SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2025
ORGANIZATION AND MANAGEMENT SERIES NO. 225

POLISH LANGUAGE MODELS IN BUSINESS AND PUBLIC SECTOR: A STRATEGIC PERSPECTIVE

Roksana ULATOWSKA

Jagiellonian University; roksana.ulatowska@uj.edu.pl, ORCID: 0000-0003-3199-483X

Purpose: The aim of this article is to analyse the first two Polish language models (Large Language Models) from the point of view of strategic dimensions of implementing national (LLMs) in the business and public sector. The study outlines both the possible benefits of implementing LLMs in Polish organisations and the challenges of their implementation.
Design/methodology/approach: The research was conducted using a qualitative approach with three techniques - literature analysis, case study analysis and semi-structured in-depth interviews. Eight interviews were conducted with key persons responsible for the management and establishment and development of the two national language models.
Findings: Both models demonstrate different approaches in building, further developing and implementing national LLM models. Regardless, both play a strategic role in Poland's digital sovereignty. The results underline the importance of LLM not only in the context of technological innovation, but of building digital capacity in business and public administration.
Research limitations/implications: Although the results of the research provide many contested insights they have their limitations. The small sample size limits the possibility of generalisation. Future research should take into account the triangulation of other methods.
Practical implications: The practical implication is the ability to used presented framework to design, development and implementation of national LLMs. The research can inspire further research on the implementation of LLMs in public institutions and business in the context of sovereign technological solutions.
Social implications: The article can encourage critical reflection on the social role of national LLMs including ethical issues and those related to the development of digital competence.
Originality/value: The article offers an innovative contribution to the development of first national language models. The research shifts the analytical perspective towards strategic, institutional and governance dimensions. The value of the research is its interdisciplinary approach, which focuses on the development and implementation of LLMs in the national infrastructure.
Keywords: Large Language Model, Digital Transformation, Technological Sovereignty, Public-Business Sector Innovation, AI in Poland.
Category: Research paper.

Introduction

For several years, there has been an increase in both the availability and extent of use of large language models (LLMs) based on deep neural networks (Zubiaga, 2024). Although the development of LLMs began with Google's creation of the BERT model, it was only with the introduction of Chat GPT in 2022 that the development of large language models worldwide expanded (Ravichandiran, 2021; Wang et al., 2024). LLMs offer their users assistance to improve their every day work and personal lives. Still, their role, principles of operation and application are only fully understood by expert communities. In the literature on the subject, LLMs are often described in technical, specialised language and in a fragmented manner, making it difficult to access the knowledge about their practical application. Advanced artificial intelligence systems are characterised by being trained on extensive text corpora with parameters that often exceed hundreds of billions (Jiang et al., 2020). The increase in computational power and the use of multiple and diverse neural network architectures to handle large datasets, allow models to perform various tasks more accurately (Wickramasekara, Scanlon, 2024). LLMs processing natural language have become capable of generating human-like text and understanding complex prompts with high precision (Wang et al., 2024).

LLMs are changing the way users interact with computers, primarily because these models are able to analyse complex linguistic structures and increasingly take context into account and formulate human-like responses (Kietzman, Park, 2024). It is worth emphasising that despite these advances, their ability to map deeper cognitive processes is still limited. Large language models show some degree of sensitivity to representations of other people's mental states (so-called false belief reasoning), but the ability to do so is still limited and does not match the human ability to understand other people's beliefs. As the results of the Trott et al. (2023) study indicate, the linguistic statistics in the LLM training data alone are not sufficient to fully represent the complex aspects of human theory of mind (Carlson et al., 2013). While humans in a test based on the False Belief Task achieved a correctness of 82.7%, the largest LLM model tested achieved a score of 74.5%. Importantly, the model's behaviour did not fully explain people's behaviour.

Large-scale LLM language models are not only used by end users as private assistants but play an important role in wider social and professional contexts. Their versatility and new functionalities are influencing their use in many different disciplines (Retkowsky et al., 2024). They also have an increasingly strong impact on business as well as public service organisations (Chatham et al., 2024, Islam, Islam, 2024). Nowadays, although large, elaborate, multifunctional language models such as ChatGPT, Deepseek, Gemini dominate, it is possible to observe the development of smaller, including national models that seek to compromise between operational efficiency and accessibility (Żmigrodzki, 2025). Among others, the following have developed their models: Netherlands (GPT-NL), Italy (Modello Italia),

France (Mistral), Taiwan (TAJDE), Singapore (SEA-LION), Saudi Arabia (ALLaM), United Arab Emirates (JAIS), Spain (MarIA) (SIfHCAI, 2024).

National language models emphasise the recognition of language-specific ambiguity and semantic complexity. They are designed so that they can reflect the specific linguistic, cultural, legal and social characteristics of a country and can be particularly attractive not only for individual users but especially for business and the public sector, administration, health, education, the judiciary (SIfHCAI, 2024). National LLMs aim to strengthen digital sovereignty and deliver services in line with the local priorities (Pandey, 2025). They are becoming a key component of local innovation ecosystems (Chiarello et al., 2024; Ferraris et al., 2025). Regardless of whether LLM models are designed for national use or any other context, they still have a strong tendency to generate (or hallucinate) false content and information (Wickamasekara et al., 2024), often biased and reinforcing existing stereotypes (Gallegos et al., 2024). As Leong and Snug (2024) emphasise, the hallucinations that LLM models generate are structurally related to the existing social prejudices.

Given their rapid development and deepening integration with human and organisational activity and their increasingly widespread use in business and the public sector, it is particularly important to conduct research related to the implementation of language models. In particular, the development of national language models, which constitute a new alternative in management strategies and the digital transformation of organisations, is under-researched. Their implementation creates both opportunities and challenges for all organisations. Therefore, this study focuses on the following research questions:

RQ1.What are the key goals, characteristics, and directions of the future changes of the first two Polish LLM models?

RQ22. How can business and the public sector use models in strategic management?

By answering these questions, the article will provide a better understanding of the importance of national language models for the digital transformation of business and public organisations, their impact on management strategies, and the challenges of their implementation in Polish organisations.

Strategic applications of language models in the business sector

Business instantly recognised the potential and importance of large language models (LLMs) as important transformational and developmental tools. Initially, LLM models accelerated the execution of operations in organisations, streamlined decisions and enhanced interactions with users and customers. As time passed and large language models became more functional, companies began to benefit more and more from their use. Automation began to streamline workflows that had previously been performed manually, enabling companies to

focus resources on higher-value tasks (Fahland, 2025; Maarouf, Feuerriegel, Pröllochs, 2025). As a result of the development of LLMs, organisations very quickly recognised their potential. The new possibilities of integrating LLMs, especially in the context of business process management, have led organisations to seek to optimise their operational processes (Estrada-Torres et al., 2024).

The development of natural language processing (NLP) capabilities has translated into analysing vast amounts of structured and unstructured data and generating human-like responses. The introduction of advanced chatbots and virtual assistants that answer questions, handle requests and claims has not only improved customer interaction and communication (Dash, Bakshi, 2019) and helped to improve customer satisfaction (Janssen et al., 2020).

In addition to the deployment of chatbots for customer service processes, the automation of business processes has significantly advanced predictive analytics, knowledge management and strategic decision-making. The ability to analyse large amounts of information quickly, has allowed companies to identify trends, generate insights and make informed decisions that drive business growth. With the democratisation of the access to information, the use of extensive data and knowledge that can be applied for market analysis, idea generation and research initiatives has become more accessible and widespread (Chang et al., 2023; Szilágyi, Tóth, 2023; Fahland et al., 2025). As a result of these processes, business organisations have improved their operational efficiency, increased their quality of service, competitiveness and innovation and, basically, reduced costs (Chen et al., 2025).

The qualitative breakthrough that has occurred in the development of language models has opened up a shuttle of new practices and competences for business that were previously attributed only to human intelligence (Makridakis et al., 2023). Chiarello and co-authors (2024) recognise the increasingly complex and multidimensional role of language models and their application in the business environment. The high flexibility of LLM models is translating more and more into increased personalisation of services, supporting decision-making, increasing operational efficiency and fostering innovation across sectors. One of the processes undergoing the change under the influence of LLM is the area of human resources management. The models contribute to automatisation of HR policies creation, talent acquisition strategies, as well as improving recruitment and selection processes and building and personalising development plans for employees. Administrative processes - day-to-day tasks, communication or team collaboration in automated environments - are strongly transformed under the influence of generative models. LLMs support tasks related to document classification, product and service customisation. They enable the creation of diverse content, including marketing content, its evaluation and selection, as well as the development of content moderation algorithms (Chiarello et al., 2024).

Even more efficiently, they can search and customise information based on user queries. By improving communication and personalising responses, they can provide more tailored experiences for customers (Chen et al., 2024). Finally, they contribute to accelerating informed decisions in a rapidly changing digital landscape (Ilagan et al., 2024).

Research conducted by experts at the Speakleash Foundation (Żmigrodzki et al., 2025) shows that 49.3% of Polish companies surveyed use both open and closed LLM. Chat GPT is by far the most widely used, while Llama, Claude, Gemini or Mistral models are used much less frequently by Polish businesses. Entrepreneurs most often use models for knowledge extraction, creative writing, content summarisation and translation. Models are also helpful in automating communication, sales support or customer service (Żmigrodzki et al., 2025, p. 49).

Large Language Model in public organisation

Advanced AI systems, especially large language models (LLMs), are also making a significant transformation in the public sector (Yang et al., 2024). LLM models are part of a permanent feature of the management and service delivery of public organisations. Their use contributes to improving administrative efficiency and making services more accessible to citizens and increasing their civic engagement (Kleiman, Barbosa, 2024; Yoon, 2024).

Advanced automation and processing facilitate the deployment of virtual assistants and chatbots that improve interactions between citizens and government offices. LLMs provide detailed and relevant answers to citizens through intelligent query processing. As noted by (Pessanha et al., 2024), the use of LLMs allows for faster and more precise service to citizens, mainly by providing more contextual answers to the questions sent. This not only reduces waiting times for information but increases the responsiveness of services (Papageorgiou, 2024) and aptly supports users in the execution of administrative procedures (Gaetano, Diliberto, 2018; Decker et al., 2023). National LLM models, which are specifically trained on local data and adapted to local language, culture or law, become particularly effective in these tasks (SIfHCAI, 2024).

Another key application area for LLM in the public sector is data analysis and the generation of insights. Language models can analyse large data sets, identify patterns and predict trends. This translates into better operational optimisation and efficiency in government operations. Administrative decisions can be made in a more informed manner and respond more quickly and accurately to societal needs (Carullo, 2023; de Curtò,2024). The public health field uses language models to identify health problems to diagnose diseases and suggest potential treatment pathways and therapies, analyse patient data, and support clinical decisions (Thirunavukarasu et al., 2023; Huo at al., 2025).

Mamalis et al. (2024) highlight the role of LLMs in regulatory management, including the interpretation and analysis of complex legal texts. The integration of language models into legal systems allows for more efficient retrieval of relevant provisions, facilitates the interpretation of legislation and supports both officials and citizens in navigating legal complexities. For example, specialised LLM models can be used in administration to automatically classify documents, analyse legal precedents or support the drafting of laws.

For example, specialised LLM models can be used by the government to automatically classify documents, analyse legal precedents or support the drafting of laws. The provision of accurate information is enabled by the Retrieval-Augmented Generation (RAG) methodology, whereby LLMs interact with customised datasets (Jeong, 2023) and generate answers based on these. LLMs are also seen as a tool to support training and education of public sector employees. Models can help government employees to better process complex data sets and interpret the legal framework (Zhang, Chen, 2025) or onboarding processes (Ramli, 2025).

LLMs can also contribute to more efficient management of public resources, including the optimisation of energy consumption in administrative buildings and urban planning. By integrating distributed data sources and improving decision-making processes, artificial intelligence supports environmental policies and urban infrastructure management strategies (Zhang, Chen, 2025).

The use of LLMs is also finding its place in educational settings and higher education. It is transformative in that it influences the development of teaching materials, the personalisation of teaching and learning processes, and the design and conduction of research (Kasneci et al., 2023; van Dis et al., 2023; Wang et al., 2024;). The role of LLMs is not only limited to the automation of educational processes, but is hugely influential in changing the paradigm of teaching and learning (Choi, Abdirayimov, 2024; Chowdhury, Chowdhury, 2024).

With the growing role of LLM and the increasing number of applications in many areas, including strategic ones, there are ethical and practical challenges. It becomes necessary to balance the benefits and risks so that the implementation of LLM allows the potential of this technology to be realised.

Methods

The research used a qualitative approach, combining three complementary methods - literature review, case analysis and in-depth individual interviews. Qualitative methodology was used because of its ability to capture the complex, contextual aspects of the phenomena under study. This is particularly important in the exploration of the dynamic processes of the development of local language models. The flexibility of the qualitative approach makes it

possible to adapt the research tools to the specificities of the cases analysed and the themes that emerge (Groenland, Dana, 2019).

The literature review included academic publications, including articles and book chapters available in Scopus, Google Scholar and Web of Science databases, and industry reports on the development and use of LLM models in management in business and academic settings. One of the key elements of the research was a case study (Schwandt, Gates, 2018), in which two Polish Large Language Models (LLMs) were examined in detail. The analysis aimed to understand what needs are fulfilled by the two models on a national basis, to learn about their characteristics, applications in the business environment and the public sector, and further development. Their importance stems primarily from the fact that both models are the first of their kind developed for Polish users and citizens. The models open up new perspectives for digitisation, process automation and technological self-reliance of the country. Seven interviews were conducted in the final stage of the research. The selection of participants was purposive, in line with qualitative research, which allowed detailed and contextual information to be obtained from key people in the projects under study (Bouncken, Czakon, Schmitt, 2025). The respondents included both the creators (authors of the idea), those leading the implementation of both projects and those involved in the project team. The criteria adopted for selecting participants took into account their level of involvement, their role in the project structure and their access to decision-making and operational information related to the development of the LLM. This method of sampling enabled diverse perspectives to be gathered, covering both strategic and technical aspects of project implementation. Due to the expert and specialised nature of the subject matter under study, as well as the limited number of people actually involved in both projects, the research sample was limited. Nevertheless, its composition ensured access to in-depth knowledge, representing different levels of responsibility and perspectives within the projects analysed.

In the case of the first analysed model, the Bielik, four people representing the project team participated in the interviews, while in the case of PLLuM, there were three people, each representing a different consortium member. The interviews were conducted online and moderated based on a prepared scenario. The average interview time was at least 60 minutes, the longest of them took 90 minutes. The respondents were informed about the ethical principles of conducting interviews and assured of anonymity. Each of the interviewees agreed to consciously participate in the interview and its recording. Then the interviews were transcribed and analysed using the Maxqda program. The thematic analysis conducted allowed for systematic coding and identification of recurring patterns and thematic categories.

The triangulation of methods contributed to obtaining a more complete picture of the problems of Polish language models, characterisation of both models, development plans and strategic applications.

Results

Application Both the Bielik and PLLuM form families of models that have been designed for the needs of different user groups. They are designed to meet the goals and needs that flow from both business and the public sector. Both models have variants ranging in size from small models to multi-billion dollar models. The Bielik model was described in the context of potential applications in business, while the PLLuM model in the public sector was described as ready-made instruments to be implemented and deployed by other companies and organisations. However, the author of the paper emphasises that both models already have implementations in both areas. The Bielik has been implemented by, among others, the Poznan Supercomputing and Networking Centre, the EU Council, the Chancellery of the Prime Minister (Poznań Supercomputing Centre, 2025), while the PLLuM model has been implemented by the Comarch S.A. centre (Comarch, 2025). Table 1 presents a summary of the most important features and characteristics of the first two models (family) of Polish Large Language models.

Table 1.
Comparison of the Polish language models LLM - the Bielik and PLLuM

Features	Bielik	PLLuM
Year of creation	2024	2025
Lead time	app. 12 months	app. 11 months
Founding bodies	SpeakLeash Foundation Cyfronet AGH	Consortium: Wrocław University of Technology (consortium leader), NASK National Research Institute, Information Processing Center - National Research Institute (OPI PIB), Institute of Computer Science Polish Academy of Sciences, the University of Łódź, and the Institute of Slavic Studies of the Polish Academy of Sciences
Objectives of the model/ family of models	Support of machine learning research processes, training of a generative, pre-trained transformer model on the basis of the collected data	Support for the development of digital competences, the development of innovation in public administration and business
Number of parameters	from 1.5 to 11 mld (various architectures)	8 to 70 mld (various architectures)
Project funding	A bottom-up initiative, a project supported by the availability of Cyfronet's computing power sponsors (Helios, Athena), Patronite servis	Ministry of Digitalisation, Public funding

Cont. table 1.

Budget	No data	14.5 mln
Access	Free of charge, open source model	Models under fully open licenses; Models under restricted licenses (not allowing commercial use)
Training data	A collection of more than 3 terabytes of texts in Polish, exclusively from open, public sources (e.g. Polish Wikipedia, scientific articles under open licences, collections from the Common Crawl and data from public institutions	Trained on approximately 150 billion tokens of mainly Polish-language text, comprising internet data, resources acquired through contracts with publishers and instructional data created manually and distilled from other models. In addition, approximately 100.000 human grade samples were used to fine-tune the model preferences

Source: Own study, 2025.

The Polish generative artificial intelligence model Bielik was created by the SpeakLeash Foundation (Bielik, 2025). One respondent noted:

> *The Bielik was created by our own efforts, thanks to a grassroots social initiative. An initiative that is independent of any entities.*
>
> *(Respondent 2)*

The idea to create a Polish language model stemmed from a growing awareness of the importance of large models and the need to create a tool adapted to the Polish language - including the cultural context. The project team noticed that foreign models could not cope with the Polish language in data processing. The first goal of the project was to create a training dataset. As one respondent emphasised: We wanted to collect one terabyte of data. That was our goal. One terabyte of data, because the scientific studies said that models start just above one terabyte.

> *For us it was such a first goal to grab a terabyte, to produce and just see this scale of what we can achieve on one terabyte of the Polish data.*
>
> *(Respondent 1)*

The second goal the project set itself was to develop benchmarks that would allow the effectiveness of the model to be measured in the context of the Polish language. This ensured that the Bielik could not only be adapted to the Polish language, but also subjected to rigorous quality testing.

> *We started doing a language nuance benchmark and that was the other goal. Once we started thinking about the Bielik and training it, we suspected we would create the biggest set of Polish benchmarks.*
>
> (*Respondent 2)*

One of the most important assumptions in creating the model was to create one that could work without the need for global AI solutions. In addition, the developers of the Bielik wanted the model to be able to be implemented locally in organisations. The work started by the team

was initially experimental, as at the time there was a lack of publicly available information on how to create large language models, and many aspects of their construction remained a business secret. The decisive moment for the development of the project was the addition of the Cyfronet team from the University of Science and Technology, which provided the Bielik with the necessary computing power. This led to the creation of the first LLM model, initially in the 1 billion parameter version, which now runs on 11 billion parameters and is able to process and store as many as 32,000 tokens simultaneously. From the beginning, the project was open source, meaning that any user can download and use the model for commercial and research purposes for free.

Now the project team of the Bielik has begun to expand the project's objectives to include cultural and social aspects. The new version of the model, the Bielik 3.0, aims not only to support business, but also to protect and document elements of the Polish culture, including dialects and regional customs:

> *We now want to address the cultural strands more, because there are fewer and fewer people who can distinguish between white borscht and sour soup, who can recognise different customs.*
>
> *(Respondent 1)*

PLLuM (Polish Large Universal Model) was created as part of a consortium of six partners from the Wrocław University of Technology (leader), the NASK National Research Institute, the Information Processing Centre - National Research Institute (OPI PIB), the Institute of Computer Science of Polish Academy of Sciences, the University of Łódź and the Slavic Institute of Polish Academy of Sciences (Polish Ministry of Digitalisation, 2025). Its main objective was to create an open-source, Polish-language model in line with the responsible development of AI systems. In addition, the project prioritised the development of a prototype application of this model for public administration in the form of a Polish-language intelligent assistant.

> *The aim of the project was to create a model that is good for the Polish language, over which we first of all have control, we know how to train it, we know how to tune it for the task we want. We are the ones who say what the model can say, what it can't say, and nobody imposes it on us from the outside. And that was the main idea, the core.*
>
> *(Respondent 1)*

The foundation of the model was an extensive database, developed in collaboration with public administrations, publishers and cultural institutions (PLLuM, 2025). The construction of the training corpus covered 100 milliards of words. It was designed to understand natural Polish, regarding its specificity, history and socio-administrative context. A key aspect of it was to ensure content security, i.e. to minimise the generation of harmful, offensive or toxic speech. PLLuM will play a key role in the digitalisation of public administration and to facilitate the navigation of complex government systems. As a prototype of a citizen assistant, it will support

the operation of the mObywatel [mCitizen] application and facilitate citizens' access to public information and services. The project resulted in a family of the PLLuM models, which consists of 18 versions. They differ in the number of parameters and their purpose. The models can be divided into three variants. The first, a base model, which has undergone language adaptation, understands Polish and can predict subsequent words. The second model, the instructional model, has been trained to perform specific tasks based on instructions. As one interviewee points out:

> *The instructional model can perform certain tasks. We refer to it as nurtured (aliveness), i.e. adapted to preferences. The model not only responds, carries out tasks, but carries them out in a certain way that we define. For example, according to norms, according to social values, according to expectations.*
>
> *(Respondent 3)*

The third of the family of the models, the dialogue model, was adapted after the nurturing process to interact and talk with users.

Application of the Bielik model family in business

The Bielik models will be useful for text analytics and process automation tasks. This makes them suitable for: a) automated document analysis - e.g. reviewing regulations, contracts, procedures; b) creating summaries - financial reports, ESG reports, market analyses; c) generating marketing content - advertising slogans and content, product descriptions. As one interviewee highlighted: *The Bielik can summarise, analyse, style answers - it does it all, but it doesn't have a built-in fact base. It does, however, make great inferences.* The model is easy to personalise, which makes it possible to tailor its operation to the specifics of a given industry - *(...) if someone gives the Bielik a description of their business and says: come up with a slogan, they can additionally choose the level of creativity of the model. (Respondent 2).*

The Bielik family of models can provide support for knowledge management in an organisation. The Bielik provides opportunities for integration with various systems, e.g. internal databases or analytical tools. According to one respondent, the model gets its wings when it is connected to a knowledge source, e.g. a library system. In this way, companies can automatically search and summarise large amounts of knowledge, create assistants (chatbots) that can support employees in data analysis and subsequent key decision-making, and use the models to automate customer interactions. A useful tool for business is the Bielik's ability to correctly interpret linguistic errors, sarcasm or colloquial style, and it is good at recognising emotions and tone of speech. These functionalities are particularly important for analysing customer feedback, automating the handling of queries or processing corporate communication texts. In this way, it can support companies in monitoring online brand and product opinions, analysing marketing campaigns and their reception or detecting disinformation. As the Bielik

offers the ability to process internal data, which is often of a confidential and sensitive nature, without having to send it to external providers. This feature appears to be crucial for companies that process medical, financial and legal records and for which data and information security is a priority:

> *The Bielik as an open source system can be downloaded onto their own infrastructure, which ensures that data is not sent to third parties and is not used for further training.*
>
> (*Respondent 4)*

By basing the Bielik family of models on the idea of open source and local operation, companies do not have to incur costs for processing data on commercial providers' servers. For the organisations that work with large data sets, this translates into considerable savings. Companies can create their own AI systems without incurring commercial licensing or token costs, and adapt their Bielik-based models to the specifics of the organisation without licensing restrictions. One respondent stated that:

> *For large databases, the cost of cloud computing can run into the millions of dollars. Using a local model like the Bielik reduces costs to the level of electricity consumption.*
>
> (*Respondent 2)*

In summary, the Bielik family of models offered as an extension and ethically active solution for AI business applications. The applicability, open source basis and application to applications that are available for outreach is truly an alternative to large language models. hidden for organisations that are subject to confidential data and large repositories of knowledge. The Bielik model not only strengthens processes such as document analysis, knowledge management and marketing automation, but is also a control part with strategic priorities, attachments like data analytics, operational scalability and cost management.

This can lead to an organisation's application, decision-making based on the provision and use of language processing technology without lack of control over their data infrastructure.

PLLuM – application of the model in the public sector

Among the wide range of 18 variants of the PLLuM family of models, licensed (non-commercial) and general-purpose models are available. Licensed models, with a larger corpus, have been trained on large datasets (approximately 150 billion tokens) and are available for research purposes only. They may be applicable in academia, language analysis and prototype implementations in the public administration. Smaller models can be commercially applied in customer automation processes, report generation, or legal text analysis. The access to a variety of models provides institutions and organisations the opportunity to choose and adapt to specific, individual needs and capabilities. These tools can support the public sector in the activities related to automating administrative processes, improving official communication or

supporting citizens in accessing public services. One argument in favour of using PLLuM models is their high linguistic quality. The models were built and then trained on the basis of organic data. This allowed the elimination of linguistic loans derived from direct translations of English-language collections.

> *There are no calques in the models, there is all the punctuation. If we want to use the model in public administration, we want it to actually use correct Polish, so that there are no calques, this negative transfer of translations, then such a model, which is sensitised to organic instructions, has a definite advantage over other models here.*
>
> (*Respondent 3)*

The texts generated by the model ensure not only linguistic correctness but security, reliability and legitimacy. In the context of administrative applications, these features take on particular importance. The models work with data that has been acquired in accordance with applicable laws, including data protection or copyright regulations.

The PLLuM models are currently in the prototype phase, but will be available for testing for the selected public institutions as early as 2025. The pilot deployments in the public administration services are planned as part of the new HIVE AI project. As part of the integration with the mObywatel [mCitizen] application, its users will be able to receive answers to the questions about official procedures. A citizen assistant, created as part of the project, will assist citizens and officials in the tasks related to searching and processing information. In this case, in the future, it will be possible to apply the RAG (Retrieval-Augumented Generation) technique. It allows the use of external databases and synthesis of answers in the same way as the regular Internet search engines do. Due to this, the model is not based solely on the knowledge recorded in neural structures, but can dynamically operate on the current data and official documents. One of the people responsible for the implementation of the HIVE AI project stated:

> *Thanks to the cooperation with additional partners, such as COI and Cyfronet AGH, it will be possible to increase computing power and improve the process of implementing the model in practice. The project will allow the entire sector to enter the key stage of the digitisation of public services.*
>
> *(Respondent 3)*

The PLLuM language model reinforces the direction of language models = tailored to the needs of Polish public administration. The wide spectrum of model variants available, enables a flexible and strategic approach of the organisation to digital transformation. The planned integration of PLuMM with the mObywatel platform, illustrates its potential to increase citizen engagement, streamline bureaucratic processes and support information accessibility. The PLLuM model offers not only a technological innovation, but also a tool that is tailored to the modernisation of the public sector.

Discussion

The study results confirm that both the Bielik and PLLuM represent complementary approaches to the development of national language models. The models are tailored to the diverse needs of both public and private sector users. The models differ in architecture, implementation goals and funding modalities, but they share a strategic role in building digital sovereignty and enhancing digital competence. The Bielik model, developed within an open and experimental environment, targets both business and public sector needs. Features such as open access and the possibility of local deployment make it an attractive tool also for small and medium-sized enterprises, or offices. In turn, the PLLuM model, developed in a consortium of research and public institutions, is tailored to the needs of the public administration and state institutions. The model offers a high level of security and compliance with the legal standards.

The study results confirm that both the Bielik and PLLuM represent complementary approaches to the development of national language models. The models are tailored to the diverse needs of both public and private sector users. The models differ in architecture, implementation goals and funding modalities, but they share a strategic role in building digital sovereignty and enhancing digital competence. The Bielik model, developed within an open and experimental environment, targets both business and public sector needs. Features such as open access and the possibility of local deployment make it an attractive tool also for small and medium-sized enterprises, or offices. In turn, the PLLuM model, developed in a consortium of research and public institutions, is tailored to the needs of the public administration and state institutions. The model offers a high level of security and compliance with the legal standards.

However, as Trott et al. (2023) indicate, current models are still limited in their ability to capture deep cognitive processes, such as understanding the beliefs of others. This means that, despite their technological advancement, the use of LLMs should include an awareness of their cognitive and epistemological limitations. According to Pandey's (2025) research, national language models are becoming key components of digital sovereignty of states, offering an alternative to corporate solutions that often fail to grasp local contexts. Based on the local data and preferences, models such as the Bielik and PLLuM can serve as a foundation for responsible and sovereign development of AI.

Summary

Both business and the public sector are subject to the changes initiated by the rapid development of artificial intelligence. AI is constantly changing our world and introducing a revolution linked to large national language models. This research indicates that national

language models, such as the Bielik and PLLuM, are key tools to support the processing and generation of content in Polish. Their use contributes to the development of digital competences and also promotes innovation in the public and private sectors. Importantly, these models play a strategic role in building technological independence and strengthening cyber security, fitting in with Poland's long-term digital sovereignty goals. National LLMs play a multi-faceted strategic role. This role involves alignment with the needs of the public sector and business, especially in strengthening management processes. With their advanced natural language processing capabilities, the models adapt Polish to the socio-cultural context and ensure high-quality information processing. Their functionalities are not only important for easing the burden of routine tasks or increasing the efficiency of employees from both sectors. Moreover, the role of national LLMs takes on particular strategic importance. In addition to providing information and analytical tools, they contribute to a country's technological independence and strategic digital sovereignty. By processing sensitive data within national borders, they provide independence from global models. Their creators and owners are often independent corporations whose models are encroaching more and more aggressively on the private and public sphere of the Polish citizens. These models provide scalable tools for operational efficiencies, customer interactions and knowledge management, especially in the sectors that require the specific Polish language. Companies and the public sector can use LLM to automate services, personalise content and make data-driven decisions.

The development of national LLM programmes, such as the Bielik and PLLuM, carries significant implications in the areas of policy and governance. For policymakers, these initiatives highlight the importance of sustained investment in localised artificial intelligence infrastructure, ethical governance and data sovereignty. National support for open but secure artificial intelligence ecosystems can strengthen strategic autonomy and innovation potential.

In the context of government applications, this is particularly important both for building citizens' trust in the actions of the government and public administration. The question of whether national LLMs should belong to the national critical infrastructure also seems important. Fostering public-private-academic partnerships can accelerate the responsible deployment of artificial intelligence and stimulate further research and development.

Unlike many non-English LLM initiatives, Poland emphasizes distinct entrepreneurial strategies to mitigate data vulnerability and ensure cultural alignment. This study comparatively analyzes entrepreneurial practices and sovereignty outcomes.

References

1. *Bielik the LLM.* The SpeakLeash Foundation. Retrieved from: www.bielik.ai
2. Bouncken, R.B., Czakon, W., Schmitt, F. (2025). Purposeful sampling and saturation in qualitative research methodologies: recommendations and review. *Review of Managerial Science*, 1-37.
3. Carlson, S.M., Koenig, M.A., Harms, M.B. (2013). Theory of mind. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(4), 391-402.
4. Carullo, G. (2023). Large Language Models for Transparent and Intelligible AI-Assisted Public Decision-Making. *CERIDAP*, *3*, 1-23.
5. Chang, A.D, Xiumin Xi., M., Zhou, Ch. (2023). *AI (ChatGPT) Democratization, Return Predictability, and Trading Inequality. Olin Business School Center for Finance & Accounting Research Paper, No. 2024/19*. Retrieved from: https://ssrn.com/abstract=4543999 or http://dx.doi.org/10.2139/ssrn.4543999
6. Chatham, M., Duncan, T., Li, W. (February 8, 2024). *Employing AI in Academia: the Role of ChatGPT in Generating Research Questions*. Retrieved from: https://ssrn.com/abstract=4721270 or http://dx.doi.org/10.2139/ssrn.4721270
7. Chen, J., Liu, Z., Huang, X., Wu, C., Liu, Q., Jiang, G., Pu, Y., Lei, Y., Chen, X., Wang, X., Zheng, K., Lian, D., Chen, E. (2024). When large language models meet personalization: Perspectives of challenges and opportunities. *World Wide Web, 27*(42). https://doi.org/10.1007/s11280-024-01276-1
8. Chen, L.C., Pardeshi, M.S., Liao, Y.X., Pai, K.C. (2025). Application of retrieval-augmented generation for interactive industrial knowledge management via a large language model. *Computer Standards & Interfaces, 94*, 103995.
9. Chiarello, F., Giordano, V., Spada, I., Barandoni, S., Fantoni, G. (2024). Future applications of generative large language models: A data-driven case study on ChatGPT. *Technovation, 133*, 103002. https://doi.org/10.1016/j.technovation.2024.103002
10. Choi, H.W., Abdirayimov, S. (2024). Revolutionizing Education: The Era of Large Language Models. *Journal of Multimedia Information System*, *11*(1), 97-100.
11. Chowdhury, G., Chowdhury, S. (2024). AI-and LLM-driven search tools: A paradigm shift in information access for education and research. *Journal of Information Science*, 01655515241284046.
12. Comarch S.A. (2024). *PCSS oficjalnym partnerem polskiej prezydencji w Radzie UE.* Retrieved from: https://www.pcss.pl/pcss-oficjalnym-partnerem-polskiej-prezydencji-w-radzie-ue/
13. Dash, M., Bakshi, S. (2019). An exploratory study of customer perceptions of usage of chatbots in the hospitality industry. *International Journal on Customer Relations*, *7*(2), 27-33.

14. de Curtò, J., de Zarzà, I., Calafate, C.T. (2024, June). LLM Multi-agent Decision Optimization. In: *KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications* (pp. 3-15). Singapore: Springer Nature Singapore.
15. Decker, H., Trang, K., Ramirez, J., Colley, A., Pierce, L., Coleman, M., ..., Wick, E. (2023). Large language model− based chatbot vs surgeon-generated informed consent documentation for common procedures. *JAMA network open*, *6*(10), e2336997-e2336997.
16. Di Gaetano, S., Diliberto, P. (2018). Chatbots and conversational interfaces: Three domains of use. *Fifth International Workshop on Cultures of Participation in the Digital Age, Vol. 2101*. Castiglione della Pescaia, Italy, pp. 62-70.
17. Estrada-Torres, B., del-Río-Ortega, A., Resinas, M. (2024, May). Mapping the landscape: Exploring large language model applications in business process management. *International Conference on Business Process Modeling, Development and Support*. Cham: Springer Nature Switzerland, pp. 22-31.
18. Fahland, D., Fournier, F., Limonad, L., Skarbovsky, I., Swevels, A.J.E. (2025). How well can a large language model explain business processes as perceived by users? *Data & Knowledge Engineering, 157*, 102416.
19. Ferraris, A.F., Audrito, D., Caro, L.D., Poncibò, C. (2025). The architecture of language: Understanding the mechanics behind LLMs. *Cambridge Forum on AI: Law and Governance, 1(11)*. https://doi.org/10.1017/cfl.2024.16
20. Gallegos, I.O., Rossi, R.A., Barrow, J., Tanjim, M.M., Kim, S., Dernoncourt, F., Yu, T., Zhang, R., Ahmed, N.K. (2024). Bias and Fairness in Large Language Models: A Survey. *Computational Linguistics, 50(3),* 1097-1179. https://doi.org/10.1162/coli_a_00524
21. Huo, B., Boyle, A., Marfo, N., Tangamornsuksan, W., Steen, J.P., McKechnie, T., Guyatt, G. (2025). Large Language Models for Chatbot Health Advice Studies: A Systematic Review. *JAMA Network Open*, *8*(2), e2457879-e2457879. https://doi/10.1001/jamanetworkopen.2024.57879
22. Ilagan, J.R., Ilagan, J.B., Basallo, C.L., Alabastro, Z.M. (2024). Exploratory prompting of large language models to act as co-pilots for augmenting business process work in document classification. *Procedia Computer Science, 237*, 420-425.
23. Islam, I., Islam, M.N. (2024). Exploring the opportunities and challenges of ChatGPT in academia. *Discov. Educ., 3, 31*. https://doi.org/10.1007/s44217-024-00114-w
24. Janssen, A., Rodríguez Cardona, D., Breitner, M.H. (2020). More than FAQ! Chatbot taxonomy for business-to-business customer services. *International Workshop on Chatbot Research and Design*. Cham: Springer International Publishing, pp. 175-189.
25. Jeong, C. (2023). Generative AI service implementation using LLM application architecture: based on RAG model and LangChain framework. *Journal of Intelligence and Information Systems*, *29*(4), 129-164.
26. Jiang, Z., Xu, F.F., Araki, J., Neubig, G. (2020). How can we know what language models know?. *Transactions of the Association for Computational Linguistics*, *8*, 423-438.

27. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274
28. Kietzmann, J., Park, A. (2024). Written by ChatGPT: AI, large language models, conversational chatbots, and their place in society and business. *Business Horizons, 67*(5), 453-459. https://doi.org/10.1016/j.bushor.2024.06.002
29. Kleiman, F., Barbosa, M.M. (2024). Management and Performance Program Chatbot: A Use Case of Large Language Model in the Federal Public Sector in Brazil. *Digital Government: Research and Practice*.
30. Leong, K., Sung, A. (2024). Gender stereotypes in artificial intelligence within the accounting profession using large language models. *Humanities and Social Sciences Communications*, *11*(1), 1-11.
31. Maarouf, A., Feuerriegel, S., Pröllochs, N. (2025). A fused large language model for predicting startup success. *European Journal of Operational Research, 322*, 198-214.
32. Makridakis, S., Petropoulos, F., Kang, Y. (2023). Large language models: Their success and impact. *Forecasting, 5*(3), 536-549. https://doi.org/10.3390/forecast5030030
33. Mamalis, M.E., Kalampokis, E., Fitsilis, F., Theodorakopoulos, G., Tarabanis, K.A. (2024). A large language model agent-based legal assistant for governance applications. *Proceedings of the International Conference on Electronic Government*. Retrieved from: https://api.semanticscholar.org/CorpusID:271971364
34. Ociepa, K., Flis, L., Wróbel, K., Gwóździej, A., Kinas, R. (2024). *Bielik 7B v0.1: A Polish Language Model - Development, Insights, and Evaluation*.
35. Pandey, P. (2025). Digital sovereignty and AI: Developing India's national AI stack for strategic autonomy. *Procedia Computer Science, 254*, 250-259. https://doi.org/10.1016/j.procs.2025.02.084
36. Papageorgiou, G., Sarlis, V., Maragoudakis, M., Tjortjis, C. (2024). Enhancing E-Government Services through State-of-the-Art, Modular, and Reproducible Architecture over Large Language Models. *Applied Sciences*, *14*(18), 8259. https://doi.org/10.3390/app14188259
37. Pessanha, G.R.G., Vieira, A., Brandão, W.C. (2024). Large Language Models (LLMs): A systematic study in Administration and Business. *RAM. Revista de Administração Mackenzie, 25*(6), eRAMD240059. https://doi.org/10.1590/1678-6971/eRAMD240059
38. *PLLuM, Polish Large Language Model* (2025). Retrieved from: https://pllum.org.pl/#consortium

39. Polish Ministry of Digitalisation (2025). *Poland builds its own artificial intelligence - PLLuM ready for action 2025.* Retrieved from: https://www.gov.pl/web/cyfryzacja/polska-buduje-wlasna-sztuczna-inteligencje--pllum-gotowy-do-dzialania
40. Poznań Supercomputing Centre (2024). *PCSS oficjalnym partnerem polskiej prezydencji w Radzie UE.* Retrieved from: https://www.pcss.pl/pcss-oficjalnym-partnerem-polskiej-prezydencji-w-radzie-ue/
41. Ramli, M.S. (March, 2025). *Enhancing job rotation onboarding in the Malaysian Public Sector: A NotebookLM Large Language Model Case Study, Authorea*. https://doi.org/10.22541/au.174172238.84876993/v1
42. Ravichandiran, S. (2021). *Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT*. Packt Publishing Ltd.
43. Retkowsky, J., Huysman, M., Hafermalz, E. (2024). Managing a ChatGPT-empowered workforce: Understanding its affordances and side effects. *Business Horizons, 67*(5), 511-523.
44. Schwandt, T.A., Gates, E.F. (2018). Case study methodology. In: N.K. Denzin, Y.S. Lincoln (Eds.), *The sage handbook of qualitative research* (pp. 600-630). Sage Publishers.
45. Stanford Institute for Human-Centered Artificial Intelligence (SIfHCAI) (2024). *The AI Index Report 2024*. FTI Delta & IBM. Retrieved from: https://aiindex.stanford.edu/report/
46. Szilágyi, R., Tóth, M. (2023). Use of LLM for SMEs: Opportunities and Challenges. *Journal of Agricultural Informatics, 14(2)*, 10-18.
47. Thirunavukarasu, A.J., Ting, D.S.J., Elangovan, K., Gutierrez, L., Tan, T.F., Ting, D.S.W. (2023). Large language models in medicine. *Nature Medicine, 29*(8), 1930-1940. Retrieved from: https://doi.org/10.1038/s41591-023-02448-8
48. Trott, S., Jones, C., Chang, T., Michaelov, J., Bergen, B. (2023), Do Large Language Models Know What Humans Know? *Cognitive Science, 47,* e13309. https://doi.org/10.1111/cogs.13309
49. Wang, Z., Chu, Z., Doan, T.V., Ni, S., Yang, M., Zhang, W. (2024). History, development, and principles of large language models: an introductory survey. *AI and Ethics*, 1-17.
50. Wickramasekara, A., Scanlon, M. (2024). A Framework for Integrated Digital Forensic Investigation Employing AutoGen AI Agents. *Proceedings of the 12th International Symposium on Digital Forensics and Security. ISDFS.* IEEE, pp. 01-06.
51. Yang, Z., Xu, S.-S., Liu, X., Zhou, Y.-M. (2024). Large Language Model–Based Critical Care Big Data Deployment and Extraction: Descriptive Analysis. *Journal of Medical Internet Research*.
52. Yoon, S. (2024). Digital Innovation in Public Administration through Intelligent Public Sector Automation (IPSA): Strategies and Challenges. *Journal of Multimedia Information System*, *11*(4), 249-260.

53. Żmigrodzki, M., Kondracki, S., Sarna, N., Chibowski, K., Sarnowski, M., Kowalski, J.M., Jędrzejewski, K., Piwnicki, Ł., Babis, N. (2025). *Użyteczność biznesowa dużych modeli językowych (LLM). Raport z badań wstępnych. SpeakLeash*. Retrieved from: https://static.android.com.pl/uploads/2025/01/RAPORT-Uzytecznosc-biznesowa-duzych-modeli-jezykowych-e-book.pdf
54. Zubiaga, A. (2024). Natural language processing in the era of large language models. *Frontiers in Artifical Intelligence, 6*, 1350306. https://doi.org/10.3389/frai.2023.1350306