information within the European Union (European Union, 2019). In alignment with this directive, the Swedish Public Sector Information (PSI) Directive facilitates the re-utilization of open government data for any interested party, thereby encouraging its exploitation and dissemination for diverse purposes. The primary objective of this initiative is to stimulate innovation and foster the development of new electronic services and products.

One could contend that prior to the current mandate for proactive publication of datasets on the Swedish open data portal (https://www.dataportal.se/), access to government information was predominantly limited to individual documents rather than comprehensive datasets. Recent trends advocating for the openness of government data emphasize its commercialization, aiming to establish a market for information. Within this framework, citizens are conceptualized not merely as consumers of information, but as developers of innovative electronic services utilizing the raw materials provided by the public sector in the form of open data.

The democratization of open data necessitates collaboration between the Swedish government and local administrations to establish projects or platforms and IT infrastructure that facilitate the effective exploitation of open data. Such initiatives would enable local communities to acquire essential skills and knowledge. Svärd’s pilot study, conducted in 2018 and published in 2021, revealed that a significant number of ordinary individuals interviewed lacked an understanding of the distinctions between the Press Act, which ensures citizens' access to public records, and the PSI directive, nor were they cognizant of their rights to utilize open data. The increasing accessibility of data is transforming our comprehension and

interaction with the surrounding world. The concept of "democratization of data" embodies the notion that a broader demographic, extending beyond experts and large organizations, can access, utilize, and derive benefits from data. Open data serves to empower individuals and communities, enabling them to make informed decisions based on the information available. Furthermore, the accessibility of open data can be enhanced through the application of artificial intelligence and digital platforms, thereby promoting greater transparency, accountability, and participation across various sectors, including government, healthcare, and education. This article is structured to include an introduction, a delineation of the research problem, a literature review, methodology, research findings, a discussion, and a conclusion.

2. Research Problem

To advance the democratization of open data, it is essential to facilitate citizen engagement and participation. Existing literature has explored both the positive and negative narratives surrounding open data; however, further research is necessary to deepen our understanding of its democratization, particularly considering recent advancements in artificial intelligence technologies that hold the potential to help societies realize this objective (Dander, 2013; Zuiderwijk & Janssen, 2013; Melin, 2016; Svärd, 2018; Schwoerer, 2022). This indicates a pressing need to reassess current initiatives and to investigate innovative approaches. The effective implementation of democratic open data initiatives is impeded by a range of challenges identified in literature. These challenges include insufficient citizen awareness, inadequate information technology infrastructure, and a lack of technical skills necessary for leveraging open data. As governmental institutions seek to enhance transparency and foster citizen engagement, it is crucial to examine the mechanisms through which these administrations can promote effective collaboration with diverse stakeholders to establish an inclusive open data framework. Understanding the strategies employed by local administrations to raise public awareness and comprehension of open data is vital, as these strategies are instrumental in promoting civic engagement and ensuring that citizens can effectively utilize the available data. Furthermore, it is essential to investigate how local administrations integrate open data considerations into their planning processes, particularly during the data creation phase. This aspect is critical for ensuring that the data produced is relevant, accessible, and beneficial to the community. Local administrations must possess the requisite information infrastructure to support the dissemination of open data, which entails examining the technological and organizational capabilities necessary for effective data sharing. Additionally, they should be equipped with information technology platforms that facilitate citizen engagement in the utilization of open data. Such platforms play a crucial role in bridging the gap between data availability and public participation. However, this ideal scenario is not currently realized, as local administrations in Sweden have engaged with the implementation of the 2010 Swedish PSI directive in varied manners and have encountered differing financial constraints (Svärd, 2018).

3. Research Method

A research methodology constitutes the systematic framework governing the processes involved in the collection, analysis, and interpretation of data relevant to a specific research inquiry. In the present study, the researcher adopted a secondary data research method, which entails the examination of data originally gathered by another entity for an alternative primary

purpose. According to Johnston (2014), like any other research method, the secondary research method must conform to established research principles, akin to those applied in studies utilizing primary data. This method necessitates the formulation of a research question, the identification of an appropriate dataset, and a thorough evaluation of that dataset (Johnston, 2014). It is essential for researchers to explore the existing body of knowledge and pinpoint gaps in understanding related to the phenomenon under investigation. Stewart and Kamin (1993) contend that secondary data offers rapid and cost-effective solutions to numerous research inquiries, serving as a critical foundation for subsequent primary research. The process typically involves an extensive review of secondary sources, which may encompass various forms of data, including official statistical studies, often conducted by governmental agencies, as well as scholarly articles and technical reports. The authors further argue that the proliferation of information, combined with a lack of standardized guidelines for the utilization of secondary sources, has led to the underutilization of this methodological approach.

Both Johnston (2014) and Stewart and Kamin (1993) advocate for the employment of the secondary data method due to the extensive array of available information sources. Johnston (2014) underscores that a critical component of secondary data analysis involves the application of theoretical frameworks and conceptual skills to effectively leverage existing data in addressing a research question. In the context of this study, the researcher identified a significant research problem aimed at elucidating the parameters necessary to achieve the objectives of open data initiatives and their democratization using artificial intelligence. A comprehensive search was conducted across databases such as Google Scholar, Emerald, and ScienceDirect, employing specific keywords to identify relevant scholarly articles. This method facilitated access to pertinent scientific literature, thereby informing the literature review section, which examines various dimensions of the research problem.

The concluding phase of the research involved a comprehensive analysis and evaluation of secondary data. Stewart and Kamin (1993) delineate a systematic evaluation framework that encompasses the following steps: (a) articulating the purpose of the study; (b) identifying the entity responsible for the collection of data; (c) determining the specific information that was gathered; (d) documenting the timeframe during which the data was collected; (e) comprehending the methodology employed in the data acquisition process; and (f) evaluating the consistency of information obtained from a single source in relation to data from other sources. The researcher meticulously adhered to these prescribed steps while analyzing the relevant literature pertinent to the study. Furthermore, the AvidNote Research tool was employed judiciously throughout the research process.

4. Research Findings

To rigorously address the research problem, the author undertook a thorough examination of the existing scientific literature by leveraging various academic databases, including Emerald, Google Scholar, and ScienceDirect. The literature search was guided by a set of targeted keywords, namely "open data and democratization," "open data and artificial intelligence," "open data initiatives and citizen participation," and "open government initiatives." This methodological approach facilitated the identification of relevant scholarly articles pertinent to the study.

4.1 Democratizing Open Data

Promoting good governance and inclusive development and access to government information is of crucial importance. Bowles et. al. (2014) posited that in the UK and in the US, the open data initiatives focus on data that is used by people in their role as consumers and not voters. Quoting Anthony Downs (1957) they contended that the Economic Theory of Democracy dictates that people have four distinct demands and these include:

- As consumers, they search out data that helps them make purchases;
- As workers, they want news that helps them do their jobs;
- Some information is simply enjoyable to consume, thereby satisfying a demand for entertainment and
- As voters, people might benefit from learning more about candidates and issues as elections draw close.

The above stated distinct information demands are also well reflected in section 4 of the 2013/37/EU.

In her 2024 analysis, Meng explored the concept of "data democratization," which refers to the increasing accessibility of data to a broader audience beyond merely experts or large organizations. This notion emphasizes the importance of making data available to all individuals, thereby fostering an inclusive environment where diverse stakeholders can utilize and benefit from information. Meng posited that the democratization of data is intrinsically linked to the advancement of democratic principles within society. She argued that open data has the potential to empower individuals and communities, enabling them to make informed decisions grounded in accessible information. Furthermore, this accessibility can enhance transparency, accountability, and civic participation across various sectors, including government, healthcare, and education. This is important because when more people can access and understand data, they can use it to inform their decisions, engage in discussions, and hold authorities accountable. Data democratization can strengthen democracy in several ways. First, it enables citizens to make informed choices, whether in voting or in community decision-making. When people have access to relevant data, they can better understand the issues at hand and advocate for their interests. Second, it can promote transparency in government and institutions, as citizens can scrutinize data and demand explanations for policies and actions. This increased transparency can lead to greater trust in democratic processes. Additionally, the article highlights that data democratization can help to empower marginalized communities whose voices might often be overlooked in traditional political discourses. By providing them with data, these communities can better represent their needs and demands, thus enhance their participation in democratic processes. In conclusion, the article emphasized that data democratization is not just about making data available; it is about fostering a culture of engagement where citizens actively use this data to influence and shape their communities and governance. Though optimistic about the potential benefits of democratizing data, she cautioned that we must be mindful of the risks involved such as privacy, data security, and the potential for misuse of information.

The authors, Lefebvre, Legner, and Fadler (2021), discussed the concept of data democratization (DD) and its significance in organizations, particularly in the context of increasing data generation and the underutilization of this data due to access and skill limitations among employees. They argued that while data is recognized as a strategic asset,

many companies struggle to leverage it effectively because non-specialists often lack both access to data and the necessary analytical skills. The authors were of the view that to achieve data democratization, organizations must provide broader access to data and tools for a wider audience, particularly non-specialists. They highlighted the importance of data catalogs as platforms that can help break down data silos and facilitate easier access to data across the enterprise. Democratization efforts should be tailored to the specific needs of different user groups and should consider the unique goals and contexts of individual organizations. They also explored the relationship between organizational culture and data democratization, suggesting that a culture that encourages information sharing and diversity is crucial for effectively utilizing data-driven insights. They proposed that democratization is not a one-size-fits-all approach; rather, it requires a nuanced understanding of the varying needs of employees and the specific objectives of data initiatives. Overall, the authors advocated for a strategic shift in how organizations approach data access and usage, encouraging them to empower employees at all levels to engage with data actively. This empowerment can lead to more informed decision-making and ultimately enhance organizational agility and effectiveness in leveraging data for competitive advantage.

Schwoerer (2022) aimed to understand whether the available data are sufficient for empowering citizens and if they meet the needs, interests, and activities of the diverse citizenry that the city serves. She therefore discussed the role of open government data (OGD) as a tool for promoting democratic values such as transparency, accountability, and citizen participation. While OGD initiatives have expanded globally since their introduction, research indicates that they have not fully realized their potential in enhancing democratic outcomes. The study identifies several challenges to the democratization of open data such as, the diverse roles of stakeholders and the specific data needs of different groups. She argued that this complexity could hinder the effective implementation and management of OGD programs. The author further highlighted the disparity in power among individuals and organizations that shape what data is made available, which can lead to unequal access and influence over the use of data. Schwoerer (2022) emphasized that merely disclosing government information is insufficient for meaningful citizen engagement. Data must be relevant and useful for fostering active participation in democratic processes. The study concludes with recommendations for cities to enhance citizen participation by increasing the availability of citizen-relevant data and improving the usability of OGD, ultimately advocating for a more inclusive approach to open data initiatives. She concluded that open data is still concentrated in the hands of private sector firms, researchers affiliated with well-resourced institutions, and established non-profits. These groups often possess more power and resources, allowing them to leverage open government data (OGD) to further their interests, while typical citizens tend to engage less with the available data. This disparity raises concerns about the inclusive and democratic potential of open data initiatives. Governments therefore must work hard to bridge the growing data divide.

Van Deursen and Van Dijk (2010) posited that the term digital divide originally referred to gaps in computer access but when computers got diffused in society, it shifted to include Internet access. They were of the view that Internet skills are critical to accessing information made available and that lack of such skills exacerbated existing inequalities. McCarthy (2016) argued that the criticism on big data divide is like that levied against the digital divide. The big data divide is about those who have access to and own large-scale distributed datasets and those who do not. Espinosa et. al. (2015) stated that non-experts are unable to take advantage of

big data like data mining experts do, which creates the big data divide. Data mining is a prerequisite if one is to discover the implicit knowledge patterns and richer insights into the data and it requires expert knowledge. A study that was published by the Swedish Local Fiber Alliance revealed that one million Swedes are digitally excluded in a country with the fastest Internet and advanced broadband developments. The factors that cause exclusion are; inequality, broadband access, education, motivation, user ability and access to computers. The groups that were affected by the digital divide included the elderly, people born outside Europe and groups with low incomes (The Local on the 27th of October 2016).

Graves and Hendler (2013) confirmed that numerous governmental organizations have adopted Open Government Data (OGD) policies aimed at enhancing public access to governmental data. This data encompasses a wide array of domains, including financial and ecological information. While these initiatives frequently report anecdotal evidence of increased efficiency and cost savings within governmental operations, the full spectrum of potential applications for OGD remains largely unexplored. They posited that a significant segment of the population stands to benefit from the utilization of OGD; however, they face barriers that prevent them from effectively engaging with this data. The primary obstacle is a pervasive lack of expertise and technical knowledge necessary for the collection, processing, integration, and interpretation of such data. To address this challenge, they advocate for the implementation of visualizations to enhance accessibility and comprehension of large datasets. Visualizations serve as an effective tool for simplifying the understanding and communication of complex information.

Evans and Campos (2013) discussed the complexities and obstacles associated with citizen engagement in open government initiatives, particularly in the context of data accessibility. The authors posited that while government agencies have made significant investments in promoting transparency and participation through open data, there remains a notable gap in effectively engaging citizens, focusing on understanding user needs, evaluating formats and venues for providing information to citizens, and examining the nature and scope of public demand for information and expectations for transparency. One of the primary challenges they identified was the limited understanding that agencies have of user needs. This disconnect leads to difficulties in providing information that is not only accessible but also relevant and comprehensible to the public. The authors further argued that the tension between supplying comprehensive data sets and catering to citizens who may lack the necessary background knowledge complicates efforts to enhance public engagement. Additionally, they pointed out that agency staff often lack the skills to repurpose data for public education, which can result in misinterpretation and manipulation of information. This raises concerns about the relevance of the data to public inquiries and the overall effectiveness of open government initiatives. The authors also noted that the integration of technology in government-citizen interactions is fraught with challenges. For instance, social media interactions often lack established guidelines, making it difficult for agencies to effectively manage citizen input. Furthermore, while agencies are expected to adopt open government principles, there is a lack of clear guidance from the Executive Office of the President, leading to inconsistent implementation across agencies. Overall, the authors underscore that simply providing data is insufficient to guarantee citizen participation. The lack of context around the information, the need for better feedback mechanisms, and the necessity for agencies to align their technological tools with strategic goals are all critical factors that influence the success of open government initiatives.

4.2 Challenges of Open Data

Zuiderwijk, Susha, Charalabidis, Parycek and Janssen (2015) concluded that despite the fact governments around the world are undertaking initiatives to open up government data, there is still paucity in research and practice regarding the success of such initiatives. Their research identified context-specific critical success factors for open data publication and use. They posited that some of these identified factors are universally applicable while others are context bound. They also confirmed that there is currently no holistic framework of critical success factors in relation to open data publication and use. The identified critical success factors included;

- Legislation, regulation and licenses;
- Strategy and political support;
- Management support and publication processes within governmental agencies;
- Training of and support for civil servants;
- Evaluation of the open data initiative;
- Sustainability of the open data initiative;
- Collaboration;
- Open data platforms, tools and services;
- Accessibility, interoperability and standards; and
- Data stewardship and the development of a management plan for it. (Stewardship here refers to assuring accuracy, validity, security, management, and preservation of information holdings).

Zuiderwijk and Janssen (2013) posited that the negative stories of open data are not well emphasized. Open data primarily benefits individuals and entities that are already in positions of privilege, specifically those who possess access to open data infrastructures, as well as the requisite hardware, software, financial resources, educational opportunities, and skills. This assertion is corroborated by Dander (2013), who contended that open data remains predominantly concentrated within the domains of governmental bodies, commercial enterprises, and academic institutions. Furthermore, empirical research conducted by Melin (2016) in a Swedish municipality substantiated this perspective, revealing that the engagement with open data initiatives was largely driven by a coalition of politicians, civil servants, and information technology professionals. Zuiderwijk and Janssen (2013) argued that in the formulation of open data policies, governmental entities seek to promote and facilitate the dissemination of government-generated data while also reaping the benefits associated with its utilization. Presently, a diverse array of open data policies exists across multiple levels of government; however, there remains a notable deficiency in systematic and structured research addressing the issues encompassed by these policies, their underlying intentions, and their tangible impacts. The findings indicated that the prevailing policies tend to be predominantly inward-focused. They suggested that enhancements to open data policies could be achieved through collaborative efforts with external organizations, an emphasis on the policy's impact, the encouragement of open data utilization, and the establishment of a culture that integrates data publication into routine operational processes. The insights garnered from this research may significantly contribute to the formulation of new open data policies and the refinement of existing ones.

Kassen (2017) asserted that Sweden is emerging as a global leader in the promotion of open data, having initiated a diverse array of platforms within this domain. He argued that open data serves as a potentially limitless reservoir of information, which can be harnessed by independent developers and technically proficient citizens to create innovative applications. He emphasized the importance of civic engagement in fostering broader community participation in open data initiatives, which, in turn, can generate public value and enhance both knowledge and financial resources within the community. He identified independent developers as pivotal stakeholders in the open data landscape, given their technical expertise and capacity to address local challenges through entrepreneurial ventures. Furthermore, Kassen posited that hackathons represent a valuable opportunity for networking among developers, designers, coders, and professionals in information and communication technology.

The Swedish Open Government Partnership Plan for the years 2019-2021 is reported to have been developed through a collaborative process involving multiple stakeholders. In 2018, a series of dialogue meetings and in-depth interviews were conducted with representatives from various organizations, including Sweden's study associations, digital civil society entities, the Church of Sweden, the Red Cross, and the Swedish Sports Confederation (The Swedish Government, 2019). These dialogues and workshops facilitated the collection of recommendations from civil society regarding the implementation of measures aimed at promoting the publication, matching, and utilization of open data.

In accordance with insights provided by a digitalization strategist cited in the Action Plan, the objectives delineated therein were successfully achieved in 2018. The findings from this initiative reveal a critical necessity for user-friendly interfaces that facilitate effective engagement with open data by civil society. Sweden's dedication to the Action Plan exemplifies its intention to promote co-creation with civil society within the context of open data initiatives. Furthermore, the Plan indicates that Sweden attained the second position in the Digital Economy and Society Index, as published by the European Union in 2019. This ranking is attributed to the digital competencies of its populace, patterns of internet usage, the integration of digital technologies, and advancements in digital governance. However, it is important to note that, according to the European Portal's Open Data Maturity in Europe report from 2018, Sweden was positioned 22nd in terms of open data maturity. This discrepancy underscores the paradox that, despite Sweden's prominent status in the development of e-government, it maintains a relatively low ranking in the domain of open data.

The Action Plan also highlighted that investments in open data required putting in place a public administration that would create attractive, citizen-centred solutions in partnership with the citizens and civil society. This would ensure continued and long-term collaboration between the public sector and civil society. It was confirmed that 52 per cent of Swedish agencies, municipalities and county councils published open data but that more municipalities than government agencies lacked the necessary skills and common standards. The plan also stated, “Much of the information that could be made available as open data is already digitized, and virtually everyone in the potential target groups is digitally competent and connected.” (The Swedish Government, 2019, p. 5). The Plan contains some recommendations such as:

- Strategic investments to boost co-creation with civil society
- Clear prioritization of a focus on open data
- Implementing thematic investments in open data

The need to draw up a national action plan, vision and mission and clear goals on open data.

Melin (2016) conducted an analysis of the myths and realities surrounding open data initiatives at the local government level in Sweden. He posited that the financial investments associated with these initiatives are substantial, and that the anticipated benefits typically include enhanced decision-making, the development of innovative products and services, as well as increased transparency and accountability. However, Melin critiqued the prevailing expectations surrounding open data, suggesting that they are often presented in an uncritical, overly simplistic, and idealized manner. In a subsequent study, Svärd (2018) explored the implementation of the Public Sector Information (PSI) directive within two municipalities in Sweden. Her research revealed significant disparities in the resource capacities and levels of e-government development between the municipalities, which subsequently influenced their respective implementations of the PSI directive. Notably, the larger municipality, equipped with a more substantial budget, successfully implemented the PSI directive and made a variety of data sets accessible on its website. In contrast, the smaller municipality, constrained by limited financial resources, was only able to publish a few documents. Svärd argued that for the PSI initiative to realize its democratic potential and effectively serve all citizens, municipalities must possess comparable capacities. Furthermore, she emphasized that the provision of high-quality Public Sector Information necessitates the adoption of a records and information continuum approach by municipalities. This perspective highlights the critical importance of proactive and comprehensive information management to facilitate the pluralization of data across diverse contexts.

Further, Svärd and Borglund (2022) examined the challenges faced by Swedish government agencies in the publication of open data, particularly in relation to compliance with the Public Sector Information (PSI) law. They highlighted the importance of a well-functioning electronic archive (e-archive) for the proactive publication of open data, as mandated by the PSI. Despite some agencies publishing open data, the study revealed that this is often done in a way that creates barriers to re-use, primarily due to a lack of interoperable information management systems and technical infrastructure. Key challenges identified included the need for better prioritization of open data publication, the fear of violating data protection regulations, and the limited understanding of the benefits of open data among agency staff. The authors concluded that without a comprehensive e-archive, the potential for effective publication and re-use of government information remains compromised. The study underscored the need for common specifications to enhance interoperability and facilitate information sharing, which have not been fully adopted by the agencies due to existing independent solutions.

Lassinantti, Bergvall-Kåreborn and Ståhlbröst (2014) explored open data initiatives in two Swedish municipalities, Stockholm and Skellefteå, using a qualitative cross-case study approach. Their findings revealed two distinct interpretations of open data: one as a platform for techno-economic growth (Stockholm) and the other as a platform for co-created societal growth (Skellefteå). These differing approaches reflect the municipalities' contextual challenges and interests, leading to diverging paths towards either open data or open government realization. The challenges faced by the municipalities in democratizing open data

included: Technical and Infrastructural Issues: Stockholm's initiative highlighted the need for digitalization and infrastructure improvements to support a growing city, which complicates the effective release and use of open data. Strategic Management Decisions: In Stockholm, a top-down approach limited engagement from broader community stakeholders, as strategic management decided what data to release without necessarily considering the needs of all potential users. Target Audience Focus: Both municipalities had different target groups for their data releases. Stockholm primarily focused on developers and companies, which may have excluded other community members from utilizing the data. Coordination with Events: The opening up of data in Stockholm was often tied to specific events like hackathons, which may hinder a sustainable and continuous engagement with the data. There was also a problem of balancing short-term and long-term strategies. Skellefteå's holistic approach seemed to slow down the process but aimed at a comprehensive long-term strategy for open government, while Stockholm's focus on immediate techno-economic growth could lead to a fragmented understanding of openness. Overall, the article emphasized that the realization of open data initiatives was influenced by local contexts, management strategies, and the relationship with data users, presenting both opportunities and challenges in democratizing access to public data.

4.3 Artificial Intelligence in Promoting Open Data Democratization

Guerrero et al. (2024) posited that systems driven by artificial intelligence (AI) possess the potential to enhance the management of public records. Such advancements would consequently streamline the publication of open data, allowing governmental administrations to strategically plan for the dissemination of this information prior to its creation. Given that public records constitute a significant component of open data, the application of AI technologies may be instrumental in facilitating data publication through the optimization of records management processes, including classification, storage, retrieval, and disposition. They identified four key tasks associated with public records management that could benefit from the integration of AI are outlined as follows:

1. **Enhanced Efficiency and Organization:** AI possesses the capability to automate essential processes such as the classification, indexing, and tagging of public records This automation not only liberates the time of records managers and archivists for engagement in more complex responsibilities but also ensures a higher degree of consistency in record-keeping practices.
2. **Improved Search and Retrieva**l: The application of AI can facilitate the analysis of extensive datasets of public records, enabling the identification of pertinent information in response to user queries. Such capabilities may significantly enhance the efficiency with which citizens and governmental entities can access the information they require.
3. **Transparency and Accountability:** AI technologies can be employed to detect and redact sensitive information within public records prior to their dissemination, thereby promoting transparency while safeguarding individual privacy Moreover, AI systems can monitor the utilization of public records thereby facilitating greater accountability of government agencies in their operations.
4. **Data Insights and Innovation:** AI-driven analytical frameworks can be utilized to scrutinize public records for emerging trends and patterns, yielding valuable insights for policymakers and service providers. The resultant data can subsequently inform the development of novel and enhanced e-government services.

They concluded that the integration of AI into public records management presents a transformative opportunity to improve efficiency, accessibility, accountability, and innovation within the realm of e-government.

Dessimoz and Thomas (2024) discuss how artificial intelligence (AI) has the potential to make access to data and knowledge more equitable and widespread, a process referred to as "democratization." Traditionally, access to important information, such as research data or advanced knowledge in fields like medicine or technology, has been limited to a select group of people, often those with advanced education or significant resources. This exclusivity creates a knowledge gap between different segments of society. AI can help bridge this gap in several ways. For example, AI tools can analyze vast amounts of data quickly and effectively. This means that even individuals who may not have technical expertise can gain insights from complex data sets. Additionally, AI technologies can be designed to be user-friendly, allowing a broader audience, including those without specialist knowledge, to engage with and understand important information. AI can facilitate personalized learning experiences, tailoring information to the individual's needs and comprehension levels. This targeted approach allows more people to learn about difficult subjects at their own pace. By making data and knowledge more accessible, AI has the potential to empower individuals, foster innovation, and promote social equity. However, the authors also note some challenges because they argue that it is crucial to ensure that AI systems are unbiased and fair. This is because, if the underlying data used to train AI is skewed, the outcomes could disadvantage certain groups instead of helping to level the playing field. In summary, the authors highlight that AI can enable more people to access important information and engage with it meaningfully, potentially leading to a more informed and equitable society.

Hussain, Kumar, and Gandomi (2024) discussed the challenges and opportunities associated with the use of small datasets in artificial intelligence (AI). The authors argued that while big data is often prioritized in research and public discourse, small datasets are equally prevalent and can provide significant advantages. They highlighted the real-world constraints faced by many organizations, such as budget limitations and the complexity of data analysis, which make small data a more practical choice in certain contexts. The authors emphasized that high-quality small datasets could yield better insights than low-quality large datasets, citing historical examples to support this claim. They argued that the current focus on big data exacerbates inequality in AI, as tech giants have easier access to vast datasets, leaving smaller organizations at a disadvantage. This creates a widening gap between the "AI haves" and "have-nots," which the authors sought to address by exploring deep learning (DL) techniques optimized for small datasets. The authors advocated for the development of AI models that require less data, which could reduce the need for extensive personal information collection and alleviate privacy concerns associated with big data. The authors also highlighted the potential for AI to advance in areas where data is scarce, such as detecting rare events or conditions. By promoting the use of small datasets, the authors argued that it is possible to democratize AI, enabling smaller organizations to compete more effectively and foster a more equitable AI landscape. In conclusion, the article calls for a shift in focus towards DL models that can effectively utilize small datasets, thereby encouraging innovation and inclusivity in AI applications across various sectors.

5. Discussion

The study has delved into the critical issue of open data democratization, emphasizing the necessity of citizen engagement and participation to achieve this goal. It highlights the existing research on both the successes and failures of open data initiatives, pointing out that while there is a growing body of literature, significant gaps remain, particularly concerning the democratization of data in the context of emerging artificial intelligence technologies. The research identifies several barriers to effective implementation, including a lack of citizen awareness, insufficient IT infrastructure, and the technical skills necessary for utilizing open data effectively. The findings underscore that merely making data available is not enough to ensure public engagement. The research indicates that the context surrounding the data, feedback mechanisms, and alignment of technological tools with strategic objectives are vital for the success of open government initiatives. Challenges of open data democratization are multifaceted, involving diverse stakeholder roles and specific data needs across different groups. The disparity in power dynamics can lead to unequal access to data, which could undermine the very goals of democratization. A more inclusive approach to open government data (OGD) initiatives will enhance the relevance and usability of data and is crucial for fostering active citizen participation.

Sweden faces several challenges in its efforts to democratize open data, despite its historical commitment to transparency and access to government information. Key issues include technical and infrastructural limitations that hinder effective data release and use. Strategic management decisions often adopt a top-down approach, limiting broader community engagement and failing to consider the diverse needs of potential data users. The focus on specific target audiences, such as developers and companies, has excluded other community members from utilizing open data. Additionally, reliance on events like hackathons for data engagement can lead to a lack of sustainable, continuous interaction with the data. Interoperability challenges should be addressed to facilitate information sharing among agencies, which has not been fully realized due to existing independent solutions. Overall, the challenges are compounded by a lack of citizen awareness, citizen-oriented IT infrastructure for data exploitation, and the necessary technical skills, indicating a need for strategic investments and collaborative efforts between the public sector and civil society to foster a more inclusive open data environment.

Artificial Intelligence (AI) has the potential to significantly democratize open data, particularly within the public sector. The democratization of data is critical for fostering transparency, accountability, and citizen engagement in governance. One of the primary ways AI can facilitate this democratization is by enhancing the accessibility and reusability of government data. Traditional barriers to data access often stem from the complexity of data sets, which can be overwhelming for individuals without technical expertise. AI tools can analyze vast amounts of data quickly and effectively, transforming complex data into understandable insights. This capability allows non-experts to engage with data meaningfully, empowering them to make informed decisions based on available information. Further, AI can streamline the management of public records, optimizing processes such as classification, storage, retrieval, and publication of data. By automating these tasks, AI can enhance the efficiency of data dissemination, ensuring that important information is readily available to the public. This proactive approach to data publication not only increases transparency but also fosters a more

participatory governance model where citizens can actively engage in decision-making processes. AI technologies can also be designed to be user-friendly, making it easier for individuals to access and interact with data. Personalized learning experiences can be created, allowing users to engage with information tailored to their comprehension levels. This targeted approach can help bridge the knowledge gap that often exists between different segments of society, promoting social equity and innovation. Therefore, while AI holds transformative potential for democratizing open data, careful consideration must be given to the design and implementation of these technologies to avoid perpetuating existing inequalities.

6. Conclusion

In conclusion, the democratization of open data necessitates an enhancement of citizen awareness regarding their rights and the accessibility of open data resources. It is imperative that individuals are equipped with the knowledge to effectively access and utilize this information. Additionally, the establishment of a robust information technology infrastructure is essential, alongside the implementation of training programs designed to bolster the technical competencies of both citizens and agency personnel. This will facilitate more meaningful engagement with open data initiatives. It is crucial to develop clear guidelines and best practices for governmental agencies to adhere to when operationalizing open government principles. Such measures will promote consistency and effectiveness across various governmental entities. To ensure that datasets align with user needs, collaboration between government agencies and community organizations is necessary. This partnership should focus on identifying the specific data requirements of diverse groups, particularly those from marginalized communities, thereby ensuring that the data disseminated is both relevant and beneficial. The integration of artificial intelligence technologies is also recommended to enhance access to and analysis of open data, thereby promoting transparency, accountability, and citizen participation in governance. AI has the potential to significantly contribute to the democratization of open data by improving accessibility, streamlining data management, and fostering greater citizen engagement. By harnessing the capabilities of AI, the public sector can not only enhance service delivery but also contribute to the creation of a more informed and equitable society. Nevertheless, ongoing research and inclusive methodologies are essential to address the challenges associated with AI-driven open data initiatives, ensuring that their benefits are equitably distributed across all segments of society.

7. Recommendations

Enhance Citizen Awareness and Education: governmental bodies should establish educational programs aimed at informing citizens of their rights concerning open data and methodologies for effectively accessing and utilizing available resources.

Improve Data Relevance and Usability: Open data initiatives should prioritize the identification and distribution of data that is pertinent and beneficial to citizens.

Foster Collaboration and Inclusivity: To mitigate the data divide, it is crucial to encourage collaborative efforts among various stakeholders, including government entities, civil society organizations, and the private sector.

Leverage Artificial Intelligence Responsibly: Governments should investigate the potential of artificial intelligence to enhance open data initiatives, while remaining vigilant about the risks of exacerbating existing inequalities.

8. Bibliography

Awasthi, Pranjal and George, Jordana J., "A case for Data Democratization" (2020). AMCIS 2020 Proceedings. 23. Available at: https://aisel.aisnet.org/amcis2020/data_science_analytics_for_decision_support/data_science_analytics_for_decision_support/23 This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL) (accessed 20 March 2025).

Bowles, N., Hamilton, J. T., & Levy, D. A. L. (2019). Introduction. In D. A. L. Levy, J. T. Hamilton, & N. Bowles (Eds.), *Transparency in politics and the media: Accountability and open government* (First edition, pp. xii–xxiii). I.B. Tauris. https://doi.org/10.5040/9780755694716

Dander, V. (2014). How to Gain Knowledge When Data Are Shared? Open Government Data from a Media Pedagogical Perspective. *Seminar.Net*, *10*(1). https://doi.org/10.7577/seminar.2377

Dessimoz, C., & Thomas, P. D. (2024). AI and the democratization of knowledge. *Scientific Data*, *11*(1), 268. https://doi.org/10.1038/s41597-024-03099-1

Espinosa, R., García-Saiz, D., Zorrilla, M., Zubcoff, J. J., & Mazón, J.-N. (2015). Enabling Non-expert Users to Apply Data Mining for Bridging the Big Data Divide. In P. Ceravolo, R. Accorsi, & P. Cudre-Mauroux (Eds.), *Data-Driven Process Discovery and Analysis* (Vol. 203, pp. 65–86). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46436-6_4

European Commission European legislation on the re-use of public sector information. Available at: https://wayback.archive-it.org/12090/20190328175642/https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information (accessed 11 January, 2025).

Evans, A. M., & Campos, A. (2013). Open Government Initiatives: Challenges of Citizen Participation. *Journal of Policy Analysis and Management*, *32*(1), 172–185. JSTOR.

G8. (2013). G8 Open Data Charter. Retrieved December 20, 2016 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/207772/Open_Data_Charter.pdf

Graves, A., & Hendler, J. (2014). A study on the use of visualizations for Open Government Data. *Information Polity*, *19*(1,2), 73–91. https://doi.org/10.3233/IP-140333

Guerrero, E., Svärd, P. Balugon, T, Saurombe, N., Hentonnen, P and Jacobs, L. (2024). The Africa-Europe Gaps of AI Regulations for Managing Public. In *Artificial Intelligence and Documentary Heritage*, Eds. Duranti, L and Rogers, C., SCEaR Newsletter 2024 Special Issue 2024. Available at: https://interparestrustai.org/assets/public/dissemination/SCEaRNewsletterSpecialIssue2024ArtificialIntelligence.pdf

Lassinantti, J., Bergvall-Kareborn, B., & Stahlbrost, A. (2014). Shaping Local Open Data Initiatives: Politics and Implications. *Journal of Theoretical and Applied Electronic Commerce Research*, *9*(2), 5–6. https://doi.org/10.4067/S0718-18762014000200003

Layne, K., & Lee, J. (2001). Developing fully functional E-government: A four stage model. *Government Information Quarterly*, *18*(2), 122–136. https://doi.org/10.1016/S0740-624X(01)00066-1

Lefebvre, H., Legner, C., & Fadler, M. (2021, December). Data democratization: toward a deeper understanding. In ICIS. *Forty-Second International Conference on Information Systems*, Austin.

McCarthy, M. T. (2016). The big data divide and its consequences. *Sociology Compass*, *10*(12), 1131–1140. https://doi.org/10.1111/soc4.12436

Melin Ulf. (2016). Challenges and Benefits in an Open Data Initiative – A Local Government Case Study of Myths and Realities. In *Innovation and the Public Sector* (pp. 111–122). IOS Press. https://doi.org/10.3233/978-1-61499-670-5-111

Meng, X.-L. (2024). *Data Democratization: An Ecosystemic Contemplation and Coordination*. https://doi.org/10.1162/99608f92.f2a44cc0

Rall, D, Fraunholz, T. and Bauer, B. (2023). AI-Democratization: From Data-first to Human-first AI. In *Proceedings of the Central European Conference and Intelligent Systems*, 34th CECIIS, September 20-22. Available at: https://www.proquest.com/docview/2876104622/fulltextPDF/3A78A99213104EDBPQ/1?accountid=146540&sourcetype=Conference%20Papers%20&%20Proceedings (accessed 21 March 2025).

Rather, I. H., Kumar, S., & Gandomi, A. H. (2024). Breaking the data barrier: A review of deep learning techniques for democratizing AI with small datasets. *Artificial Intelligence Review*, *57*(9), 226. https://doi.org/10.1007/s10462-024-10859-3

Schwoerer, K. (2022). Whose open data is it anyway? An exploratory study of open government data relevance and implications for democratic inclusion. *Information Polity*, *27*(4), 491–515. https://doi.org/10.3233/IP-220008

Stewart, W. D. and Kamins, A. M. (1993). Secondary Research: Information Sources and Methods, Second Edition. Sage Publications Inc, California.

Svärd, P. (2018). Public Information Directive (PSI) implementation in two Swedish municipalities. *Records Management Journal*, *28*(1), 2–17. https://doi.org/10.1108/RMJ-04-2016-0012

The Local Sweden. (2016, October 27). *A million Swedes are digitally excluded: Report*. The Local Sweden. https://www.thelocal.se/20161027/a-million-swedes-are-digitally-excluded-report

Ubaldi, B. (2013). *Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives* (OECD Working Papers on Public Governance 22; OECD Working Papers on Public Governance, Vol. 22). https://doi.org/10.1787/5k46bj4f03s7-en

Van Deursen, A., & Van Dijk, J. (2011). Internet skills and the digital divide. *New Media & Society*, *13*(6), 893–911. https://doi.org/10.1177/1461444810386774

Zuiderveen Borgesius, F., Gray, J., & van Eechoud, M. (2016). Open Data, Privacy, and Fair Information Principles: Towards a Balancing Framework. *Berkeley Technology Law Journal*, *30*(3), 2073–2131. https://doi.org/10.15779/Z389S18

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, *31*(1), 17–29. https://doi.org/10.1016/j.giq.2013.04.003

Zuiderwijk, A., Susha, I., Charalabidis, Y., Parycek, P., & Janssen, M. (2015). Open data disclosure and use: Critical factors from a case study. In P. Parycek & N. Edelmann (Ed.), *CeDEM 2015: Proceedings of the International Conference for E-Democracy and Open Government 2015* (pp. 197–208). Edition Donau-Universität Krems; DiVA. http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-50203