GPT-NL: Towards a Public Interest Large Language Model*

Tom Barbereau[1,*,†] and Lieke Dom[2,†]

[1] *Dutch Organization for Applied Scientific Research (TNO), Anna van Buerenplein 1, 2595 DA Den Haag, The Netherlands*

Abstract

Artificial intelligence (AI) is evolving into a transformative force across various sectors. Large language models (LLMs) are at the forefront of this transformation. To date, LLMs are predominantly controlled by multinational, private companies that are misaligned with public interests. This monopolization threatens sovereignty. OpenAI's ChatGPT is paradigmatic for this observation. In response, the Dutch government has invested in projects that foster sovereign innovation in numerous sectors. This paper introduces one of these projects – GPT-NL – aimed at developing a Dutch LLM and fostering a supportive ecosystem involving academic institutions, companies, and government. Though formally GPT-NL's development is directly aligned with the European Union's Guidelines on Trustworthy AI, in the scope of this academic work we focus on its alignment to requirements for public interest AI (Züger & Asghari, 2023). As such, GPT-NL serves as a case on public interest AI for future research that provides knowledge to practitioners.

1. Introduction

Artificial intelligence (AI) is rapidly developing from a promise to a game changer in many industries. One of the most important developments in recent years is the widespread introduction and adoption of large language models (LLMs). OpenAI's ChatGPT is the leading platform, used by millions of users across the world. (Provided the limited scope of this paper, we refer the reader to literature on LLMs [see 1].)

LLMs are almost exclusively developed and owned by American and Chinese multinational technology giants, who are known to shield access from users, are closed off to researchers' scrutiny, and are a thorn in the foot of regulators [2,3]. This form of monopolisation creates a situation in which there is uncertainty about safeguarding values and posing questions to confidentiality, privacy, intellectual property and compliance with (inter)national legislation or policy frameworks such as the European Union's AI Act [4]. In short, sovereignty is under threat [5].

In response to this threat, the Dutch government has committed significant resources towards projects that strive to strengthen sovereignty and strategic autonomy [6]. Amongst others, support is being allocated to research on hydrogen propulsion, the development of biobased building materials, and in sectors of food safety and smart industry. One these projects is GPT-NL, an initiative aimed at training a Dutch-English LLM of the Netherlands aligned with public interests as deducted from European regulations and Constitutional fundamental rights. The project can be seen as a direct response to the rise and dominance of ChatGPT [7].

In this paper, we introduce GPT-NL – a joint initiative by the Dutch Organization for Applied Scientific Research (TNO), the cooperative association of Dutch educational and research institutions (SURF), and the Netherlands Forensic Institute (NFI) – launched in November 2023. The primary goal of GPT-NL is to develop, strengthen and perpetuate public interests (1) with its own Dutch language model and (2) the formation of an associated ecosystem of academic institutions, private companies, government, and end users. The model's development is firmly rooted in fundamental rights and aligns with the EU's Guidelines for Trustworthy Artificial Intelligence [8]. With GPT-NL, we introduce a nascent case of value to future research and practitioners working on public interest AI.

This paper is structured as follows. In Section 2, we set the scene and provide the background to GPT-NL. In Section 3, we introduce GPT-NL and map it to the requirements for public interest AI [9]. In Section 4, we discuss this mapping exercise and reflect on the relationship between the EU's Guidelines and the requirements. Here, we provide recommendations to practitioners, introduce future research questions, and discuss limitations. In Section 5, we conclude with an outlook for GPT-NL.

---

2. Background

In the last decade, one of the most disruptive technologies was the large-scale introduction of LLMs, with the leading platform ChatGPT from the company OpenAI. ChatGPT managed to conquer the world within a few weeks alone [10]. Alphabet's Google quickly followed suit; and so did other multinational technology providers [11]. These LLMs can perform a large variety of tasks – they answer questions, write essays, structure texts, and perform repetitive tasks autonomously – with a surprising accuracy. LLMs are poised to find application in numerous industries and consequently affect labour markets [2].

At present and given the (uncontrollable) rise of LLMs, sovereignty and strategic autonomy is increasingly under pressure [12]. While until recently groundbreaking research into AI mainly took place within the academic sphere, the dependence on computing power and troves of data has caused a shift towards developments in private companies based in the United States and China [13,14]. The LLMs developed by these companies are closed off to public scrutiny, hoard collected data, and feature inherent biases. Indeed, the generated outputs may:

1. contain information whose origin and authenticity can be disputed [4,15],
2. contain information that may contain stigmatizing opinions and judgments [16,17], and
3. pose risks of manipulation [18].

Generally speaking, these private models are in no way aligned with public interests [12,19,20]. Furthermore, in spite of their widespread availability – making it attractive for businesses and governments to interweave their products and services with them 'on-demand' [2] –, numerous economic risks are posed by these (e.g., vendor lock-in's).

In response to these dynamics, the Dutch government selected GPT-NL as part of its plan to fund projects that foster sovereignty [6]. The project plan "Facility for a sovereign Dutch language model" was submitted in May 2023 and awarded at the end of November 2023. An amount of 13,5 million euros has been made available for the project. GPT-NL is also included in the government-wide vision on generative AI

of the Netherlands where it plays a key role in the "development of open […] LLMs in line with public values" [21]. GPT-NL consists of two phases:

Phase 1: The first phase of the GPT-NL project is to develop, strengthen and perpetuate digital sovereignty through a government-based Dutch language model [7,22]. Aside from the model, an ecosystem of private, governmental, and public interest groups is formed to contribute to and oversee the model [23,24]. Academic institutions are also represented, in particular with regards to the former.

Phase 2: In the second phase, GPT-NL focuses specifically at facilitating application-oriented research into the use of LLMs in sensitive contexts. This can include, for example, developing advanced solutions to make government communication accessible to people with low literacy, improve forensic investigation, automatically processing large amounts of text (such as case law and intercepted communications), enabling natural interaction between humans and computers or robots in healthcare, and acting as an assistant when drawing up a program of requirements for defence purchases. All of these are applications with high requirements regarding confidentiality, stigmatization, sensitivity to errors and transparency. Dependence on foreign (commercial) parties is undesirable in this regard [7].

At the time of writing, GPT-NL is in the first phase of development, which means that there is a dominant focus on data collection. This paper aims to introduce GPT-NL as a case for future research into the alignment to public interests.

3. Public interest and GPT-NL

To realize and align GPT-NL with public interests, the model is formally built in line with the European Union's Guidelines for Trustworthy AI [8]. The guidelines put forward a framework under which systems must meet certain requirements to be deemed trustworthy. The framework covers two out of three vertical axis which should be met throughout the system's entire life cycle ("Ethical AI" and "Robust AI").[2] The framework consists of three levels in decreasing order of abstraction: (1) foundations (adherence to principles based on fundamental rights), (2) realization (implementation of key requirements), and (3) assessment

---

[2] The last of the three axis, "Lawful AI", is not covered by the guidelines because it "proceed[s] on the assumption that all legal rights and obligations that apply to the processes and activities involved in developing, deploying and using AI systems remain mandatory and must be duly observed" (p. 6).

(operationalization of the key requirements). Given the limited scope of this paper, we cannot present a detailed overview of the operationalization of these key requirements for GPT-NL. Instead, we subsequently consider the process of development where certain requirements are to be met to refer to an AI system as "serving the public interest" [9].

In Züger & Asghari (2023) an alternative lens to the EU's Guidelines is proposed; that of public interest in political theory.[3] On the basis of latent literature, five requirements are derived for an AI system to serve the public interest. It needs to (1) have a public (not profit-oriented) justification, (2) serve equality, (3) require a deliberation / co-design process, (4) follow key technical standards, and (5) be open for validation. In the following, we map these five requirements to GPT-NL.

3.1. Public (not profit-oriented) justification

Since it cannot be guaranteed that current LLM-providers comply with Dutch and European law, the government-wide vision on generative AI of the Netherlands describes that it is not possible to use these American or Chinese systems in public organisations [21]. GPT-NL offers a solution to this problem. It is targeted to be compliant with Dutch and European laws (n.b. the General Data Protection Regulation (GDPR) and the AI Act), as well as adhering to transparency requirements (n.b. in alignment with the Dutch Open Government Act). To this end, it will be built from scratch such that compliance and transparency can be realized in every step of the development process. It will then be (legally) possible for public parties to benefit from and innovate with generative AI. Arguably, by compliance with EU law and by avoidance of international companies, GPT-NL reflects the idea of public justification.

When public organisations have legal access to LLMs, it can be applied to reap opportunities (e.g. more efficient work processes, transcription, etc). Using GPT-NL will not only allow to reap opportunities, which is in the public interest of the Netherlands; it also aims to contribute to digital sovereignty of both the Netherlands and Europe. When Dutch public organisations have access to their own LLM, one becomes less dependent of alternatives from abroad. Furthermore, by sharing the lessons learned with European member states, Europe will be able to connect expertise and increase its knowledge position on AI, making the Union less dependent from third-party services.

3.2. Serving equality and human rights

The general criticism is that ethical development of AI is reduced to fairness and bias [28]. When it comes to serving equality, GPT-NL adds important values to the practice, namely equal accessibility to the technology and reciprocity. By proxy of the alignment with European regulation, human rights are upheld in GPT-NL.

The technology will be made accessible for everyone by developing a cost-effective modus operandi. The licences will vary depending on the intensity and goal of use (commercial versus non-commercial). No profits will be made; the fee will solely be used to cover operational costs of and reinvestments in GPT-NL. Beyond the licence, GPT-NL aims to create a reciprocal relationship with not only organisations who use the model, but also society more broadly. As the model will partially be trained with datasets provided by Dutch municipalities who wish to use GPT-NL, and because the data originates from the cities and citizens of those municipalities, by reciprocity it is only fair that GPT-NL will not financially benefit from it. Bias will be avoided by proxy of the provided data, which is expected to be representative of Dutch society.

Central to the EU's Guidelines on Trustworthy AI is the notion of AI 'made in Europe'. It implies that any AI systems developed according to the Guidelines is grounded in human rights and the foundational values of the EU — respect for dignity, democracy, equality, and rule of law amongst others [9,29]. This explicitly holds true for GPT-NL by extension of the application of European regulations.

3.3. Deliberative and participatory design process

When it comes to the design process, there are two aspects that highlight GPT-NL's deliberative and participatory way of working the consortium and the data provision.

The consortium behind GPT-NL consists of TNO, SURF and NFI. These organisations bring together important knowledge and resources. Each carries a specific role, ensuring checks-and-balances between one another. TNO has extensive expertise working on privacy

---

[3] We opt for the framework of Züger & Asghari (2023) over others given the workshops' focus on "Public Interest AI". For alternatives refer to Leslie, 2019; Paraman & Anamalah, 2023; Lu et al., 2024 among others.

and anonymization such that it can support data providers in data preparation. Others within the organisation study the possible societal harms and biases that can be embedded in AI [see 29]. SURF has a network within the academic and educational sector. Innovation in this sector is of direct interest to the public. Via SURF, GPT-NL has access to Snellius (the Dutch national supercomputer) as well as the necessary technical expertise to train AI-models on high-performance computation clusters.[4] Snellius is ranked high on the Green500 list,[5] this way it is possible to develop GPT-NL while keeping the environmental footprint low. Lastly, with NFI as a partner within the consortium, GPT-NL directly has a use case to test the model when it is operational. The NFI already built experience with working with LLMs [refer to 30].

The data provision is also an example of how GPT-NL will be built in participation with others. Organisations or people can share their data with GPT-NL on a voluntary basis. GPT-NL will offer corresponding preparation tools and anonymization methods to these organisations. This way, they can clean the dataset themselves and make sure that no personal, confidential, or sensitive data is distributed without consent or legitimate interest. Furthermore, the cooperation between GPT-NL and the data providers also means that GPT-NL can help data providers with the development of future use cases.

3.4. Implementation of technical safeguards

GPT-NL will use a state-of-the-art model architecture. Training and building the model is done from scratch. This helps to strive for high quality of data, system accuracy, and the safeguarding of data privacy.

Current LLMs are pre-dominantly developed by American companies, which means that these LLMs are trained on English data. While some training data used is in other languages (like Dutch), compared to English it is only a minor portion. If used, it is fair to assume that the (Dutch) data is simply scraped from the internet, and is of lower quality [1] and not in line with copy-right legislation. By building a model from scratch, in cooperation with voluntary data providers, GPT-NL aims to build a model trained on high quality Dutch (and English) data from consenting providers.

This contributes to system accuracy. Although LLMs like ChatGPT or Gemini also function in Dutch, most of the Dutch is translated which means that cultural aspects and linguistic subtilties are not embedded in the

system. By training a Dutch language model on data from the Netherlands and training the English version on English data about (what happens in) the Netherlands, the system will be more accurate when it is used.

At last, because the development of GPT-NL happens in accordance with the GDPR, the AI Act, and intellectual property law, the highest technical standards are to be used. However, given the current stage of development, no information on which standards are used can be made publicly available.

3.5. Openness to validation

Transparency is the main driver that distinguishes GPT-NL from other, profit-oriented LLMs. GPT-NL aims to be transparent where possible. For example, it is not possible to publish actual datasets since there is always a residual risk that personal or sensitive data is still in these. With regards to public interest, the values of transparency and privacy cannot be satisfied. Therefore, GPT-NL opts to publish data sheets. Data sheets reveal performance data, technical specifications, and other information on useability. They do not reveal sensitive information [31] or, in this case, traces to the provided training data. Additionally, GPT-NL publishes decisions that relate to the process of development. For example, commitments are listed on which GPT-NL aims to report on a quarterly basis [32].

Another example of GPT-NL's openness to validation is in reference to the Dutch Algorithm Register ("Algoritmeregister"). This register helps to increase transparency about the use of algorithms across governmental organisations. It makes it easier for citizens, researchers, and media to hold the government accountable for algorithmic usage [33]. When governmental organisation or public parties build systems and tools using GPT-NL, these will also be listed in the register. As soon as it will be used in the Netherlands, GPT-NL itself will also be added to the register by these parties. While to date it is not mandatory for organisations to publish their algorithms, this will become legally binding by the end of 2024 [34].
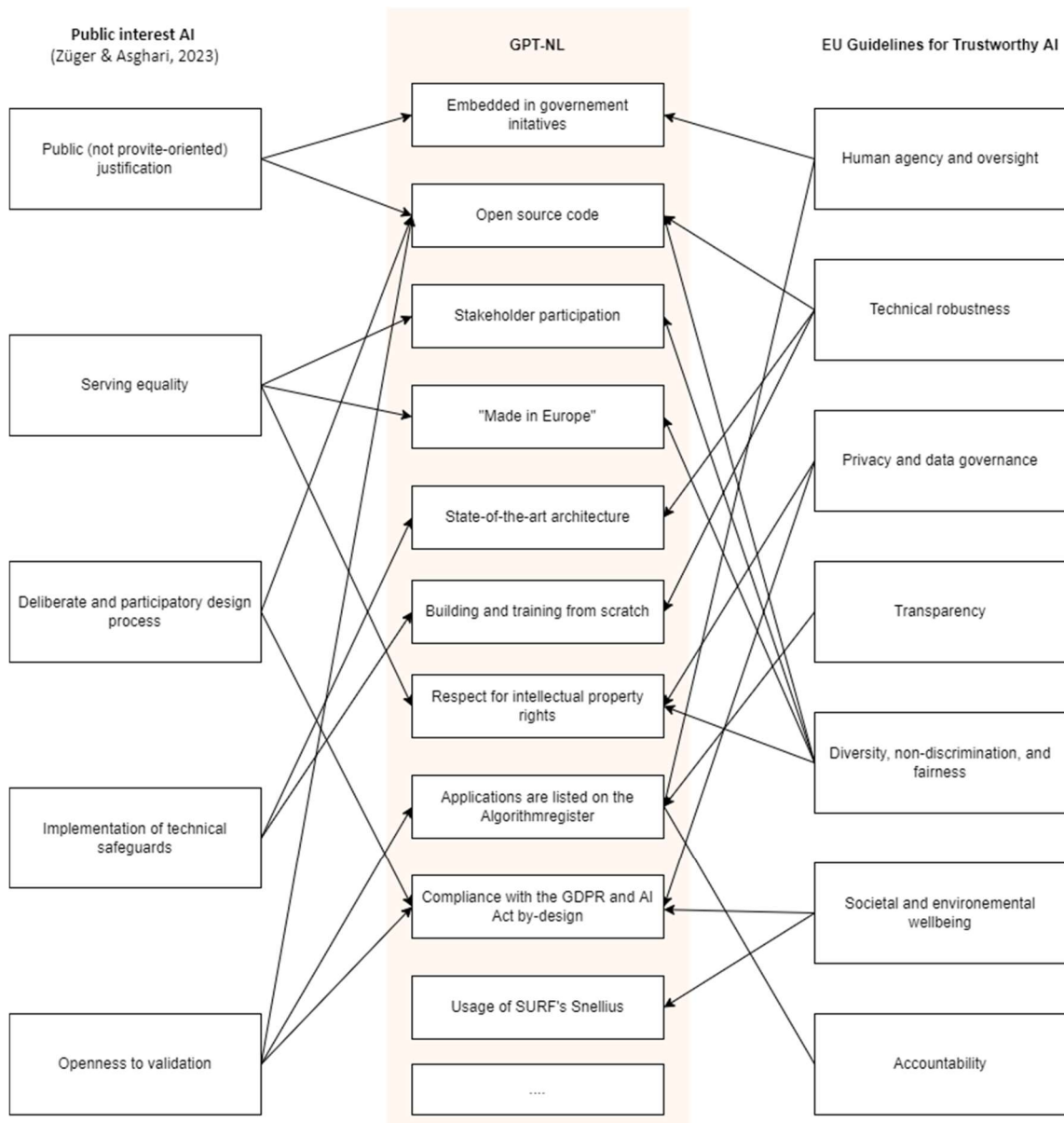

Figure 1: Visual mapping of frameworks to GPT-NL.

4. Discussion

For GPT-NL, as seen in the previous section, there is considerable overlap between the EU's Guidelines for Trustworthy AI [8] and the requirements for public interest AI [9]. Given the scope of this paper, we merely provide a tentative, visual mapping between these two frameworks using the case of GPT-NL (**Figure 1**).[6] The figure demonstrates numerous connections and overlaps between the EU's Guidelines for Trustworthy AI and the requirements for public interest AI. This is to be expected [9].

GPT-NL is a concrete case for the operationalizations of public interest AI guidelines. Despite the value provided by the EU's Guidelines and the public interest AI framework, additional research is needed around challenges in their operationalization. There can be no single framework that offers a simplified recipe of doing so – in particular, with regards to questions of ethics

---

[6] Note that the last box for GPT-NL acts as placeholder accounting for future iterations of the project.

[35]. Identifying tensions and controversies in operationalization is an ambitious, yet necessary topic of research.

With GPT-NL, we have here seen how the institutional and political landscape – composed of entities strongly embedded within initiatives led by the Dutch government – require alignment with normative frameworks [see also 29,36,37]. We hope that researchers can continue to take concrete cases – such as GPT-NL – to derive generalizable knowledge and feed back into existing theoretical research. Best practices will be of considerable value to practitioners.

This paper exhibits several limitations characteristic of early-stage, case-centric research. GPT-NL began last year and is at an early stage of development. The absence of detailed information on the iterative process and the planned roadmap leaves uncertainty regarding the project's trajectory and scalability [38]. Conversely, this leaves room for the consortium to publish a transparent developmental framework to map the commitments made to fully realize the potential of the GPT-NL project in subsequent iterations. Similarly, questions related to the publication of data sheets will be raised from the perspective of data providers. It is quintessential to ensure that these are compliant with the GDPR [see 31].

As authors who are (in-)directly involved with the project, we are aware of our own confirmation bias. To account for such bias, we disclosed our affiliation transparently and believe to have introduced GPT-NL in an objective manner. This paper does not aim to promote the project; instead, it aims to provide a case for future research on public interest AI.

5. Outlook

Looking ahead, the ambition for GPT-NL is to give public organisations access to a LLM to accelerate innovation all while remaining sovereign. It will be made possible to go beyond pilots and experiments and improve processes that are of public interest. By sharing knowledge and building an ecosystem around GPT-NL, the knowledge position of the Netherlands will increase. In the future, GPT-NL aims to share knowledge across the European Union, therefore supporting Europe's sovereignty at large.

Acknowledgements

References

[1] J. Yang, H. Jin, R. Tang, X. Han, Q. Feng, H. Jiang, S. Zhong, B. Yin, X. Hu, Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond, ACM Transactions on Knowledge Discovery from Data 18 (2024) 160:1-160:32. https://doi.org/10.1145/3649506.

[2] Y.K. Dwivedi, N. Kshetri, L. Hughes, E.L. Slade, A. Jeyaraj, A.K. Kar, A.M. Baabdullah, A. Koohang, V. Raghavan, M. Ahuja, H. Albanna, M.A. Albashrawi, A.S. Al-Busaidi, J. Balakrishnan, Y. Barlette, S. Basu, I. Bose, L. Brooks, D. Buhalis, L. Carter, S. Chowdhury, T. Crick, S.W. Cunningham, G.H. Davies, R.M. Davison, R. Dé, D. Dennehy, Y. Duan, R. Dubey, R. Dwivedi, J.S. Edwards, C. Flavián, R. Gauld, V. Grover, M.-C. Hu, M. Janssen, P. Jones, I. Junglas, S. Khorana, S. Kraus, K.R. Larsen, P. Latreille, S. Laumer, F.T. Malik, A. Mardani, M. Mariani, S. Mithas, E. Mogaji, J.H. Nord, S. O'Connor, F. Okumus, M. Pagani, N. Pandey, S. Papagiannidis, I.O. Pappas, N. Pathak, J. Pries-Heje, R. Raman, N.P. Rana, S.-V. Rehm, S. Ribeiro-Navarrete, A. Richter, F. Rowe, S. Sarker, B.C. Stahl, M.K. Tiwari, W. van der Aalst, V. Venkatesh, G. Viglia, M. Wade, P. Walton, J. Wirtz, R. Wright, "So, what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy, International Journal of Information Management 71 (2023) 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642.

[3] M. Goos, M. Savona, The governance of artificial intelligence: Harnessing opportunities and mitigating challenges, Research Policy 53 (2024) 104928. https://doi.org/10.1016/j.respol.2023.104928.

[4] N. Helberger, N. Diakopoulos, ChatGPT and the AI Act, Internet Policy Review 12 (2023). https://policyreview.info/essay/chatgpt-and-ai-act (accessed July 4, 2024).

[5] C. Stolwijk, M. Punter, P. Timmers, J. Rabbie, D. Regeczi, S. Dalmolen, Towards a sovereign digital future – the Netherlands in Europe, TNO, 2024.

[6] Ministerie van Economische Zaken en Klimaat, 14 Nederlandse onderzoeksfaciliteiten krijgen € 184 miljoen ondersteuning, Rijksoverheid (2023). https://www.rijksoverheid.nl/actueel/nieuws/2023/11/02/14-nederlandse-onderzoeksfaciliteiten-krijgen-184-miljoen-euro-ondersteuning (accessed July 8, 2024).

[7] L. Verhagen, Nederland ontwikkelt antwoord op ChatGPT: AI-taalmodel GPT-NL, de Volkskrant (2023). https://www.volkskrant.nl/tech/nederland-ontwikkelt-antwoord-op-chatgpt-ai-taalmodel-gpt-nl~b3c293bf/ (accessed July 8, 2024).

[8] High-Level Expert Group on AI, Ethics Guidelines for Trustworthy Artificial Intelligence, European Commission, 2019. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[9] T. Züger, H. Asghari, AI for the public: How public interest theory shifts the discourse on AI, AI & Soc 38 (2023) 815–828. https://doi.org/10.1007/s00146-022-01480-5.

[10] K. Grace, S. Jackson, The history of OpenAI, from the early days with Elon Musk to the ChatGPT maker being put on blast by Scarlett Johansson, Business Insider (2024). https://www.businessinsider.com/history-of-openai-company-chatgpt-elon-musk-founded-2022-12 (accessed July 5, 2024).

[11] N. Tiku, G.D. Vynck, W. Oremus, Big Tech was moving cautiously on AI. Then came ChatGPT., Washington Post (2023). https://www.washingtonpost.com/technology/2023/01/27/chatgpt-google-meta/ (accessed July 8, 2024).

[12] H. Hirsch-Kreinsen, T. Krokowski, Trustworthy AI: AI made in Germany and Europe?, AI & Soc (2023). https://doi.org/10.1007/s00146-023-01808-9.

[13] H. Roberts, E. Hine, M. Taddeo, L. Floridi, Global AI governance: barriers and pathways forward, International Affairs 100 (2024) 1275–1286. https://doi.org/10.1093/ia/iiae073.

[14] P. Terzis, Law and the political economy of AI production, International Journal of Law and Information Technology 31 (2023) 302–330. https://doi.org/10.1093/ijlit/eaae001.

[15] L. Floridi, M. Chiriatti, GPT-3: Its Nature, Scope, Limits, and Consequences, Minds & Machines 30 (2020) 681–694. https://doi.org/10.1007/s11023-020-09548-1.

[16] B. Fecher, M. Hebing, M. Laufer, J. Pohle, F. Sofsky, Friend or foe? Exploring the implications of large language models on the science system, AI & Soc (2023). https://doi.org/10.1007/s00146-023-01791-1.

[17] J. Rutinowski, S. Franke, J. Endendyk, I. Dormuth, M. Roidl, M. Pauly, The Self-Perception and Political Biases of ChatGPT, Human Behavior and Emerging Technologies 2024 (2024) 7115633. https://doi.org/10.1155/2024/7115633.

[18] M. Sax, Between empowerment and manipulation: the ethics and regulation of for-profit health apps, Wolters Kluwer, Alphen aan den Rijn, 2021.

[19] B.C. Stahl, D. Eke, The ethics of ChatGPT – Exploring the ethical issues of an emerging technology, International Journal of Information Management 74 (2024) 102700. https://doi.org/10.1016/j.ijinfomgt.2023.102700.

[20] I. Ulnicane, W. Knight, T. Leach, B.C. Stahl, W.-G. Wanjiku, Framing governance for a contested emerging technology: insights from AI policy, Policy and Society 40 (2021) 158–177. https://doi.org/10.1080/14494035.2020.1855800.

[21] Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Overheidsbrede visie Generatieve AI, Rijksoverheid, 2024. https://www.rijksoverheid.nl/documenten/rapporten/2024/01/01/overheidsbrede-visie-generatieve-ai.

[22] W. van der Louw, Nederland zet met beperkte middelen in op eigen AI-taalmodel, Trouw (2023). https://www.trouw.nl/binnenland/nederland-zet-met-beperkte-middelen-in-op-eigen-ai-taalmodel~b6faaf00d/ (accessed July 8, 2024).

[23] GPT-NL, Alleen samen bouwen we GPT-NL, GPT-NL (2023). https://gpt-nl.nl/samenwerken/ (accessed July 8, 2024).

[24] GPT-NL, Data Acquisition Pipeline, GPT-NL (2023). https://gpt-nl.nl/planning/ (accessed July 8, 2024).

[25] D. Leslie, Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector, The Alan Turing Institute, 2019. https://doi.org/10.5281/zenodo.3240529.

[26] P. Paraman, S. Anamalah, Ethical artificial intelligence framework for a good AI society: principles, opportunities and perils, AI & Soc 38 (2023) 595–611. https://doi.org/10.1007/s00146-022-01458-3.

[27] Q. Lu, L. Zhu, X. Xu, J. Whittle, D. Zowghi, A. Jacquet, Responsible AI Pattern Catalogue: A Collection of Best Practices for AI Governance and Engineering, ACM Comput. Surv. 56 (2024) 173:1-173:35. https://doi.org/10.1145/3626234.

[28] K. Hao, How Facebook got addicted to spreading misinformation, MIT Technology Review (2021). https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation/ (accessed July 9, 2024).

[29] I. Georgieva, C. Lazo, T. Timan, A.F. van Veenstra, From AI ethics principles to data science practice: a reflection and a gap analysis based on recent frameworks and practical experience, AI Ethics 2 (2022) 697–711. https://doi.org/10.1007/s43681-021-00127-3.

[30] Nederlands Forensisch Instituut, NFI-er schrijft internationaal handboek over toepassing kunstmatige intelligentie bij forensisch onderzoe, Overheid (2023). https://doi.org/10/18/nfi-er-schrijft-internationaal-handboek-over-toepassing-kunstmatige-intelligentie-bij-forensisch-onderzoek.

[31] T. Gebru, J. Morgenstern, B. Vecchione, J.W. Vaughan, H. Wallach, H. Daumé, K. Crawford, Datasheets for datasets, Communications of the ACM 64 (2021) 86–92. https://doi.org/10.1145/3458723.

[32] GPT-NL, Onze commitments, GPT-NL (2023). https://gpt-nl.nl/commitments/ (accessed July 8, 2024).

[33] Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Handreiking Algoritmeregister, Overheid, 2023. https://www.digitaleoverheid.nl/document/handreiking-algoritmeregister/.

[34] Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Stand van zaken Algoritmeregister, (2022). https://open.overheid.nl/documenten/ronl-391fc340bc62f9c0b60f7408f9d780aa6be5168e/pdf (accessed July 9, 2024).

[35] B. Mittelstadt, Principles alone cannot guarantee ethical AI, Nat Mach Intell 1 (2019) 501–507. https://doi.org/10.1038/s42256-019-0114-4.

[36] J. Morley, L. Floridi, L. Kinsey, A. Elhalal, From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices, Sci Eng Ethics 26 (2020) 2141–2168. https://doi.org/10.1007/s11948-019-00165-5.

[37] L. Weigl, T. Barbereau, G. Fridgen, The construction of self-sovereign identity: Extending the interpretive flexibility of technology towards institutions, Government Information Quarterly 40 (2023) 101873. https://doi.org/10.1016/j.giq.2023.101873.

[38] F.W. Geels, Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study, Research Policy 31 (2002) 1257–1274. https://doi.org/10.1016/S0048-7333(02)00062-8.