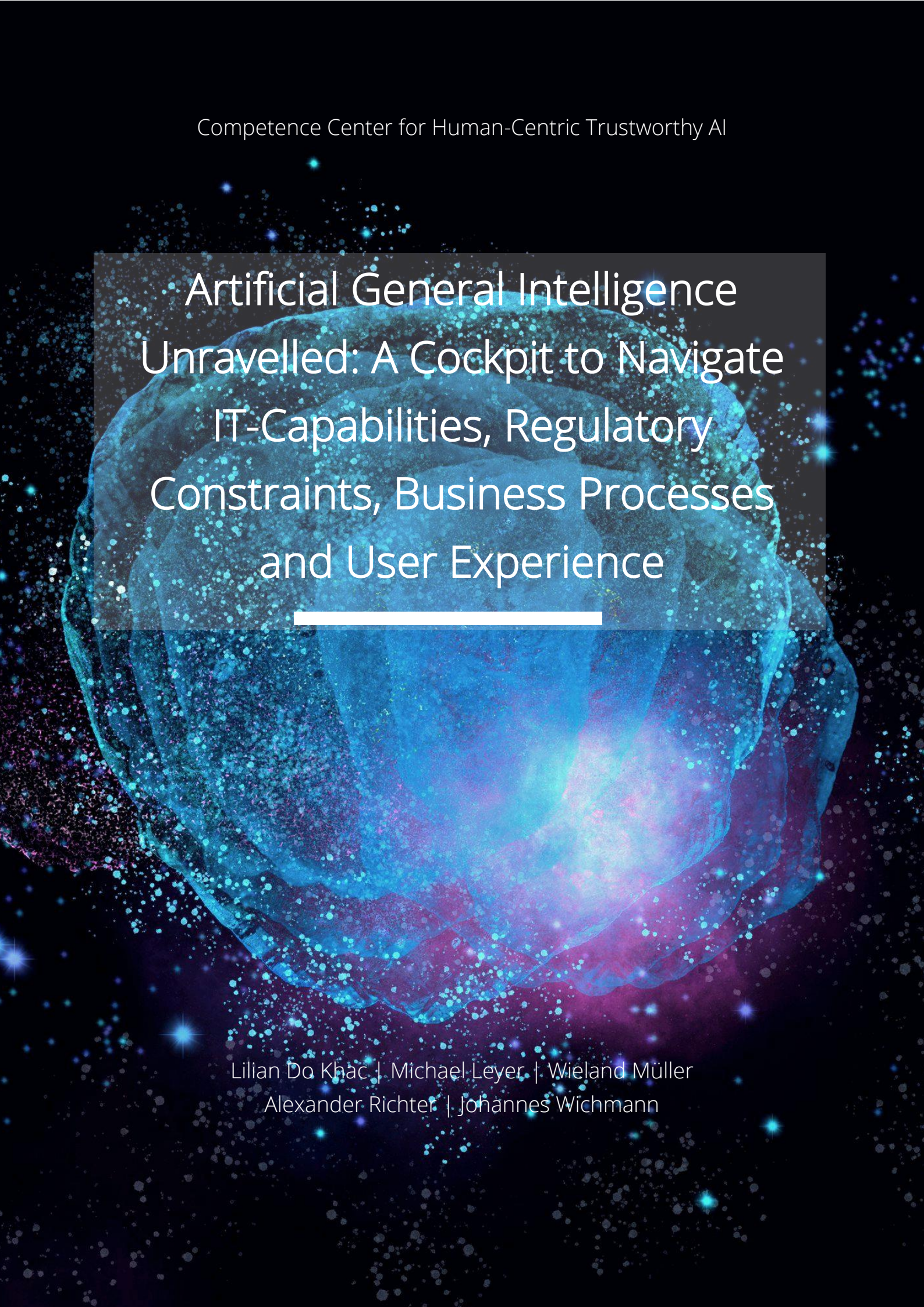
Competence Center for Human-Centric Trustworthy AI



Artificial General Intelligence Unravelled: A Cockpit to Navigate IT-Capabilities, Regulatory Constraints, Business Processes and User Experience

Lilian Do Khac | Michael Leyer | Wieland Müller
Alexander Richter | Johannes Wichmann

Artificial General Intelligence Unravelled: A Cockpit to Navigate IT-Capabilities, Regulatory Constraints, Business Processes and User Experience

Whitepaper. V1.0. 22.9.2023

Authors

Lilian Do Khac, M.Sc. RWTH
Doctoral Candidate (External) at Group Digitalization and Process Management, University of Marburg
University of Marburg
Prof. Dr. Michael Leyer
Chairholder Research at Group Digitalization and Process Management, University of Marburg
Adjunct Professor at School of Management, Queensland University of Technology, Brisbane
University of Marburg
Wieland Müller, M.Sc.
Doctoral Candidate at Chair for Service Operations, University of Rostock
University of Rostock
Prof. Dr. Alexander Richter
Professor at Wellington School of Business and Government, Victoria University of Wellington
VU Wellington
Dr. Johannes Wichmann
Post-Doc at Group Digitalization and Process Management, University of Marburg
University of Marburg

Layout: Oliver Behn| Lilian Do Khac

About the Competence Center for Human-Centric Trustworthy AI

The concept of human-centric AI emphasizes the importance of designing AI technologies with a focus on empowering humans. It involves ethical considerations, transparency, fairness, privacy, and accountability in AI development and deployment.

You can find out more about the Competence Center for Human-Centric Trustworthy AI at https://www.uni-marburg.de/en/fb02/research-groups/business-administration/digitalization/research/competence-center-for-human-centric-trustworthy-ai

Members
Prof. Dr. Alexander Richter
VU Wellington
Prof. Dr. Florian Möslein
University of Marburg
Prof. Dr. Michael Leyer
University of Marburg
Prof. Dr. Sabrina Schneider
Management Center Innsbruck

Preamble

ChatGPT has put AI on top of managerial agendas in 2023. Management consultants and researchers seem to agree:

AI will disrupt most occupations and organisations.

Almost a decade ago Brynjolfsson & McAfee [1] delineated how AI will transform businesses; In the last couple of months ChatGPT and related platforms gave many of us a better understanding of what we can expect. Now, many companies are trying to grasp these newly introduced capabilities to transfer them into their business processes to reap efficiency gains or to create new products and services. However, this sensemaking process needs to consider regulatory constraints, and IT-capabilities and the human side: Expectations from customer and employee side. Though we are not there yet to enjoy or fear fully outgrown Artificial General Intelligence (AGI), the fundamental elements such as laying down the suitable hosting strategy of such models in line with compliance with regulation are the challenges of today.

In this whitepaper we provide insights necessary to navigate the challenges of integrating AGI into your businesses. We provide a toolkit that facilitate discussions and decisions necessary to unleash the power of AGI without underestimating related risks. Finding one's own strategy within such an unprecedented speed of technological disruption along with highly fuzzy concepts around ethic and responsibility demands for a radically new intercourse to uncertainty.

AI asks for more interdisciplinary skills that change common assumptions.

The future of doing business is changing right now and a new way with quite other partners and way of doing things materializes.

Key Points

- AI and foundation models develop with an unprecedented speed and challenge organisations around the globe: we witness changing customer expectations, expectations of employees and new requirements to IT-capabilities.
- At the same time digital regulation, particularly for AI, is evolving and expected to be present in the second half this current decade. Therefore, the development of these regulations must be monitored closely and applied in far-sight to avoid issues with compliance.
- The provision of such powerful AI models in business applications demands for a high-performance computing infrastructure and this means a significant investment decision. Decision-makers must derive from their desired business vision what they can integrate feasibly and how much risk they can afford to take.
- The integration of modern AI capabilities will fundamentally change job designs and is expected to lead to increases in productivity. The skills of human and machine collaboration as well as enabling factors for such are becoming a key competitive advantage

[1] Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton. https://books.google.de/books?id=WiKwAgAAQBAJ

1. Artificial general intelligence (AGI)

Artificial general intelligence (AGI) describes the extent of application areas by a respective AI model[2]. Many traditional AI models are limited in terms of their applicability (e.g. AlphaGo was designed for gaming). Recent developments, particularly large language models (LLMs), show that modern AI models can be applied to a large set of tasks[3].

Such fundamental models are trained without specific application case in mind but for the purpose of general application and can be applied to various tasks if the expected data input format is provided.

On a meta level three particular meta-use-cases can be derived for LLMs: (1) Domain Knowledge Agent, (2) Document Processing, (3) Text manipulation & generation (see figure 1).

The Domain Knowledge Agent is one of the most popular downstream applications for fundamental models as was demonstrated by OpenAI through the release of ChatGPT in November 2022. As such it is trained to interact in an anthropomorph way with human beings to answer knowledge retrieval related questions. It can deal with very heterogenous ways how knowledge is represented and how it is requested, making it very useful for dealing with the high variety in ways to communicate knowledge needs as a human.

Document processing in turn focuses on the heterogenous way particular information is depicted in documents and retrieving from this variety of representation the relevant information effectively or classifying these respectively.

Lastly LLMs also showed remarkable performance on generating text such as for software coding or editorial work.

Key terminology

" 'foundation models means an AI model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks" (EU AI Act Article 3 (1), p. 137).

" 'general purpose AI system' means an AI system that can be used in and adapted to a wide range of applications for which it was not intentionally and specifically designed" (EU AI Act Article 3 (1d).

[2] Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal of Artificial General Intelligence*, *5*(1), 1-48. https://doi.org/doi:10.2478/jagi-2014-0001

[3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Currently we are safe to say that we are amidst a fourth industrial revolution that is moving towards the age of AI. Whereas the previous industrial revolutions mainly focused on a holistic picture, thereby treating everything within it equal and centralized, today's industrial revolutions are aiming towards a more purpose driven approach which is more human-centric and in collaboration with machines.

Research[4] and politics stress the human-machine aspect as well as augmentation of humans with AI is highlighted. It is underlined that the sheer automation of tasks will not result in the same performance increases compared to previous industrial revolutions. Unlike narrow AI, which focuses on specific tasks or problems, AGI aims to develop systems that can learn from few examples, reason abstractly, and adapt to new situations. The focus is clearly placed on hybrid systems and the interplay of social and technical world. Therefore, the goal with artificial general intelligence would not be to enhance how we do things today.

The goals of AGI are to create machines with human-like intelligence that are able to understand, learn and adapt to a wide range of tasks. Some key objectives of AGI include replicating human intelligence to perform tasks requiring human-level intelligence, such as reasoning, problem-solving, decision-making, and understanding [5]. AGI seeks to create systems that can adapt and learn across various domains and tasks without extensive domain-specific data for training [6]. AGI has the potential to revolutionize various sectors, including education[7], healthcare and more[8] but also poses significant risks that need to be carefully managed, such as loss of control, development of unsafe targets and existential threats [9]. Also, ensuring the safety and ethical use of AGI is a critical goal. Developers must consider the potential risks and ensure that AGI systems adhere to ethical principles, including avoiding harm to humans and respecting privacy[10].

Addressing these issues and achieving the goals of the AGI requires a multidisciplinary approach involving computer scientists, engineers, social scientists, philosophers and policy makers.

[4] Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *The Academy of Management Review*, *46*(1), 192-210. https://doi.org/10.5465/amr.2018.0072

[5] Latif, E., Mai, G., Nyaaba, M., Wu, X., Liu, N., Lu, G., Li, S., Liu, T., & Zhai, X. (2023). Artificial general intelligence (AGI) for education. *arXiv preprint arXiv:2304.12479*.

[6] Ibid.

[7] Ibid.

[8] Becker, S. M., & Wright, C. B. (2021). Update on the Status and Impact of the National Eye Institute Audacious Goals Initiative for Regenerative Medicine. *J Ocul Pharmacol Ther*, *37*(3), 144-146. https://doi.org/10.1089/jop.2020.0015

[9] McLean, S., Read, G. J. M., Thompson, J., Baber, C., Stanton, N. A., & Salmon, P. M. (2023). The risks associated with Artificial General Intelligence: A systematic review. *Journal of Experimental & Theoretical Artificial Intelligence*, *35*(5), 649-663. https://doi.org/10.1080/0952813X.2021.1964003

[10] European Parliament. (2017). *Civil Law Rules on Robotics, https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html, retrieved on 12.09.2023*.

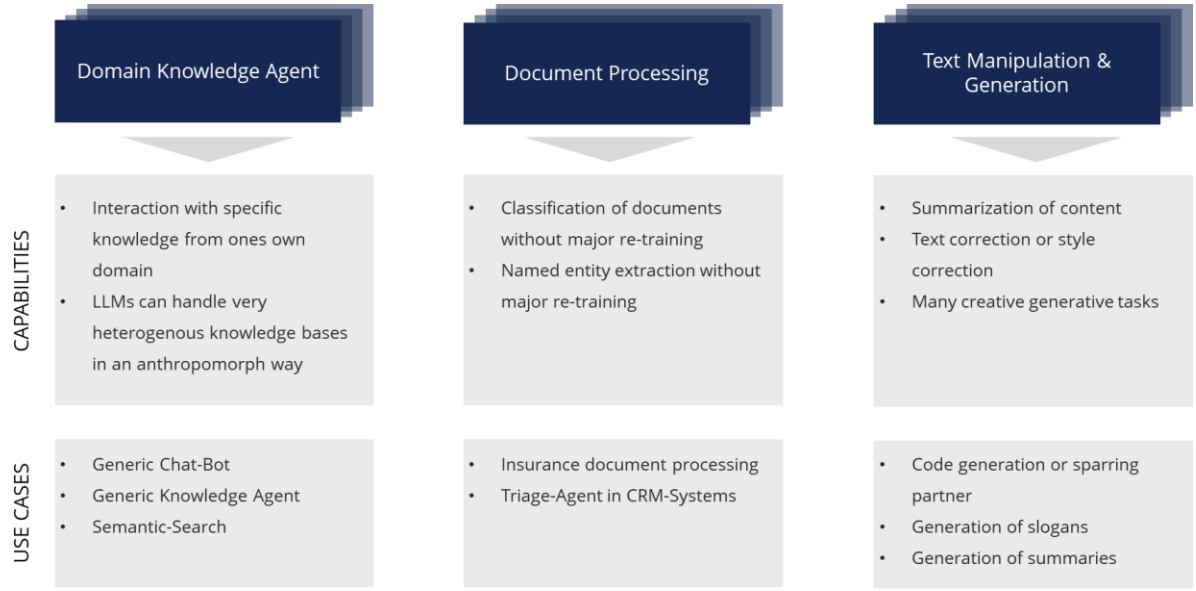


Figure 1 | Capabilities and Use-Cases of Large Language Models (LLMs)

2. AI regulation in the EU

Offerings of digital products and services with in-build- AI-capabilities within the European Union (EU) market are subject to laws. Three types of regulations considering AI products and services are particularly relevant: (1) EU-AI-Act (AIA)[11], (2) General Data Protection Regulation (GDPR)[12], (3) Copyright law[13].

According to the DRAFT Compromise Amendments on the AIA (Draft) as of 9.5.2023 the particularly addresses in Article 28b the provision of fundamental models. As such providers of fundamental models have obligations regarding transparency and diligence when putting their models into service or making them available. Among the obligations EU lawmakers require the reduction and mitigation of reasonably foreseeable risks, careful treatment of data sets used in those models, ensuring quality during the whole model lifecycle, environment sustainability, technical documentation and a quality management that ensures adherence to what has been enumerated before. Further, providers of foundation models are obliged to register within the EU Databases for cataloguing such models and to have audits by authorities conducted on their products.

The GDPR act applies to the processing of personal identifying information (PII) and the AIA is an attempt to regulate AI[14, 15] . The scope of the GDPR includes all PII of EU citizens or residents within the EU and outside of the EU when they are subject to data processing[16]. Since 2018, when the US Cloud (Clarifying Lawful Overseas Use of Data) Act came into effect, there has been legal uncertainty with its compatibility to the GDPR. The Cloud Act explicitly states that Cloud providers such as AWS, Google Cloud, and Microsoft must provide U.S. authorities access to PII stored or processed within their products, regardless of the geographic location of their compute resources. Under GDPR there must be a particular legal basis to lawfully process PII of (a) European citizens. Up until now this dispute has not been resolved. For a couple of LLM use cases there is some significant interaction with PII and therefore the GDPR is a significant legal risk that must be considered by companies aiming for utilizing LLMs. Within the DRAFT Compromise Amendments on the AIA (Draft) as of 9.5.2023 the AIA refers to existing privacy and data protection rules that must be considered during AI applications[17]. The

[11] European Parliament. (09.05.2023). DRAFT Compromise Amendments on the Draft Report EU-AI-Act, retrieved from ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf (europa.eu) on 25.06.2023. In.

[12] European Union. (04.05.2016). Regulation (EU) 2016/679 General Data Protection Regulation (GDPR) on EUR-Lex - 32016R0679 - EN - EUR-Lex (europa.eu), retrieved on 25.06.2023. In.

[13] For further regions see: https://dai.ki/navigating-ai-governance-a-comprehensive-look-at-existing-and-new-eu-and-us-ai-regulations/

[14] Schuett, J. (2023). Risk Management in the Artificial Intelligence Act. *European Journal of Risk Regulation*, 1-19. https://doi.org/10.1017/err.2023.1

[15] Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021). *Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI* Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event, Canada. https://doi.org/10.1145/3442188.3445923

[16] see GDPR Article 3 (1)

[17] see Article 4a, 1c, Draft Compromise Amendment AIA

existing privacy and data protection rule in the EU is the GDPR.

Most foundation models are provided through the cloud computing platforms of U.S. firms, therefore GDPR involvement must be evaluated. Should such case arise aspects of sovereignty have to be considered.

Under Article 4 of the EU Copyright Directive (EU DSM Directive 2019/790) the usage for "text and data mining" purposes of publicly available works is restricted in accordance with what a right holder explicitly and in machine-readable manner stated on their works. EU copyright law (and if respectively translated nationally) therefore can be taken as basis for deriving the legal base for the application of LLMs, such as for semantic searches, but not when such data went into the training of such foundation models unauthorized. It is unclear from an AI point of view how copyright law protects the issue of training foundation models with such data. Within the DRAFT Compromise Amendments on the AIA (Draft) as of 9.5.2023 the AIA states that providers of foundation models must "[…] document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law" [18]. One particularly significant source of training data for training LLMs is comprised of web data that has been crawled (such as Common Crawl).[19,20,21,22,23]

Most foundation models were trained on massive datasets crawled from the internet. It is not guaranteed that copyrighted data might have slipped into the collection and therefore would be subject to copyright law infringement. If so, then it would affect all foundation models (at least LLMs) equally.

[18] see Article 28b, 4c, Draft Compromise Amendment AIA
[19] https://docs.aleph-alpha.com/docs/introduction/model-card/
[20] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877-1901.
[21] Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., & Gehrmann, S. (2022). Palm: Scaling language modeling with pathways. *arXiv preprint arXiv:2204.02311*.
[22] Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., Castagné, R., Luccioni, A. S., Yvon, F., & Gallé, M. (2022). Bloom: A 176b-parameter open-access multilingual language model. *arXiv preprint arXiv:2211.05100*.
[23] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., & Azhar, F. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

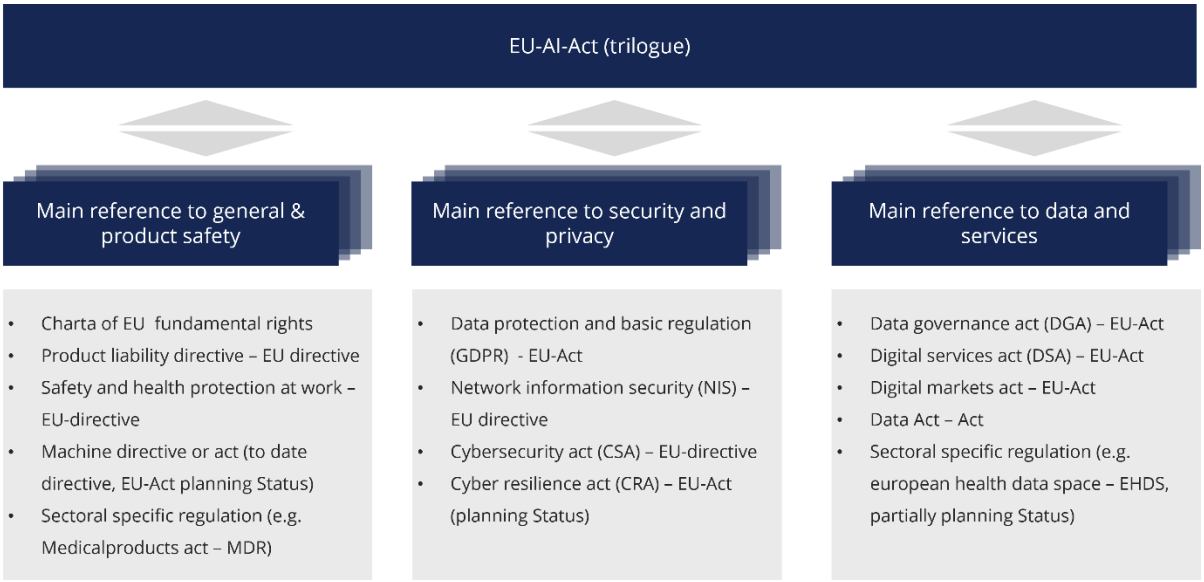


Figure 1 | Digital Lawmaking in Europe, Source: adapted from DIN/DKE (2022) German Standardization Roadmap on Artificial Intelligence, 2nd Edition, p. 23

3. User experience and satisfaction

User experience and satisfaction regarding AGI are complex and multifaceted, reflecting both the potential benefits and concerns associated with this technology.

The change from an offer- driven to a demand-driven market is a key driver of challenge that competitive organizations meet. As of the current industrialization paradigm more customer oriented, that is personalization, requirements must be met. It is likely to be one of the key activities for the upcoming industrialization phases. As such, expectations are not static but are dependent on customers experience, knowledge and desires towards AGI[24]. To evaluate their expectations, customers handle levels of expectations as well as expectation zones to end up in positive or negative disconfirmation[25]. If expectations are too high, it is likely that the performance ends up in an expectation zone that is below of what is acceptable to the customer. Performances in low expectation zones lead to negative disconfirmation and ultimately hinder customer satisfaction with AGI[26].

Depending on the field, different expectations emerge for an AGI, but primarily the AGI should solve complex problems[27] and meet its attributes such as "intelligent", "creative", "autonomous" and "flexible"[28]. For customer satisfaction, the alignment of expectations and perceived performances is mandatory.

As AGI continues to advance, there are also growing concerns among customers. Perceived trust and information privacy concerns influence the confirmation of expectations of customers towards AI and have to be considered[29]. Customers are also expressing concerns about AGI-powered systems potentially misusing personal data and the associated risks of cyberattacks and security breaches. Furthermore, there is an underlying worry that AGI-powered systems could perpetuate or even exacerbate existing biases and discrimination, particularly in critical areas like hiring and lending[30].

In addition to these concerns, customers may exhibit reluctance to trust AGI-powered systems if they lack an understanding of how these systems operate and make decisions[31]. Hence, AGI developers have to adhere to European regulations towards AGI as presented in section 2 to reduce information privacy concerns of customers towards their AGI. Further, by improving their interactions

[24]Oliver, R. L. (2010). *Satisfaction: A Behavioral Perspective on the Consumer: A Behavioral Perspective on the Consumer* (2. ed.). Routledge. https://doi.org/https://doi.org/10.4324/9781315700892

[25] Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, *17*, 460 - 469.

[26]Teas, R. K., & DeCarlo, T. E. (2004). An Examination and Extension of the Zone-of-Tolerance Model: A Comparison to Performance-Based Models of Perceived Quality. *Journal of Service Research*, *6*(3), 272-286. https://doi.org/10.1177/1094670503259408

[27] Monteith, S., Glenn, T., Geddes, J., Whybrow, P. C., Achtyes, E., & Bauer, M. (2022). Expectations for Artificial Intelligence (AI) in Psychiatry. *Current Psychiatry Reports*, *24*(11), 709-721. https://doi.org/10.1007/s11920-022-01378-5

[28] Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1-37. https://doi.org/doi:10.2478/jagi-2019-0002

[29]Brill, T., Munoz, L. S., & Miller, R. J. (2019). Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, *35*, 1401 - 1436.

[30] Pentina, I., Xie, T. L., Hancock, T., & Bailey, A. (2023). Consumer-machine relationships in the age of artificial intelligence: Systematic literature review and research directions. *Psychology & Marketing*, *40*(8), 1593-1614. https://doi.org/10.1002/mar.21853

[31] Ibid.

with customers, AGI developers are able to raise customers' perceived trust towards their AGI[32].

Such measures will be beneficial for AGI developers, as high customer satisfaction is known to be a key contributor of companies' success in terms of acquiring and retaining customers, premium pricing, increased customer value, and positive word-of-mouth communications[33].

[32]Stock, R. M., & Bednarek, M. (2014). As they sow, so shall they reap: Customers' influence on customer satisfaction at the customer interface. *Journal of the Academy of Marketing Science*, *42*(4), 400-414. https://doi.org/10.1007/s11747-013-0355-4

[33]Anderson, E. W. (1998). Customer Satisfaction and Word of Mouth. *Journal of Service Research*, *1*(1), 5-17. https://doi.org/10.1177/109467059800100102 , Bearden, W. O., & Teel, J. E. (1983). Selected determinants of consumer satisfaction and complaint reports. *Journal of Marketing Research*, *20*(1), 21-28. https://doi.org/10.2307/3151408 , Cronin. Jr, J., & Taylor, S. (1992). Measuring Service Quality - A Reexamination And Extension. *The Journal of Marketing*, *56*, 55-68. https://doi.org/10.2307/1252296 , Reinartz, W. J., & Kumar, V. (2003). The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing*, *67*(1), 77-99. https://doi.org/10.1509/jmkg.67.1.77.18589 ;

4. IT-Capabilities for AGI

GPT-3, through ChatGPT, essentially fuelled the appetite for foundation model applications that we witness nowadays. However, the provision options and emerging issue of sovereignty created a very complex decision-space for decision makers. Depending on the hosting paradigm (see **Figure *2***) there are various options how to integrate a foundation model into one's own internal IT-ecosystem.

One way is to use foundation models is to integrate them via the given hyperscaler (AWS, Azure and OpenAI, GCP) as IaaS/PaaS/AIaaS consumption, if there are no doubts concerning the location and thereby control of the data (see left column in **Figure *2***).

The available hyperscaler offering can be also deployed on-premises but for that the current requirements demand to deploy the whole cloud-stack of the respective cloud provider. This rarely is an economical solution since not everything of the cloud-stack is needed to run a foundation model. Further, platforms such as Huggingface (though underlying infrastructure needs to be chosen for which hyperscaler) or in Korea HyperCLOVA are available. Then again, the issue of control over the data is not given.

Other providers from whom to source a foundation model are mosaicML (U.S.A), AI21Labs (Israel), co:here (Canada), and Aleph Alpha (Germany). Despite Aleph Alpha, having their own high-performance compute cluster in Germany, all other providers are not necessarily under the provision of EU privacy law and therefore some issues related to specific applications are to be expected. Some foundation models such as the open-sourced ones and Aleph Alpha Luminous can be hosted on-premises without needing to deploy a full cloud-stack. But still there is a requirement to be fulfilled regarding the compute infrastructure to be able to run such a foundation model on-premises. These would be usually comprised of a high-performance compute cluster with some powerful GPU resources for running such models (see HPE, Dell etc.) which in turn represent a significant investment that needs to be made.

Other options that would avoid having to pay licensing costs for deploying an AI solution on-premise are Falcon (United Arab Emirates) or BLOOM (Europe). LLaMA 2.0 is not associated with a typical open-source licence but rather with a more commercially inspired one. These options have been trained already for 4 million EUR (assumed for Falcon and LLaMA 2.0) to 7 million EUR (BLOOM) and provided as open-source (or commercially free usage to a certain point of users for LLaMA 2.0), also for commercial use. Still the investment needs for a suitable on-premises infrastructure must be made and there is no way around it.

If one chooses to have more ownership regarding the foundation model such as running an open-source model on-premises then in accordance to expected EU-AI-Act your role within the supply chain potentially would be "provider" of AI and you move up the risk ladder. Along with it a set of transparency obligations as depicted in Section 2. Otherwise, if you do not substantially change a foundation model within your AI-application then you are potentially a "deployer" of AI with less associated risk.

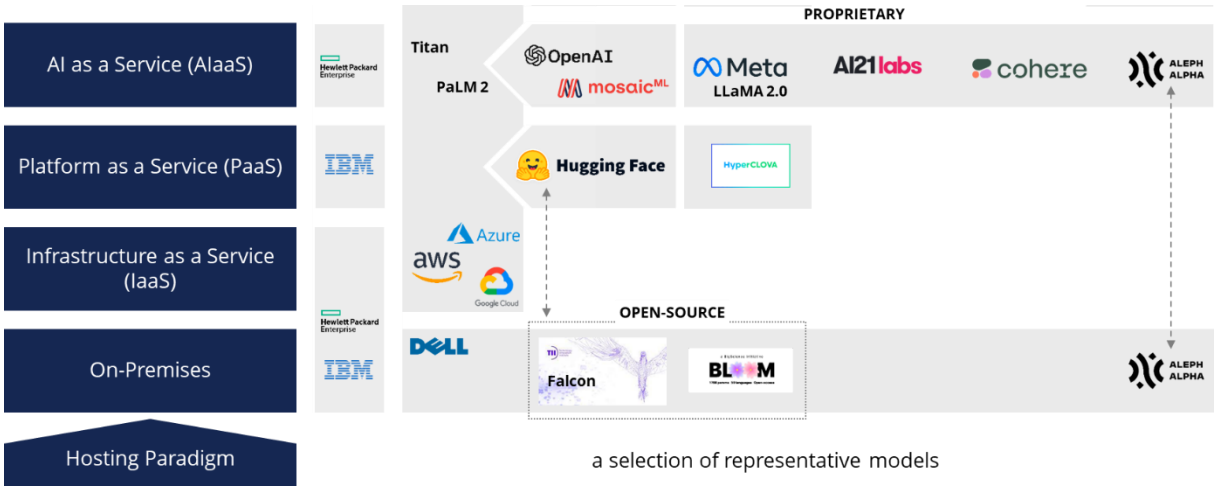


Figure 2 | Hosting Strategies for Foundation Models – A selection of representative models (provision models and/or foundation model)[34]

Key terminology

" '**provider'** means any natural or legal person, public authority, agency or other body develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge." (EU AI Act Article 3 (2))

" '**deployer** means any natural or legal person, public authority, agency or other body using an AI system under its authority, except where the AI system is used in the course of a personal non-professional activity" (EU AI Act Article 3 (4))

[34] Disclaimer: For this whitepaper Lilian Do Khac is primarily affiliated to Philips-University of Marburg in Germany. To finance her PhD engagement she works at an IT-Consulancy Firm, which integrates everything mentioned in this figure in Germany. There she holds among other roles also the Partnermanagement Role for maintaining the technology partnership with Aleph Alpha.

5. Business Processes

Recently, human-machine collaboration between AGI and humans is on the rise. Next to individuals that use AGIs in the form of LLMs for various purposes, such as research [35] , AGIs are important for management purposes as well. For instance, Haesevoets et al.[36] conducted an empirical study concerning the use of AGI in managerial decision-making. They determined that human managers prefer a share of decision-making in which they have a majority vote. Overall, they estimated that human managers prefer to perform 70% of managerial decision by themselves and 30% by AGIs.

Despite the current status, the meaning of AGIs for decisions and ultimately occupations will increase. In their publication "What Can Machines Learn and What Does it Mean for Occupations and the Economy?" Brynjolfsson et al. [37] build on previously conducted analyses for the potential of applying ML methods on tasks for various occupations. From there they build the "suitability for machine learning" (SML) metric that indicates tasks which can be transferred to machines and ML methods.

According to the authors the technological change which will be introduced with ML methods differs significantly in comparison to former technological changes towards automation. They identified that in many occupations there are some potentials to use ML methods for certain tasks, so AI is essentially pervasive and not tied to specific industries or occupation types.

Further, just a few occupations have SML across all the tasks, which would imply together with beforementioned that jobs are expected to rather change with AI but not necessarily to be redundant through AI. Therefore, they conclude that the main change and challenges that would come with AI in occupations and for the human resources is the redesign of jobs and not just only the introduction of further automation. It would imply a more interwoven human and machine interactional space that requires better interfaces between both parties to function well. These findings are repeated and underlined for specific investigations regarding foundation models such as GPT-4 through the International Labour Organization (ILO) in their recent working paper. [38] Therefore, organizations should

[35]Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., . . . Wright, R. (2023). Opinion Paper: "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*, 102642. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2023.102642

[36] Haesevoets, T., Cremer, D. D., Dierckx, K., & Hiel, A. V. (2021). Human-machine collaboration in managerial decision making. *Comput. Hum. Behav.*, *119*(C), 11. https://doi.org/10.1016/j.chb.2021.106730

[37] Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What Can Machines Learn, and What Does It Mean for Occupations and the Economy? *AEA Papers and Proceedings*, *108*, 43-47. https://doi.org/10.1257/pandp.20181019

[38] Gmyrek, P., Berg, J., Bescond, D. (2023). *Generative AI and Jobs: A global analysis of potential effects on job quantity and quality.* (ILO Working Paper 96, Issue.

enable their workforce through trainings and necessary guardrails to work optimally as well as compliantly with an AI.

It is expected that attributes such as empathetic and emotional intelligence, also referred to as "feeling economic"[39], will become more important for the future workforce. Research distinguishes former industrial revolution into physical economy (mechanical age), thinking economy (our current age), and feeling economy (future age). Within the thinking economy the main driver of economics is to think effectively whereas in the feeling economy the soft aspects will dominate. AI will enhance regarding the thinking part and will collaborate on that basis with humans who will take over the feeling part (e.g. communication with people, establishing relationships etc.).

Organizations that succeed will have trained their employees to interact with an AI such that performance is increased but also complacency and de-skilling, which is a risk regarding the human part, is mitigated. Organizations must balance the training for their experts and juniors, that will have different attitudes towards the AI agents.

[39] Huang, M.-H., Rust, R., & Maksimovic, V. (2019). The Feeling Economy: Managing in the Next Generation of Artificial Intelligence (AI). *California Management Review*, *61*(4), 43-65. https://doi.org/10.1177/0008125619863436

6. Framework

If the compliance requirements of the operating environment are defined by high compliance requirements, then the expectations on the IT-capabilities are likely to be restricted on on-premises and on sovereign stack at maximum. If otherwise, with less sensitive data in the processing, the options with hyperscalers or on a sovereign stack are likely. These two dimensions reflect the aspect of sovereignty.

If user experience requirements are very price sensitive or time sensitive (quick processing of documents for instance) then a high degree of automation is required. Otherwise, if a very well customized service or product is expected by customers then a human and machine collaboration, that is contrary to a full-automation, become very likely. These two dimensions reflect the aspect of satisfaction.

If compliance is fulfilled according to the necessary degree this also affects user experience and reflects the aspect of trust. The requirements on business processes define the required IT-capabilities and therefore reflect the aspect of integration.

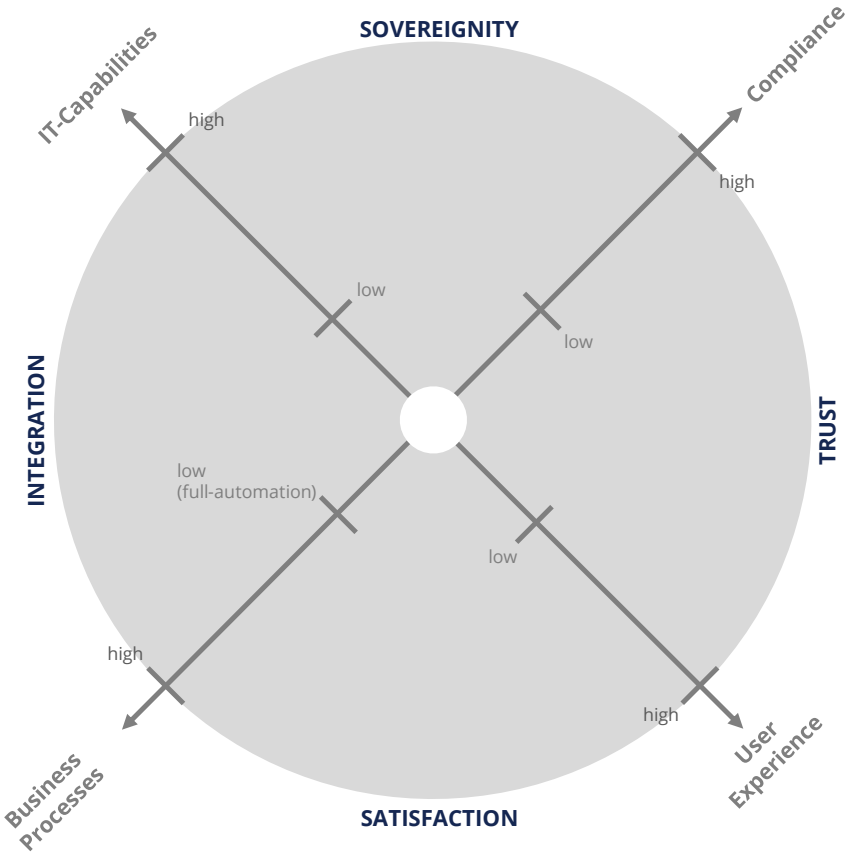


Figure 3 | Management Cockpit

7. Summary and Outlook

AGI will pave the way to the age of AI and it disrupts occupations and businesses of today. We see the unprecedented speed with which AI diffuses into the market and also the new emerging challenges around the globe. At the same time digital regulations is maturing and sets the ultimate normative boundaries, thereby forming a complex business environment. The provision of such powerful AI models is not only a matter of hardware capability but also about compliance capability, or in other words the responsible intercourse with AI.

AGI are not yet another digital tool to be integrated and applied. It is not about just starting to use AI and next generation AI in respective business processes; rather it involves the consideration of

- regulation,
- IT-capabilities,
- and human challenges

such as user experience and business processes. Clearly more interdisciplinary skills need to be considered and common assumptions are outdated.

These dynamics need to be steered by keeping an eye on

- sovereignty,
- trust,
- satisfaction (both, customer and employee) and
- optimal integration

into one's own organization.

The instantiation of these dimensions will be a constantly changing and evolving one. Companies can use the framework to keep an eye on these developments and structure them accordingly. Only then, they will be able to keep up with the radical changes and not loose oversight in the multitude of aspects to be considered.

In the upcoming months we will investigate this framework on the field within depth research to harden the dimensions and scales. We expect to find different patterns of this framework related to the industry or size of a company and thereby an answer how things might need to be.

Table of Contents

Related/Other work of the Competence Center

Tajabadi, M., Grabenhenrich, L., Ribeiro, A., Leyer, M., Heider, D. (forthcoming), Sharing data with shared benefits: An AI perspective, Journal of Medical Internet Research.

Gesk, T., Leyer, M., (2022), Artificial intelligence in public services: When and why citizens accept its usage, Government Information Quarterly, 39. Jg., Nr. 3, Artikel 101704.

Leyer, M., Schneider, S. (2021), Decision augmentation and automation with artificial intelligence. Threat or opportunity for managers?, Business Horizons, 64. Jg., Nr. 5, S. 711-724.

Schneider, S., Leyer, M. (2019), ME or Information Technology? Adoption of artificial intelligence in the delegation of personal strategic decisions, Managerial and Decision Economics, 40. Jg., Nr. 3, S. 223-231

References

Anderson, E. W. (1998). Customer Satisfaction and Word of Mouth. *Journal of Service Research*, *1*(1), 5-17. https://doi.org/10.1177/109467059800100102

Bearden, W. O., & Teel, J. E. (1983). Selected determinants of consumer satisfaction and complaint reports. *Journal of Marketing Research*, *20*(1), 21-28. https://doi.org/10.2307/3151408

Becker, S. M., & Wright, C. B. (2021). Update on the Status and Impact of the National Eye Institute Audacious Goals Initiative for Regenerative Medicine. *J Ocul Pharmacol Ther*, *37*(3), 144-146. https://doi.org/10.1089/jop.2020.0015

Brill, T., Munoz, L. S., & Miller, R. J. (2019). Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, *35*, 1401 - 1436.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877-1901.

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton. https://books.google.de/books?id=WiKwAgAAQBAJ

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What Can Machines Learn, and What Does It Mean for Occupations and the Economy? *AEA Papers and Proceedings*, *108*, 43-47. https://doi.org/10.1257/pandp.20181019

Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., & Gehrmann, S. (2022). Palm: Scaling language modeling with pathways. *arXiv preprint arXiv:2204.02311*.

Cronin. Jr, J., & Taylor, S. (1992). Measuring Service Quality - A Reexamination And Extension. *The Journal of Marketing*, *56*, 55-68. https://doi.org/10.2307/1252296

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., . . . Wright, R. (2023). Opinion Paper: "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*, 102642. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2023.102642

European Parliament. (09.05.2023). DRAFT Compromise Amendments on the Draft Report EU-AI-Act, retrieved from ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf (europa.eu) on 25.06.2023. In.

European Parliament. (2017). *Civil Law Rules on Robotics, https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html, retrieved on 12.09.2023*.

European Union. (04.05.2016). Regulation (EU) 2016/679 General Data Protection Regulation (GDPR) on EUR-Lex - 32016R0679 - EN - EUR-Lex (europa.eu), retrieved on 25.06.2023. In.

Gmyrek, P., Berg, J., Bescond, D. (2023). *Generative AI and Jobs: A global analysis of potential effects on job quantity and quality*. (ILO Working Paper 96, Issue.

Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal of Artificial General Intelligence*, *5*(1), 1-48. https://doi.org/doi:10.2478/jagi-2014-0001

Haesevoets, T., Cremer, D. D., Dierckx, K., & Hiel, A. V. (2021). Human-machine collaboration in managerial decision making. *Comput. Hum. Behav.*, *119*(C), 11. https://doi.org/10.1016/j.chb.2021.106730

Huang, M.-H., Rust, R., & Maksimovic, V. (2019). The Feeling Economy: Managing in the Next Generation of Artificial Intelligence (AI). *California Management Review*, *61*(4), 43-65. https://doi.org/10.1177/0008125619863436

Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021). *Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI* Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event, Canada. https://doi.org/10.1145/3442188.3445923

Latif, E., Mai, G., Nyaaba, M., Wu, X., Liu, N., Lu, G., Li, S., Liu, T., & Zhai, X. (2023). Artificial general intelligence (AGI) for education. *arXiv preprint arXiv:2304.12479*.

McLean, S., Read, G. J. M., Thompson, J., Baber, C., Stanton, N. A., & Salmon, P. M. (2023). The risks associated with Artificial General Intelligence: A systematic review. *Journal of Experimental & Theoretical Artificial Intelligence*, *35*(5), 649-663. https://doi.org/10.1080/0952813X.2021.1964003

Monteith, S., Glenn, T., Geddes, J., Whybrow, P. C., Achtyes, E., & Bauer, M. (2022). Expectations for Artificial Intelligence (AI) in Psychiatry. *Current Psychiatry Reports*, *24*(11), 709-721. https://doi.org/10.1007/s11920-022-01378-5

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, *17*, 460 - 469.

Oliver, R. L. (2010). *Satisfaction: A Behavioral Perspective on the Consumer: A Behavioral Perspective on the Consumer* (2. ed.). Routledge. https://doi.org/https://doi.org/10.4324/9781315700892

Pentina, I., Xie, T. L., Hancock, T., & Bailey, A. (2023). Consumer-machine relationships in the age of artificial intelligence: Systematic literature review and research directions. *Psychology & Marketing*, *40*(8), 1593-1614. https://doi.org/10.1002/mar.21853

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *The Academy of Management Review*, *46*(1), 192-210. https://doi.org/10.5465/amr.2018.0072

Reinartz, W. J., & Kumar, V. (2003). The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing*, *67*(1), 77-99. https://doi.org/10.1509/jmkg.67.1.77.18589

Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., Castagné, R., Luccioni, A. S., Yvon, F., & Gallé, M. (2022). Bloom: A 176b-parameter open-access multilingual language model. *arXiv preprint arXiv:2211.05100*.

Schuett, J. (2023). Risk Management in the Artificial Intelligence Act. *European Journal of Risk Regulation*, 1-19. https://doi.org/10.1017/err.2023.1

Stock, R. M., & Bednarek, M. (2014). As they sow, so shall they reap: Customers' influence on customer satisfaction at the customer interface. *Journal of the Academy of Marketing Science*, *42*(4), 400-414. https://doi.org/10.1007/s11747-013-0355-4

Teas, R. K., & DeCarlo, T. E. (2004). An Examination and Extension of the Zone-of-Tolerance Model: A Comparison to Performance-Based Models of Perceived Quality. *Journal of Service Research*, *6*(3), 272-286. https://doi.org/10.1177/1094670503259408

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., & Azhar, F. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1-37. https://doi.org/doi:10.2478/jagi-2019-0002