Artificial Intelligence in the Operational Management of Large-scale IT Systems

Perspectives for eu-LISA

Research and Technology Monitoring Report
July 2020

Legal notice

This report has been produced to provide up-to-date information to support the discussion on the future development and operation of large-scale IT systems operated by eu-LISA. Any views expressed in the report are entirely those of the author acting in his capacity as Capability Building Officer responsible for research and development, as well as other contributors, and are not necessarily the views of the Agency itself.

This report is public. Reproduction is authorised, except for commercial purposes, provided that the source is acknowledged.

Any direct or indirect references to products provided by specific vendors or to vendors themselves should not be taken as an endorsement by the Agency; references are provided purely for the purposes of illustrating the arguments made in the report.

All web links referred to in the report were confirmed to be functional on 26 June 2020.

Contact

To contact the main author for further information, please email: research@eulisa.europa.eu.

For enquiries regarding further use of the report or the information contained herein, please contact: communications@eulisa.europa.eu

ISBN 978-92-95217-67-6

DOI:10.2857/58386

Catalogue number: EL-04-20-018-EN-N

www.eulisa.europa.eu

Acknowledgements

The report is the result of collaboration of eu-LISA staff from across the Agency. Aleksandrs Cepilovs led the design and drafting of the report. The following colleagues contributed to the report (in alphabetical order): Grzegorz Borowski, Maria Bouligaraki, Jean-Claude Capoferri, Wulf Hemmerle, Ulrika Hurt, Alphonse Janson, Fabien Kuntzmann, Jean Letzelter, Philipp Mohrdiek, Donka Raytcheva, Padraig Ryan. Michael Evans provided editorial support. Anna Platonova and Alessandra Falcinella carried out the overall coordination of the work.

Preamble

According to Article 14 of its establishing regulation[1], eu-LISA shall monitor developments in research and technology relevant to the operational management of the large-scale IT systems operated by eu-LISA[2]. In line with Article 14 of the establishing regulation and Article 15 of the Revised Memorandum of Understanding between eu-LISA and the European Commission[3], eu-LISA shall publish at least one research and technology monitoring report per annum. This report, focusing on the possibilities for application of artificial intelligence in the operational management of large-scale IT systems, is the first deliverable towards fulfilling the obligation in 2020.

The report is intended for distribution to interested parties within EU/EEA Member States, the European Commission, EU agencies and other EU institutions. The report provides a high-level overview of the recent developments in artificial intelligence and their relevance to the core business of eu-LISA. The focus of this report is explicitly on the possibilities for application of artificial intelligence technologies in the context of eu-LISA business processes.

[1] Regulation (EU) 2018/1726 of the European Parliament and of the Council of 14 November 2018
[2] Eurodac, SIS II and VIS as of March 2020. ETIAS, EES and ECRIS-TCN are being developed at the time of writing of this report.
[3] Annex to the Commission Decision C(2019) 8941 final

Table of contents

Executive summary

Over the last few years, the debate around AI in the EU has picked up pace. In view of the implications of AI across the economy and society, and also the benefits that can be reaped from the widespread adoption of AI across all sectors of the economy, including the public sector, the EU has initiated the development of a European approach to AI. This coordinated EU approach to AI aims to put the EU forward as the leader in technology. It also aims to balance the advancement of technology and the need to protect and promote the public interest. To ensure this balance, the development of trustworthy and secure AI in the EU should be grounded in the common European values, such as respect of fundamental rights. In order to achieve these objectives, the Commission has launched a number of initiatives, including additional funding for research and innovation in AI, a coordinated plan on AI aimed at improving cooperation on AI in the EU, as well as the high-level expert group on AI focusing on the development of recommendations on AI-related ethical, legal and societal issues.

It is with this in mind that eu-LISA has been developing its approach to AI and the specific role the Agency can play in advancing the adoption of AI in the EU. First and foremost, the Agency can support the development and adoption of AI in the domain of borders, migration and security by, for example, supporting the necessary computational infrastructure for the development and testing of AI tools for the key stakeholders. Second, the Agency is well placed to benefit from the deployment of AI solutions in the context of the operational management of large-scale IT systems, in particular focusing on the improvement of performance, stability and the overall quality of the service it provides to stakeholders.

With the entry into force of the regulation establishing the Entry/Exit System (EES), eu-LISA has been mandated to develop the new system, which incorporates a component for automated biometric matching, which will rely on machine learning techniques for biometric matching. The scope for application of AI is not only limited to that, but includes a wide range of applications for AI which do not have an immediate effect on individuals. This report therefore largely focuses on applications where ethical and legal considerations are either not relevant or secondary; namely on those applications where the implementation of solutions relying on AI or ML technologies can have an immediate effect on the performance of eu-LISA as an IT service provider.

First, significant performance improvements can be attained by **deploying AI or ML solutions in the context of IT infrastructure and service management** (e.g. in application of performance monitoring, IT infrastructure and network monitoring and diagnostics, IT event analytics or IT service desk operations), in particular by increasing the availability of IT and network infrastructure, while at the same time reducing the workload of the staff responsible for it (e.g. by taking over some of the repetitive routine tasks). Such efficiency gains pave the way for devoting additional effort to further improvement of operations and services.

Second, **conversational agents, virtual assistants and chatbots** are another important class of AI tools in the context of eu-LISA core business. Chatbots can be effectively deployed in a number of areas, including providing training for future users of the EES, providing support to the users of ETIAS, as well as integrating chatbots as the first contact point for IT service desks. Use of chatbots or virtual assistants can improve the overall quality of services delivered by eu-LISA, speed up the adoption of new systems by users and improve the quality of information submitted by users using online systems.

Protection of IT systems from cyber threats is another domain where AI and ML can be put to effective use. At the most basic level, protecting digital infrastructure relies on cyber threat intelligence in order to identify vulnerabilities, as well as new attacks. Therefore, introducing automation to cyber threat intelligence using machine learning is an important first step. The second level of application of AI in cyber security is in network analysis and intrusion detection. And the next level is the creation of autonomous cyber security systems capable of threat detection and response without the need for immediate human intervention.

Finally, machine learning can be effectively used to **optimise the energy performance of data centre**

infrastructure, by improving the energy performance of the IT infrastructure and cooling systems. With the increasing complexity of data centres where electrical and mechanical equipment operate in constant interaction with non-linear patterns, the use of conventional methods to improve energy performance is no longer feasible. Statistical and more complex machine learning frameworks are useful in the development of energy efficiency models for data centres, allowing equipment performance to be optimised, taking into account real-time data.

Although AI has a great transformative potential, it does not come without challenges and risks. Therefore, organisations considering deployment of AI in either internal operations or services should bear in mind the following factors:

- Development of machine-learning based system will require large data sets in order to train the machine learning (ML) algorithm. In order to ensure that the AI system performs with a relatively low level of errors, the training data sets need to be of very high quality, potentially requiring use of real data.
- Development and deployment of AI based on machine learning algorithms requires substantial computational resources. Therefore, development of shared infrastructure that can be used for development and testing of AI solutions, such as the EuroHPC, will be necessary.
- AI systems – both rules-based and ML-based – are often not static, as they operate in a continuously changing environment. In order to take into account the changes in the environment, AI systems must be continuously updated if they are to remain accurate. This has direct implications for procurement of such systems, as well as the potential need for internal staff to support the continuous maintenance of such systems.
- It is important to look beyond the headlines and consider where AI can have significant effects without substantial barriers to its implementation. Such 'low hanging fruits' can help legitimise AI within the organisation and gain management support. They are also a good way to acquaint staff with AI systems, which will help with gradually developing and deploying more sophisticated systems where necessary.

This report is the first step for eu-LISA in its exploration of the full potential of AI for enhancing the performance of the Agency. We will continue to monitor the developments in AI and look deeper into some of the use cases as part of our research and technology monitoring activities in the future.

1. Introduction

Over the recent decades, digitalisation has had a transformative effect on human existence, and has also been an important driving force in public sector transformation, enabling the streamlining of public service delivery. However, digital transformation should not be perceived as a destination. Instead, it should be viewed as a process of continuous development, constantly on-boarding new technologies, such as artificial intelligence (AI) which has become one of the key variables in the equation. Although AI was mostly confined to science fiction, laboratories and academic discussions just a couple of decades ago, today we rely on it on a daily basis – sometimes unknowingly. The exponential growth in big data analytics and artificial intelligence applications has been fuelled by a combination of two factors, namely the availability of large-scale data sets and the growing availability of distributed computing power.

When discussing artificial intelligence, it is important to distinguish it from simple automation based on pre-programmed algorithms, which has existed for a long time. For the purpose of this discussion, we define AI as a concept that refers to a machine or a computer programme capable of observing its environment, learning and, based on the knowledge gained, proposing decisions or taking intelligent action. AI is more than simple automation in that it has the ability to function as an autonomous actor capable of engaging in activity that has not been explicitly pre-programmed.

The growth of computing power and availability of data have led to the development of new technologies and approaches in artificial intelligence, which already allow machines to outperform humans in specific tasks. Although in its current state of development AI is still far from human intelligence, as a general purpose technology, it has great potential to transform the way organisations conduct their business, and also societies more generally. In particular, AI offers great promise for the public sector at a time when public sector organisations are facing growing demands on their services in an increasingly complex environment.

Over the last few years, the debate around AI in the EU has picked up pace for a number of reasons. First, this is largely due to the realisation that AI has been deployed on a very large scale by social media and other platforms, posing significant risks to democracy and having a significant effect on other aspects of human existence[4]. Second, it has been widely recognised that the EU is lagging behind the US and China in terms of both development and deployment of AI solutions in practice. Although the EU has been one of the leaders in terms of basic AI research, it has trailed behind the US and China in reaping the benefits from this research[5]. Last is the expected positive economic effects (but also potential economic disruption) that can be caused by automation using AI.

One of the essential characteristics of AI is that development of these technologies requires massive amounts of data. Today it is often private companies operating online platforms that possess data on the necessary scale, rather than any single state. In this regard European enterprises are lagging behind, especially when it comes to consumer applications of AI, including online platforms[6]. However, Europe is leading in terms of digitalisation of industry (i.e. Industry 4.0) and the development of low-power computing devices, particularly relevant to the Internet of Things (edge computing) and high-performance computing[7]. A common European approach to data use and development of AI is thus necessary in order to avoid conflicting national initiatives and attain the necessary scale.

---

[4] https://www.technologyreview.com/2020/01/08/130983/were-fighting-fake-news-ai-bots-by-using-more-ai-thats-a-mistake/ or Graham et al. (2020) *Artificial intelligence in hiring: Assessing impacts on equality*. Institute for the future of work.

[5] Renda A. (2019) *Artificial Intelligence. Ethics, governance and policy challenges. Report of the CEPS Task Force.* Brussels: Centre for European Policy Studies. Available online: https://www.ceps.eu/wp-content/uploads/2019/02/AI_TFR.pdf

[6] (ibid.)

[7] COM(2020) 65 final. White Paper On Artificial Intelligence – A European approach to excellence and trust. Available online: https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

In view of the implications of AI across the economy and society, and also the benefits that can be drawn from the widespread adoption of AI across all sectors of the economy, including the public sector, the EU has initiated the development of a European approach to AI[8]. This coordinated approach aims to put the EU forward as the leader in technology. It also aims to balance the advancement of technology and the need to protect and promote the public interest. To ensure this balance, the development of trustworthy and secure AI in the EU should be grounded in the common European values, such as respect of fundamental rights. In order to achieve these objectives, the Commission has launched a number of initiatives, such as additional funding for research and innovation in AI, a coordinated plan on AI[9] aimed at improving cooperation on AI in the EU, as well as the high-level expert group on AI[10] focusing on the development of recommendations on AI-related ethical, legal and societal issues[11].

In its white paper on AI[12], the Commission proposed developing an ecosystem of excellence, with the aim of facilitating the development and widespread uptake of AI across the EU, including in both private and public sectors. In order to advance the development of the ecosystem, the Commission proposed a comprehensive approach addressing the following areas:

- Coordination of financing targeting the development of AI technologies across the EU Member States, with the aim of attracting over €20bn of investment in AI over the next decade[13].
- Coordination of effort in research and development in the AI domain, in particular advancing the development of AI in those areas where Europe is already leading, such as industrial automation, transport and logistics, finance, energy, forestry and agriculture.
- Development of centres of excellence will be supported through the existing and new financial instruments, including the Digital Europe Programme and the Horizon Europe, among others.
- Development of skills to both address the shortage of competences and skills necessary for the development of new AI-based technologies (as well as up-skilling employees to be fit to work with AI) and re-skill those employees whose jobs will be affected by automation.
- Support the adoption of AI not only within large corporations but also by SMEs, through such initiatives as Digital Innovation Hubs[14] and AI-on-demand platform[15].
- Securing access to data and computing infrastructures, including by investing in further development of the European super-computing infrastructures and in quantum computing. These priorities are developed further in the European data strategy accompanying the white paper on AI[16].
- Lastly, the adoption of AI by the public sector, which is well placed to benefit from the adoption of AI. First, the public sector is responsible for the provision of a wide range of data-driven public services which, while taking the necessary precautions and ensuring compliance with the relevant regulations, can gain in terms of both efficiency and quality. Widespread adoption of AI is especially relevant in healthcare and transport – two areas where the public sector is prevalent in Europe – where technologies are available for large-scale deployment. AI is also increasingly relevant for the protection of public safety and security, especially cybersecurity.

The advances in AI have great potential to transform business processes and the operation of organisations in various domains, including Justice and Home Affairs (JHA). In particular, application of AI systems could

[8] COM(2018) 237 final. Artificial Intelligence for Europe. Available online: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51625
[9] https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence
[10] https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence
[11] https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence
[12] COM(2020) 65 final. White Paper On Artificial Intelligence – A European approach to excellence and trust.
[13] COM(2018) 237 final. Artificial Intelligence for Europe.
[14] https://ec.europa.eu/digital-single-market/en/digital-innovation-hubs
[15] https://www.ai4eu.eu/
[16] https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

enhance decision making in policing and law enforcement, judicial affairs, border management and migration. AI can also help improve processes, leading to greater effectiveness and efficiency of operations on the ground. However, certain challenges remain that need to be overcome before widespread implementation of AI becomes feasible for public sector applications, especially when AI-based decisions have legal implications[17]. These legal constraints, including those directly related to data protection requirements, need to be addressed by the JHA community before AI systems can be developed and deployed at scale in the respective areas.

It is with this in mind that eu-LISA has been developing its approach to AI and the specific role the Agency can play in advancing the adoption of AI in the EU. First and foremost, the Agency can support the development and adoption of AI in the domain of borders, migration and security by, for example, supporting the necessary computational infrastructure for the development and testing of AI tools for the key stakeholders. Second, the Agency is well placed to benefit from the deployment of AI solutions in the context of the operational management of large-scale IT systems, in particular focusing on the improvement of performance, stability and the overall quality of the service it provides to stakeholders.

With the entry into force of the regulation establishing the EES[18], eu-LISA has been mandated to develop the new system which incorporates a component for automated biometric matching, which will rely on machine learning techniques for biometric matching. In light of this, in this report we focus on the technologies relevant to the operational management of large-scale IT systems and biometric matching, leaving other applications of AI outside of the scope of this report.

Given the complexity of AI technologies and of the related challenges that need to be addressed, AI will be one of the priority technologies for the Agency. Assessment of AI technologies and their relevance is especially relevant in the context of the strategic goals set by the Agency for the period of 2018-2022, in particular with the aim of becoming the principle EU ICT technology hub and centre of excellence, as well as continuous development of an efficient and agile organisation. This report is a first step in the Agency's efforts in research and technology monitoring in the field of big data analytics and AI in the longer term. The Agency will continue monitoring the developments in research, technology and innovation in AI and the applicability of new technologies within the core business of the Agency as part of research and technology monitoring activities.

The report is structured as follows. In the first section, we set the scene for the discussion with a brief overview of the history of AI, including its most recent developments. We then provide an overview of the key concepts relevant to the discussion on AI and machine learning. In the second section, we provide a concise overview of the relevant initiatives for piloting or implementing AI or machine learning in the public sector context, with a specific emphasis on the relevant cases from the JHA domain. In the third section, we address the possibilities for the application of AI in eu-LISA, focusing on the operational management of large-scale IT systems. We then provide descriptions of example use cases and the relevant technological solutions. In the final section of the paper, we discuss the solutions proposed and provide some preliminary conclusions, as well as outline next steps for further investigation.

[17] https://ec.europa.eu/jrc/en/publication/legal-and-regulatory-implications-artificial-intelligence-case-autonomous-vehicles-m-health-and-data
[18] REGULATION (EU) 2017/2226 of 30 November 2017.

2. Basic AI concepts

2.1. A brief overview of AI methods and applications

Although AI has permeated many aspects of our daily lives by changing the way we work, travel, shop or find partners, it did so apparently despite of users' lack of understanding of what was going on inside the proverbial 'black box'. In spite of the general suspicion towards AI widespread among non-specialist and techno-pessimist audiences, in the opposing camp there is a strong belief in the promise of AI to help build a better world. It can be argued that AI, just like any technology before it, including electricity, genetic engineering or the internet, can be harnessed for the benefit of humanity at large, assuming that the necessary preventive measures have been put in place constraining the space for opportunity to use technology against humanity.

First, it is important to distinguish between narrow or specific AI and general AI, sometimes also referred to as machine superintelligence – artificial intelligence equalling or surpassing the level of human intelligence.

- **Narrow AI** can be defined as the application of AI in a narrowly defined domain, performing a strictly defined functions or tasks. All currently available AI technologies can be defined as narrow AI, given that they can only carry out specific tasks, albeit often better than humans (e.g. solving complex mathematical problems, sorting large quantities of data, or playing chess or Go).
- **General AI** or machine superintelligence can be defined as artificial intelligence that is equal to or surpasses human intelligence in all domains of human activity.

Although science fiction has been concerned with the effects of general AI for a few decades already, most scientists believe that despite the major progress in artificial intelligence, such as the development of artificial neural networks and deep learning, we are still at least a few decades away from the development of general purpose AI[19]. As argued elsewhere, the fact that an artificially intelligent system has reportedly passed the Turing test[20] does not mean that AI has reached the level of human intelligence; it only points to the deficiencies of the test itself[21].

Second, it is important to distinguish between the different approaches or stages of evolution of AI, starting with symbolic AI, to machine learning, to artificial neural networks and deep learning. Symbolic or classical AI was the dominant paradigm of AI during the three decades from the mid-1950s to the late 1980s. At the core of these systems is a computer programme based on expert knowledge. Such systems follow step-by-step 'if-then-else' procedures in order to produce the intended outcomes. In the case of relatively simple processes, expert systems can rely on true/false values to provide inputs for automatic decision making. In more complex cases, such as medical diagnoses or stock market trading platforms, expert systems rely on fuzzy logic to deal with a larger number of variables and values, where the decision is made on the basis of a 'truth value' within the range of 0 to 1[22]. The 'human in the loop' principle applies to such systems by default, since the systems essentially follow human decision-making procedures. Since such systems are developed using relatively simple code, the decision-making processes in such systems can be easily subjected to human audit.

[19] National Science and Technology Council (2016) Preparing for the future of artificial intelligence. Washington, DC: Executive Office of the President. Available online: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf

[20] Turing test is a test of a machine's ability to exhibit intelligent behaviour equivalent to or indistinguishable from that of a human being, named after the English computer scientist Alan Turing.

[21] The Eugene Goostman chatbot developed by a team of Russian and Ukrainian programmers passed the test in 2014 by tricking non-expert judges using a number of techniques, such as an encoded personality (of a 13-year-old Ukrainian boy with imperfect English), as well as using jokes or changing the subject of conversation (akin to experienced politicians) whenever the question asked did not correspond with a pre-programmed response. This announcement was received with widely shared scepticism by AI experts. Mitchell (2019) Artificial Intelligence. A Guide for Thinking Humans. New York: Farrar, Straus and Giroux.

[22] https://www.controleng.com/articles/artificial-intelligence-fuzzy-logic-explained/

Although expert systems are very useful in reducing the workload necessary for making decisions, such systems can only be applied to narrowly and precisely defined problems with a limited number of variables. Hence, application of such systems to real-world problems with a very large number of variables and values becomes technically challenging, as with the increasing number of variable, the complexity of rules grows exponentially and requires extensive effort in order to formalise the necessary knowledge. Another disadvantage of expert systems is that they can only be improved with human intervention and not by themselves. Nevertheless, expert systems are still widely applied in business environments where humans perform repetitive tasks in well-defined environments; those are relatively robust, efficient and allow for fairly simple human auditability.[23]

The next stage in the development of AI can be broadly described as data-driven AI, which involves machine learning algorithms – including artificial neural networks and deep learning – that are able to improve their performance without human supervision by relying solely on the analysis of training data.

Although in theory training of machine learning systems using human interventions is possible, in practice such training is not feasible due to the very large number of human interventions that would be necessary to train an ML system. Therefore, machine learning algorithms are applied to automate the learning process. Different approaches are used for training ML systems, including supervised learning, semi-supervised learning, reinforcement learning and unsupervised learning, each of which has its own advantages and disadvantages[24].

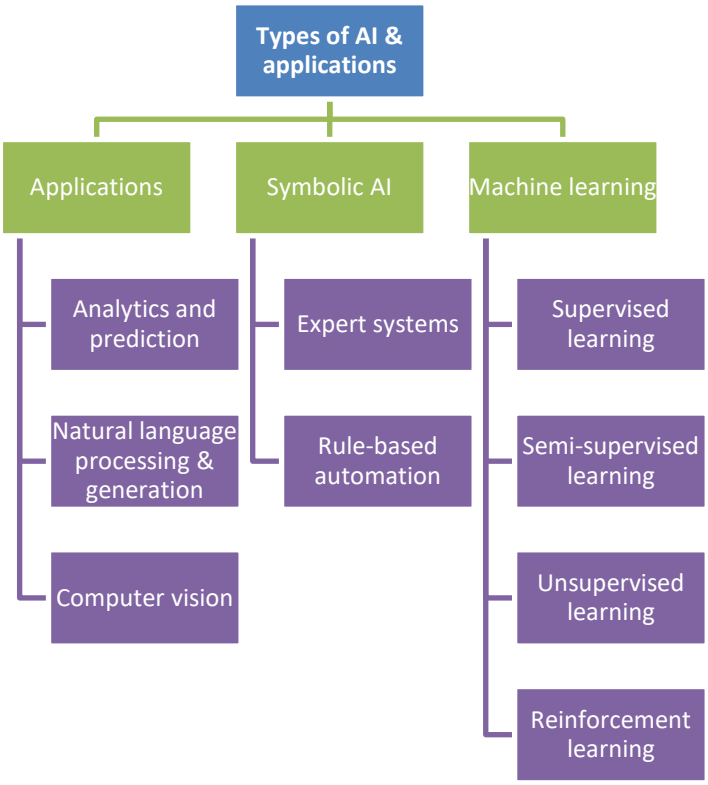

Figure 1: Types of AI and applications (source: authors)

The discussion above, and even a cursory review of the literature on artificial intelligence, suggests that data is the key element in the development of efficient and effective machine learning algorithms. Data intensity is the fundamental distinction between the symbolic AI of the 1980s and the contemporary approaches to AI based on machine learning (and specifically neural networks and deep learning).

[23] Boucher (2019) How Artificial Intelligence Works. European Parliamentary Research Service. Available online: https://www.europarl.europa.eu/at-your-service/files/be-heard/religious-and-non-confessional-dialogue/events/en-20190319-how-artificial-intelligence-works.pdf

[24] For an overview of the basic concepts in machine learning see Burkov (2019) The Hundred-Page Machine Learning Book

2.2. Data as the foundation of modern AI

As mentioned above, modern AI technologies based on machine learning are very data intensive. Although substantial work has been taking place in the development of machine learning algorithms, in particular focusing on the development of neural networks and deep learning methods, most recent advances in artificial intelligence have been caused by two interrelated factors, namely the availability of data and the availability of computing power to process this data.

Although the restrictions imposed by the GDPR on the use of data might constrain the development of AI technologies within the EU, they will also create a unique environment for the development of AI technologies, which will integrate personal data protection by default. Resolving the contradiction between the need for privacy protection and the availability of data for the development of AI is one of the challenges on the road to developing European AI. In addition to the high-level policy challenges, there are more practical challenges related to the availability and quality of data used for AI training purposes. In order to tackle these challenges systematically and build a good foundation for AI, Monica Rogati has proposed a derivative of Maslow's hierarchy of needs, aptly called 'The AI Hierarchy of Needs'[25].

At the base of the pyramid is **data collection**. Each organisation exploring applying AI, be it in internal processes or in service delivery, should consider what data is necessary and what is available. If some of the data is not being collected, it should be considered whether such data can be collected and at what cost. The next step in the data pyramid includes **data storage and data flows**. Organisational structures, data storage and sharing policies and data protection requirements all affect the possibility of using data for basic analytics and training AI. After we have resolved the data storage and data flow issues, as well as defined policies on data use for analytics purposes, the data needs to be prepared. **Data preparation** involves the following: cleaning and identification of potential anomalies in the data; data anonymisation when necessary; definition of features of the data set and data segmentation; definition of metrics for business analytics or features for the ML model; and last but not least, data labelling.

The final step before launching AI in production would be to test the development of algorithms on a testing data set to identify possible issues that may affect the results when the tools are rolled out in production mode. As argued in the focus paper on data quality in the development of machine learning and artificial intelligence published by the European Union Agency for Fundamental Rights[26], low data quality can have a significant impact on a number of fundamental rights, most importantly the right to non-discrimination. Testing AI solutions on representative data sets enables such biases and errors in systems for automated decision making to be reduced, if not completely avoided. Specific challenges related to training of automated biometric matching systems will be addressed in more detail below.

Contemporary public sector organisations are awash with data in both structured and unstructured forms. In fact, until the emergence of social media and other digital platforms, public sector organisations had a virtual monopoly over the collection, storage and processing of personal and non-personal data on a very large scale. In addition to collecting, public sector organisations produced large amounts of structured and unstructured data (e.g. operational or legislative). However, a significant share of this data still does not exist in a machine-readable format. Therefore, during the early stages of development of AI systems for use in the public sector, substantial human expert resources would be necessary to render this data useful for training of AI or processing by machines.

---

[25] https://medium.com/hackernoon/the-ai-hierarchy-of-needs-18f111fcc007

[26] European Union Agency for Fundamental Rights (2019) Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights. Available online: https://fra.europa.eu/en/publication/2019/artificial-intelligence-data-quality

Although data quality has been widely recognised as an essential precondition for the development of AI and machine learning algorithms, in particular when they are intended for application in the public sector, due to the novelty of the field there are no standards defining data quality for AI training or for standard descriptions of data sets. This is one of the fields in which the European Commission and EU agencies can make a substantive contribution.

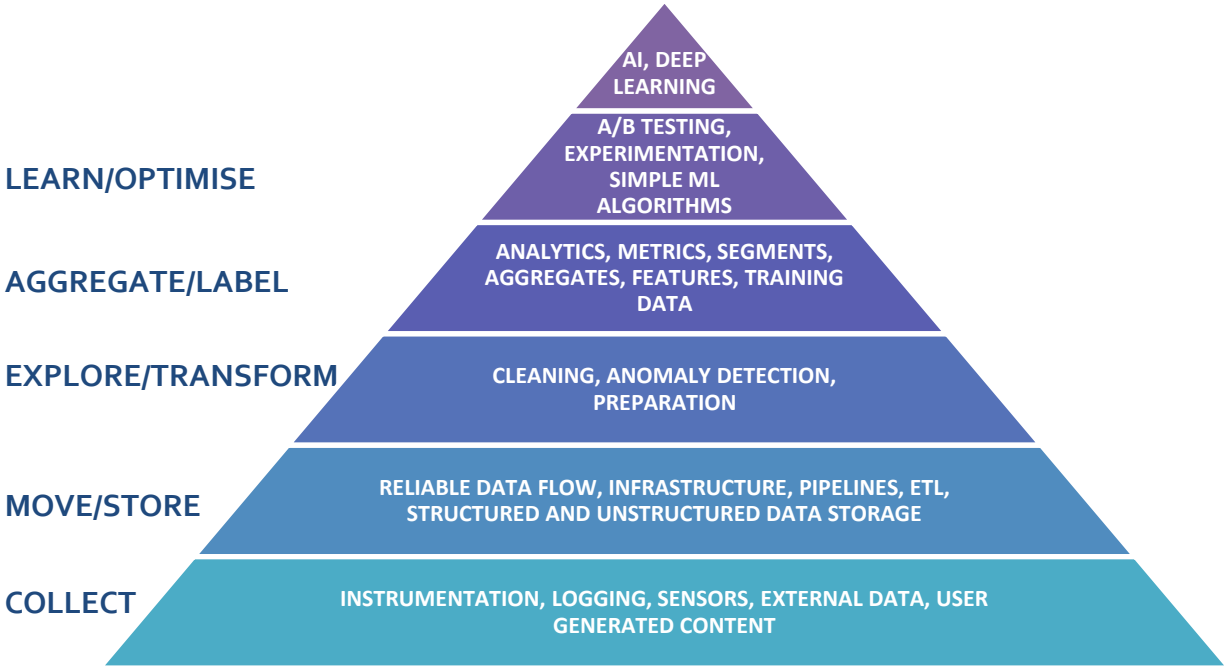

Figure 2: The AI hierarchy of needs (adapted from Monica Rogati)

As already mentioned, contemporary approaches to AI are highly data and computation-intensive. The recent explosion in AI research and deployment of AI-based solutions has led to concerns related to not only algorithmic decision making and data protection, but also the long-term sustainability of modern AI[27], especially in light of climate change. In part as a response to the concerns regarding energy efficiency of AI based on deep learning, but also addressing the need for on-device AI, new methods for optimising the performance of artificial intelligence algorithms are being developed today. Some of these approaches focus narrowly on optimisation of the performance of AI systems with the aim of reducing the amount of processing power needed to deploy an algorithm without a significant negative effect on its performance. For example, researchers at MIT have developed an approach for efficient video understanding[28], which improves the performance of video analysis algorithms deployed, for example, in autonomous driving. In the domain of natural language processing, Google and Huawei have recently improved the performance of the well-known Google language model BERT, which in its large version includes approximately 340 million data parameters. Using the knowledge distillation technique[29], Google was able to improve the efficiency of the model 60-fold, while only marginally reducing its accuracy[30]. These approaches, however, focus on narrowly improving the performance of existing deep learning algorithms.

Another stream in improving the performance of AI systems, including improvement of their efficiency, takes a different approach aimed at the development of AI algorithms that try to imitate humans in their operation. Such approaches also aim to produce algorithms that can serve as a pathway towards general artificial

[27] https://www.technologyreview.com/s/614700/the-computing-power-needed-to-train-ai-is-now-rising-seven-times-faster-than-ever-before/
[28] Lin et al. (2019) TSM: Temporal Shift Module for Efficient Video Understanding. Available online: https://arxiv.org/pdf/1811.08383.pdf
[29] An approach in which a master model trains a smaller apprentice model until it reaches the performance of the master model.
[30] https://www.technologyreview.com/f/614473/tiny-ai-could-supercharge-autocorrect-voice-assistants-on-your-phone/

intelligence. Researchers from Vicarious AI have developed a computer vision model which is 300-fold more data efficient at resolving text-based CAPTCHAs than deep neural networks. The proposed model is capable of generalising from a small number of training examples in performing a range of visual cognition tasks[31]. The proposed approach, incorporating inductive biases from systems neuroscience in the development of computer vision systems, can lead to the development of generalisable machine-learning models operating with high data efficiency. The authors suggest that building on this work and combining neural networks with structured probabilistic models may eventually lead towards general purpose AI systems. In the near term, however, new approaches to improving the efficiency of AI systems could lead to the development of high quality on-device AI applications, which may be of direct relevance to addressing the challenges of the JHA community, such as highly precise biometric on-device recognition and identification.

3. AI applications relevant to eu-LISA: opportunities and challenges

If you have followed the news over the last couple of years, you have no doubt seen at least a few headlines about AI outperforming humans or the potential for AI to displace millions of humans from their jobs. While the former sounds very promising and the latter rather alarming, the reality is likely somewhere in between. As already described, existing AI performs extremely well, indeed exceeding human performance, but only at a very narrowly defined task, e.g. playing Go or recognising patterns in text or visual media. Whenever the task becomes more complex and requires contextual understanding, the machine will likely either fail at the task or perform substantially below humans. It is therefore unlikely that the current approaches to AI, predominantly based on machine learning, will be able to replace humans entirely in the near future[32]. Where AI can indeed have a significant effect is in augmenting human performance by either providing support with information processing and providing inputs to decision making, or performing routine tasks and freeing up time for humans to engage in activities that require human involvement.

Davenport and Ronanki[33] offer a useful categorisation of AI technologies based on business capabilities:

- **Process automation:** automation of physical or digital tasks – primarily administrative back office and financial activities – using robotic process automation. These tools can be used, for example, in processing requests received by an IT service desk.
- **Cognitive insight:** detecting patterns in large data sets and across numerous data sets, and interpreting their meaning. Cognitive insight systems are largely based on some kind of machine learning algorithms and are used for predicting outcomes based on data.
- **Cognitive engagement:** customer engagement using natural language processing and generation. These systems often take the form of a chatbot or an automated voice-controlled assistant (e.g. Apple's Siri, Google Assistant or Amazon's Alexa).

In addition to the above, **computer vision** is being increasingly used in the public sector. Computer vision has been successfully applied in autonomous driving, detection and diagnostics of cancer in healthcare, detection and monitoring of oil spills in the oceans, as well as safety and security applications including border management. In the table below we provide a brief overview of AI applications in the public sector (see Annex I for a more detailed overview of possible applications).

---

[31] George et al. (2017) A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs. Available online: https://science.sciencemag.org/content/358/6368/eaag2612

[32] Some notable exceptions include autonomous driving and some areas of medical diagnostics where AI is already on a par with, or better than, humans.

[33] https://hbr.org/2018/01/artificial-intelligence-for-the-real-world

In this section of the report, we look at some of the applications of AI that are relevant to the operational management of large-scale IT systems. We first discuss computer vision and its possible applications. We then provide an overview of natural language processing and generation technologies, which is followed by a brief discussion of AI systems for decision support and decision making. We close this section with a discussion on computational infrastructure necessary for the development and deployment of AI.

3.1. Computer vision

Computer vision is perhaps one of the more common applications of AI, and the application that has attracted considerable global attention. Computer vision uses pattern recognition techniques for identifying objects in visual media (both static and moving images). The current paradigm in computer vision is largely based on supervised AI, which uses large data sets of human-labelled images for training. Deep learning models for computer vision underwent a period of very rapid development in recent years, in particular for object recognition and localisation tasks. The main challenge, however, is that these models require the existence of large-scale human-labelled data sets. The production of human-labelled training data sets is costly, in particular where images not only need to be labelled, but where more description of the images needs to be provided (e.g. location of certain elements in the image and description of those elements).

The recent advances in the development of unsupervised deep learning techniques have enabled unlabelled data sets to be used. One such technique is Tensorflow Graphics[34], which relies on computer vision in conjunction with computer graphics and a back-propagation training algorithm in an approach called analysis by synthesis. The computer vision system extracts the parameters of the scene in the image, and the computer graphics system reproduces the image on the basis of the extracted parameters. Comparison of the original and the reproduced images informs the system whether the parameters of the image were identified correctly or not, and adjusts the computer vision system accordingly. This approach allows self-supervised training of computer vision systems, but relies on large volumes of data and computation in order to achieve high performance, and is therefore not particularly energy efficient.

Computer vision models are already widely applied in the analysis of medical imaging and autonomous vehicles as well as in defence and security (e.g. autonomous aerial vehicles). In consumer products, computer vision has been widely applied in authentication systems for personal devices, autonomous vehicles, home security equipment and consumer robots (e.g. life companions).

The most relevant use case for the application of computer vision technologies within the scope of the Agency's core business is the application of facial recognition technologies (FRT) at border crossings for automated identification and biometric matching of travellers, as laid down in the regulation establishing the EES[35].

[34] https://www.tensorflow.org/graphics/overview
[35] Regulation (EU) 2017/2226, 30 November 2017.

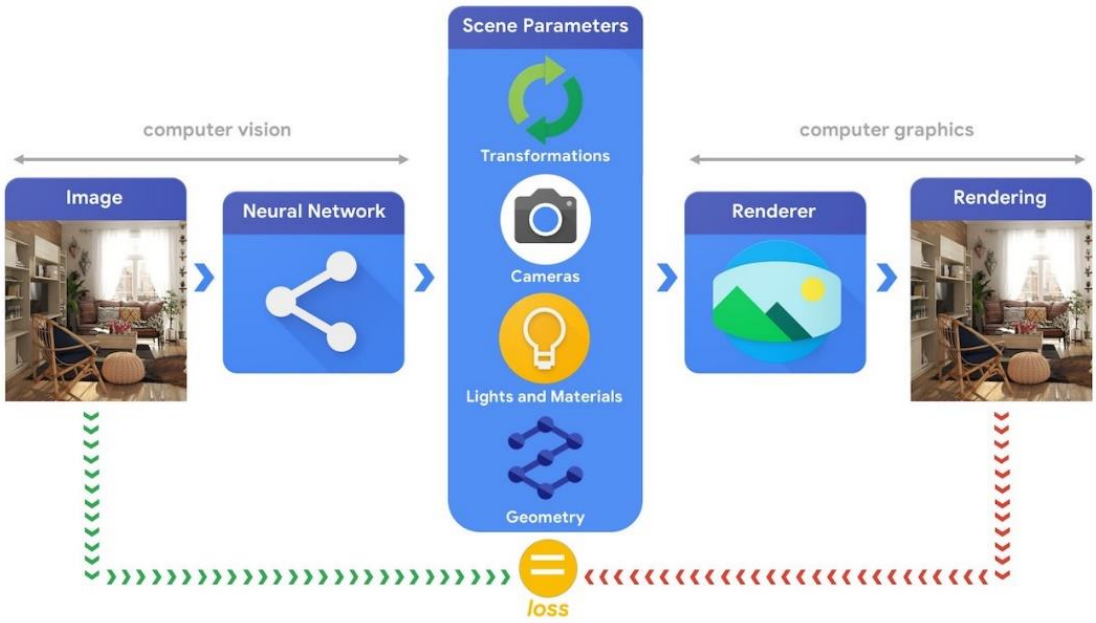

Figure 3: Conceptual representation of the Tensorflow graphics system (source: https://www.tensorflow.org/graphics/overview)

Computer vision systems for facial recognition can be implemented in two scenarios. The first is authentication of an individual, whereby a facial image is captured and compared in a one-to-one scenario. One-to-one scenarios, where a captured image is compared to the image stored on an identity document or in a database with a link to the identity document, are relatively simple and do not require significant processing power, and can therefore be easily implemented on a mobile device (e.g. smartphone). Other challenges related to the deployment of FRT for authentication are related to the acquisition of high quality images to ensure high quality of operation of the facial recognition systems. Facial image capture is often performed under conditions which are far from perfect, especially when in uncontrolled environments at land border crossing points. The quality of the captured images and therefore also the results of automatic facial recognition can be affected by insufficient lighting, weather elements or movement of the subject.

The second scenario is identification of individuals at border crossing points. This scenario relies on the 'one-to-many' identification, whereby a facial image of an individual is matched against a database containing a significant number (millions) of biometric templates. This requires both significant processing power and high-speed network connectivity to ensure high quality and speed of biometric recognition and identification, making such scenarios challenging especially in places where facial recognition and identification is carried out using mobile devices. As mentioned above, solutions for optimisation of neural network performance are being developed, allowing for the application of AI systems based on neural networks to be deployed on mobile devices, such as smartphones or portable cameras. Improving the speed of automated identification based on FRT while reducing the demands on computational and network connectivity is also possible by creating galleries of facial images. Deployment of such approaches is possible in those cases where passenger traffic is highly predictable, as in airports and ferry ports, for example.

In public sector environments, decisions made on the basis of automated data processing may have direct implications for the ability of an individual to exercise his or her fundamental rights. Therefore, the main challenge for the implementation of FRT is to ensure that the systems are highly precise and unbiased. The implementation of FRT at border crossings may significantly improve the performance of border checks in terms of both effectiveness and efficiency. However, given the quality of the FRT systems available on the market today, implementation of FRT for identification is a difficult balancing act between ensuring the safety and security of the general public and protection of privacy and non-discrimination of an individual. Although

the quality of facial recognition systems has improved markedly, FRT systems are still prone to errors, especially when processing images of certain minority groups[36]. The National Institutes of Standards and Technology (NIST) in the US have been performing face recognition vendor tests (FRVT) to identify best performing FRT systems, and to identify the specific deficiencies in the performance of these systems, such as dependence on high image quality or weak performance with certain demographic groups (e.g. racial minorities).

A number of studies analysing the demographic effects of FRTs asserted that that these technologies, as with other AI-based solutions, can be biased; in particular, their performance can depend on each demographic group[37].One of the focal points of the most recent FRVTs performed by NIST was therefore the effect of demographic characteristics on the performance of FRT[38]. The overall conclusion of the testing exercise performed by NIST was that all the facial recognition systems tested showed some demographic differentials; however, there was a very significant variability between the systems, with the generally more precise algorithms showing lower demographic differentials. As expected, the performance of facial recognition algorithms on a specific demographic depends on the data set used for training of the specific facial recognition algorithm. Furthermore, performance of algorithms is also affected by gender and age, with higher error rates for women, the elderly and children.

The challenge outlined above can be addressed by using representative data sets for training of facial recognition algorithms. The training data sets should represent all relevant features of the population of subjects that will potentially be processed by the system. There are two approaches to providing representative data sets for training:

- Creating **synthetic** data sets with the necessary characteristics representative of the population;
- Using **real** data sets in full or as a representative sample.

There are advantages and disadvantages to both approaches. Whenever personal data is concerned, use of real data sets for algorithm training can be impossible due to the limitations imposed by the regulatory framework on data protection. Studies have shown that machine learning models, such as deep neural networks, can be subject to generative model-inversion attacks with the intent to infer information about the training data set[39]. Synthetic data sets solve this issue; however, numerous studies have shown that algorithms trained on synthetic data sets work well with synthetic data but have significantly higher error rates when tested with real data. There is therefore a trade-off between the need to protect privacy and personal data and the need to provide highly precise systems. Whenever AI-based systems may affect the fundamental rights of individuals subject to the system, training of ML systems using real data sets should be preferred. However, the use of real data sets containing personal data for training of ML systems must comply with the relevant regulatory requirements and ensure high levels of security for the data used.

Beyond identification and authentication, FRTs can be used in other scenarios, such as attempts to capture human emotions to either detect deception at border crossing points, as claimed by the iBorderCtrl project[40],

[36] European Union Agency for Fundamental Rights (2019) Facial recognition technology: fundamental rights considerations in law enforcement. Available online: https://fra.europa.eu/node/37737

[37] Buolamwini J.(2017) Gender shades: Intersectional phenotypic and demographic evaluation of face datasets and gender classifiers. Technical report, MIT Media Lab; Cook, S., J. Howard, Y. Sirotin, J. Tipton, A. Vemury. (2019) Demographic effects in facial recognition and their dependence on image acquisition: An evaluation of eleven commercial systems. IEEE Transactions on Biometrics, Behavior, and Identity Science 1(1); Garvie, C., A. Bedoya, and J. Frankle. (2018)The perpetual line-up: Unregulated police face recognition in America. Technical report, Georgetown University Law School, Washington, DC; Grother, P., M. Ngan, K. Hanaoka. (2019) Face recognition vendor test (frvt) part 1: Verification. Interagency Report DRAFT, National Institute of Standards and Technology. https://nist.gov/programs-projects/frvt-11-verification; Krishnapriya, K. S., K. Vangara, M. C. King, V. Albiero, K. Bowyer. (2019) Characterizing the variability in face recognition accuracy relative to race. CoRR, abs/1904.07325. http://arxiv.org/abs/1904.07325.

[38] Grother et al. (2019) Face Recognition Vendor Test (FRVT). Part 3: Demographic effects. Available online: https://doi.org/10.6028/NIST.IR.8280

[39] Zhang et al. (2020) The Secret Revealer: Generative Model-Inversion Attacks Against Deep Neural Networks. Available online: https://arxiv.org/abs/1911.07135

[40] EU-funded Horizon 2020 Project iBorderCtrl claims to have devised a system for automated border security, which, in addition to automated processing of travellers, incorporates an automated real-time lie detector, relying on FRTs for identification of small eye and facial muscle movements.

or analyse human behaviour 'in the wild', as claimed by the SEWA project[41] [42]. These efforts, while somewhat interesting from an engineering or scientific point of view, have caused some concerns in civil society among groups concerned with social justice in the increasingly digital world. In the case of iBorderCtrl, the main point of contention is the pseudo-scientific basis of lie detection, in particular when it comes to inferring human emotions from facial movements[43]. There is, however, a broader issue with the use of such technologies in the public sector, namely their compatibility with the key principles for the development of AI in the EU put forward by the independent HLEG on AI: human-centric approach, ethics by design, respect for human autonomy, prevention of harm, fairness and explicability[44].

Furthermore, recent research shows that computer vision systems, in particular those based on convolutional neural networks, can be relatively easily misled by implanting certain information into the processed images[45]. Given that such systems based on neural networks operate as black-box systems, and understanding their behaviour requires high level of expertise in the AI auditing capability, their use at scale can lead to adverse outcomes.

The limitations outlined above need to be addressed before FRTs can be applied by public authorities at scale, especially in uncontrolled environments where performance of these systems tends to be further affected by the environmental conditions (i.e. weather, light, movement, etc.), to make sure that people's fundamental rights are respected.

3.2. Natural language processing and generation technologies

Along with computer vision, language technologies such as language comprehension, speech recognition and natural language generation have seen significant development over the last decade. One of the notable examples where AI has been successfully applied is machine translation (e.g. eTranslation[46] developed for public sector institutions in the EU). Natural language processing (NLP), which is essential for speech recognition and language comprehension, relies on machine learning algorithms for interpreting language captured as audio or text. Natural language generation (NLG), in turn, relies on machine learning algorithms trained on vast amounts of text in order to produce text or speech in response to a request from a human. These technologies are essential in a number of domains, from analysis of unstructured text-based data, to development of virtual assistants, such as chat-bots. In fact, the application of NLP and NLG technologies is perhaps one of the low hanging fruits with relatively low barriers for implementation and relatively high potential impact on the performance of service-oriented organisations with frequent customer interactions. When applied strategically, AI systems focused on natural language processing or generation can significantly improve service delivery including through quick automated responses to simple queries, sorting and classifying information for more effective processing by humans or augmenting human experts in addressing more complex requests by providing relevant information from the existing knowledge base when needed.

Natural language processing and generation are directly relevant to the core business of eu-LISA. First and foremost, NLP and NLG can be deployed in combination with a triaging algorithm for IT service desk automation, responding to requests for which responses are available in the knowledge base and classifying unanswered requests for more efficient response by the service desk team. Natural language processing and generation can also be effectively used in virtual assistants for new systems to be developed and operated by

[41] EU-funded Horizon 2020 Project SEWA aims to create a 'robust technology for machine analysis of facial, vocal and verbal behaviour'.
[42] https://www.euractiv.com/section/digital/opinion/the-eu-is-funding-dystopian-artificial-intelligence-projects/
[43] Feldman Barrett et al. (2019) Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements. Available online: https://journals.sagepub.com/eprint/SAUES8UM6gEN8TSMUGFg/full
[44] https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top
[45] See for example http://www.evolvingai.org/fooling
[46] For more information on eTranslation please visit: https://ec.europa.eu/info/resources-partners/machine-translation-public-administrations-etranslation_en

eu-LISA, helping new users to familiarise themselves with the new systems, for example. Other possible use cases for natural language processing include information security (analysis of online sources to identify new exploits or other threats) or human resources (automated pre-screening of job applications).

Conversational agents, or chatbots, as they are more widely known, is one of the more discussed forms of NLP and NLG technologies with potentially great benefits in the public sector. With the advancement of these technologies, the public sector has started testing chatbots with the aim of improving service delivery and customer satisfaction, reduce the workload of staff and improve the efficiency of service delivery. The most common version of the chatbot can be found on many websites today. This is the most simplistic implementation that relies on text-based interaction and uses simple statistical techniques to respond to questions using a database with answers linked to standard requests in the background. This type of conversational agent is only capable of answering the most common requests. Specialised or closed-domain chatbots respond to keywords in a narrowly defined area of expertise but cannot engage in an open conversation. Substantial research has been done recently on open-domain conversational agents, introducing different approaches in combination with neural networks in order to imitate human-like conversation, such as knowledge-based, retrieval-based and rule-based systems. Although these approaches have produced comparatively good results[47], their performance in multi-turn open-domain conversations is still far from perfect, as often the responses produced by chatbots may be meaningless. In a recent study Google conducted an experiment evaluating the performance of an open-domain conversational agent based on a 2.6B parameter end-to-end neural network. In the paper, the authors report on an experiment in which they tested the conversational agent on multi-turn static (data set) and interactive (human) conversations and came to the conclusion that its performance was significantly higher than that of other conversational agents. Furthermore, they concluded that it is possible to rely only on neural networks and still achieve good results, without the need to create complex handcrafted frameworks[48]. In addition to the single-domain chatbots, which are already relatively widespread, there are more sophisticated implementations designed to cover a wider range of issues through one communication channel, which are based on a network of intelligent virtual assistants[49].

Before considering the implementation of systems based on NLP or NLG, certain limitations need to be taken into account. First, high quality chatbot or virtual assistant technologies currently available in the market are all single domain. Although the underlying ML models are transferable between domains, those need to be retrained on domain-specific data sets. Second, training NLP and NLG systems requires large amounts of high quality training data covering the relevant domain[50]. Therefore, in order to develop an effective and precise system, the organisation needs to invest significant resources in order to provide high quality training data. NLP systems can still be used where large amounts of training data are not available, for example to sort and classify text-based information (e.g. incident reports). When used in combination with rules-based algorithms, NLP systems can also be used for triaging requests.

The third important limitation to take into account is language. Certain languages, such as English, have a significant corpus of textual information that can be used to train neural networks, whereas other minor languages, such as Estonian, have a far more limited corpus of data that can be used for training, which may negatively affect the quality of the algorithm. Therefore, although multilingual systems would be most relevant in the EU, the potential qualitative differences across languages mean that preference should be given to the

[47] See for example MILABOT (https://arxiv.org/abs/1709.02349); Microsoft XiaoIce (https://arxiv.org/abs/1812.08989); or Cleverbot (www.cleverbot.com).
[48] Adiwardana et al. (2020) Towards a Human-like Open-Domain Chatbot. Available online: https://arxiv.org/pdf/2001.09977.pdf
[49] See for example, the chatbot system implemented in Finland to consult entrepreneurs starting a new business in Finland: http://startingupsmoothly.fi/
[50] If the system will operate as an automated assistant at an IT help desk, it needs to be trained on customer requests addressed to the IT helpdesk and responses from the IT helpdesk. To make the algorithm even more accurate, it would need to be trained on the data originating from the same organisation, because IT help desks in different organisations deal with varying systems and requests which may mean that concepts can have different meanings.

most widely used languages in which the systems perform most accurately (e.g. English). Additional languages can be added when the system performance in the specific language improves.

There is a wide range of commercially available tools for automated translation, including natural language recognition, natural language generation, machine translation, and so on. In addition to the commercially-available technologies, CEF Digital offers an automated translation tool eTranslation[51] enabling delivery of multilingual online services for public administrations. The eTranslation tool allows translation of text-based documents and plain text in more than 24 languages. One of the advantages of using the tool is that it is being constantly developed to improve translation quality, widening the coverage for language pairs and domains. This specifically addresses the challenge of more minor languages which have been at a disadvantage when it comes to the development of automated translation tools.

Despite these caveats, natural language technologies in combination with other AI tools are valuable tools for any service organisation that has frequent interactions with service users. In addition, where chatbots or virtual assistants do not process personal data, they have few legal limitations and can be implemented within a reasonably short time. The widespread availability of natural language technologies, as well as their relatively low implementation cost (assuming that training data is available or not too costly to produce), suggest that these technologies can be prioritised as one of the first steps in automation. For eu-LISA these technologies are particularly relevant for IT service desk automation and for providing information and training for the users of the new systems it develops, in particular ETIAS.

3.3. AI for decision making and decision support

Like any other public sector organisation, eu-LISA faces myriad challenges related to operation in a complex environment. First, users of the systems developed and operated by eu-LISA have increasingly high expectations of the performance and availability of the systems. Second, eu-LISA, along with all other public sector organisations, is operating under tight budgetary constraints. Third, stemming from the tight budgetary constraint and the overall competition for highly capable professionals, eu-LISA operates in conditions of tight staffing constraints. Fourth, eu-LISA operates systems requiring high levels of security at a time when adversaries are using increasingly sophisticated strategies and methods for targeting security vulnerabilities. Last is the fact that eu-LISA operates in a very data-intensive environment. The availability of sophisticated analytical tools is therefore paramount to the efficient and effective operation of the organisation.

Decision making and decision support systems based on AI are therefore especially relevant for eu-LISA operations. Although the discussion around artificial intelligence has in recent years focused largely on computer vision applications, the application of artificial intelligence in decision support predates computer vision by several decades[52]. Early decision making and decision support systems were based on symbolic artificial intelligence and called expert systems. Expert systems encode expert knowledge in algorithms which then can be deployed in decision making or decision support systems, and have been applied in a very wide range of domains from system diagnostics, financial decision making, process monitoring and control to medical diagnosis. These systems have two main advantages. First, they are very efficient and do not require significant computing power or energy to run. Second, they are relatively transparent and can thus be subjected to human audit. However, this same advantage is simultaneously a disadvantage in that expert systems can only be applied in clearly defined scenarios with a fairly limited amount of independent variables affecting the outcome, since each step in the system needs to be programmed manually. As a result, any change in the external environment that affects the system in question must be programmed manually. Also, the more complex such algorithms are the more they are prone to failure and difficult to audit. Despite some of these

[51] https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eTranslation
[52] See e.g. Turban, E. (1988) Decision support and expert systems: Managerial perspectives. New York: Macmillan.

disadvantages, expert systems are still widely applied in a variety of scenarios and can be successfully applied in the context of operational management of large-scale IT systems, such as those managed by eu-LISA.

Today, decision making and decision support systems often rely on machine learning, including deep neural networks. Given the ability of machine learning algorithms to process large amounts of information in a relatively short amount of time, such systems are already being widely applied in healthcare to support doctors in diagnostic work across a wide spectrum of medical specialisations. Some studies suggest that the performance of automated diagnostic systems used in healthcare is already on a par with, or even better than, that of trained radiologists[53]. In other domains, however, automated systems for decision support and decision making have been deployed with varying results. Thus, for example, automated decision making systems used in the judicial system in the US have led to a system that is less fair towards defendants[54]. In a similar fashion, there is ample evidence that credit institutions using AI for risk assessment often inadvertently discriminate against certain groups[55]. The main source of bias in ML systems is the training data sets used to develop these systems. As described above, in order to reduce the error rate of AI systems, training data sets must be representative of the populations who will be subject to automated processing by the systems. It is therefore important to ensure that the systems are free of any potential bias if they are to be applied for automated decision making or decision support in the public sector. Since the exact origins of bias in the machine learning model is often not known and difficult to identify, some scholars suggest that whenever AI is used to make high stakes decisions affecting human beings, interpretable models should be used, in place of black-box neural networks[56].

There are, however, many opportunities for the implementation of decision making and decision support systems beyond those which have direct implications for citizens that still can deliver tangible benefits to the organisation. One domain where predictive analytics facilitates better decision making is cyber security. ML in combination with text mining can be used to identify potential vulnerabilities in the systems by scanning internet resources for relevant information and providing support to cyber security experts. ML can also be used to monitor system performance to identify anomalies resulting from cyber attacks[57] as well as for data centre and network automation[58]. These latter use cases, which are directly relevant to the core business of eu-LISA, will be examined further below.

3.4. Datacentre infrastructure for machine learning

As discussed above, different types of machine learning algorithms can be deployed depending on the specific use case. The common feature shared by most ML approaches is that they often require advanced high-performance data centre infrastructures for both training and deployment. Training machine learning algorithms requires very large data sets, hence also significant data storage capacity. Although personal data protection should be a default requirement independent of the sector, it is particularly relevant in the public sector, especially in the context of sensitive personal data processed by JHA organisations. Therefore, data used for training machine learning algorithms must be processed in highly secure data centre infrastructures[59]. The costs of such infrastructure and the skills necessary for maintaining it are often prohibitively high for a single organisation to bear. In the EU, the costs can even be prohibitively high for a single Member State, especially

[53] https://blogs.plos.org/speakingofmedicine/2018/11/28/better-medicine-through-machine-learning-whats-real-and-whats-artificial/
[54] https://www.nytimes.com/2017/06/13/opinion/how-computers-are-harming-criminal-justice.html
[55] https://www.brookings.edu/research/credit-denial-in-the-age-of-ai/
[56] https://www.nature.com/articles/s42256-019-0048-x
[57] https://emerj.com/ai-sector-overviews/artificial-intelligence-and-security-applications/
[58] Rafique and Velasco (2018) Machine Learning for Network Automation: Overview, Architecture, and Applications. Available online: https://doi.org/10.1364/JOCN.10.00D126
[59] We are assuming here that training machine learning algorithms on production data is legally possible, which is not currently the case.

smaller Member States. In addition, substantial efficiencies can be gained by deploying such data centre infrastructures on a very large scale[60].

There is a wide array of options available for private sector organisations to utilise on-demand infrastructure for AI training. A number of industry leaders, such as Amazon Web Services and Google Cloud AI, provide a range of services for machine learning depending on the needs of a specific company, ranging from pre-trained machine learning algorithms for specific applications (e.g. recommendation systems), to a full set of services necessary for training and deployment of custom machine learning algorithms, including data storage, computational infrastructure, analytics and security. Given the data protection requirements, certain constraints need to be considered when deploying public cloud-based solutions in the public sector. Hence, considering the ambitions expressed by the Commission to support the broad roll-out of AI in the private and public sectors[61], it is important to consider the need to develop dedicated AI infrastructure for the JHA domain where data protection requirements are often stricter than in other areas. There are three possible approaches. The first is to rely on the infrastructure available within the individual Member States and use additional solutions to make this infrastructure available across the EU, ensuring that it is utilised as efficiently as possible[62]. The second is to rely on existing centralised high-performance computing infrastructure, such as the one developed by the EuroHPC Joint Undertaking[63]. The third approach would be to develop centralised AI infrastructure and AI solution testing capabilities for the JHA domain, which would include the necessary computational and data storage infrastructure, and data science and AI expertise to develop, test and deploy relevant AI solutions. eu-LISA could contribute its expertise in the operational management of large-scale IT systems in the development and maintenance of such infrastructure.

The AI4EU project was launched in early 2019 with the support of the H2020 programme. Its aim is to build the first **European platform for AI on-demand**. The project will create an ecosystem of stakeholders involved in AI across all EU MS. The project will also design **a platform for on-demand AI, including high-performance computing resources, AI components and data sets.** Furthermore, the project will test the platform in several industry-run pilots, enabling real-life applications. More here: https://www.ai4eu.eu/about-project

[60] See, for example, the case of Facebook described in detail here: https://research.fb.com/wp-content/uploads/2017/12/hpca-2018-facebook.pdf
[61] COM(2018) 795 final
[62] For example by using virtualisation software for AI infrastructure, such as that developed by run.ai
[63] EuroHPC Joint Undertaking is an initiative between the EU and European countries to develop advanced supercomputing capabilities with a total of 1 billion euros of funding (https://eurohpc-ju.europa.eu/).

4. Use cases for the application of AI in eu-LISA

4.1. AI-enabled IT service management and IT operations

Although improving the efficiency of IT service management operations using automation of machine intelligence seems like a low hanging fruit, this area has been somewhat lagging in terms of deployment of AI solutions[64]. However, IT operations in both private and public sectors face a number of challenges, which already in the near to mid-term is likely to stimulate more interest in AI-based solutions.

- **First,** user expectations towards service providers continue to grow, resulting in increasingly tight SLA requirements. Users today expect both very high levels of system availability and an increasingly short time for incident resolution. At the same time, public sector organisations are under constant pressure to improve effectiveness and efficiency, which often results in severe budget constraints.
- **Second,** overall complexity and interdependencies between different integrated components and services which make incident resolution increasingly time-consuming.
- **Third,** the manual systems for IT service management (in particular for IT service desks) currently in place are often insufficient for today's fast-paced, dynamic and complex IT service operations. Responding to thousands of incident alerts – a significant number of which are inevitably false alarms – places IT service desks under unnecessary pressure and may result in 'alert fatigue'. Freeing up part of the service desk capacity by partially automating incident management will allow these resources to be dedicated to other tasks that require increased human judgment, such as problem management, for example.[65]

The challenges outlined above, coupled with budgetary constraints, put organisations under increasing pressure to improve cost-effectiveness. Some of these challenges can be addressed by introducing automation into IT operations.

When it comes to AI in support of IT infrastructure and service management, its application can be roughly divided into three categories.

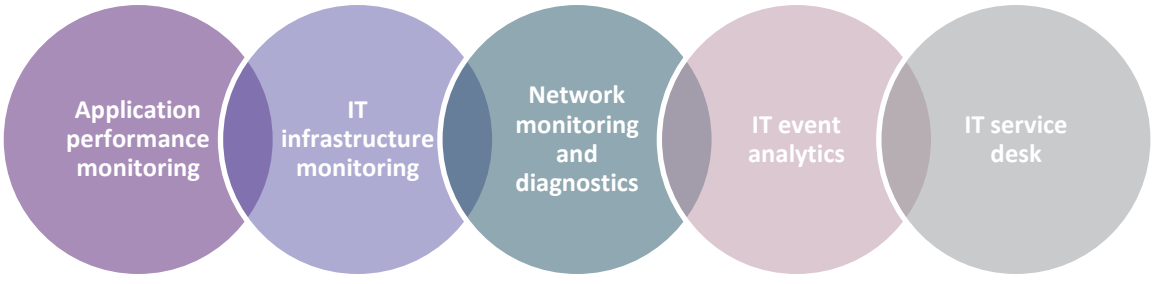

Figure 4: Opportunities for AI in IT service management

- **Anomaly detection systems** using ML algorithms can be applied to application performance monitoring as well as IT infrastructure and network monitoring. Tools using ML algorithms for pattern

[64] https://www.bcg.com/publications/2019/artificial-intelligence-coming-information-technology-operations.aspx
[65] (ibid.)

detection can help identify spikes in the use of applications or abnormal traffic in networks. Anomaly detection tools have already been tested in the context of data centre operations and proven to be successful[66].

- Another implementation with a potentially significant effect on IT operations, and in particular in the work of the IT helpdesk is **noise reduction.** Machine learning algorithms could be deployed to analyse incidents and alerts and sort those depending on their business impact, at the same time filtering out false positives. This could significantly decrease the workload of IT operations and service desk staff, allowing them to focus on resolving major incidents or manage problems.
- Machine learning algorithms could also be used for **triaging and alert correlation.** Transversal alerts affecting various IT services could be merged into one incident, speeding up triage. By analysing historical data on alerts, including their topology, time and context, the solution could identify relationships between incidents and unify those into unique instances whenever feasible. This information can then be used for **fault detection and prediction.**

Looking specifically at IT service management, Kubiak and Rass[67] provide a comprehensive overview of data-driven approaches across the ITSM processes, and identify the following areas for the implementation of data-driven approaches: classification of IT service tickets; online failure prediction; IT infrastructure event analysis.

Although **classification of IT service tickets** may appear rather mundane, a number of challenges are associated with this activity. First, in organisations maintaining large-scale IT infrastructures, the number and diversity of tickets is so large as to make manual classification inefficient or even unfeasible. Second, tickets often contain a mixture of human and machine-generated text. Third, machine-generated text may contain vocabulary specific to the monitoring agent. Fourth, text contained in user-generated requests may contain spelling mistakes. Last, sorting and allocation of tickets to a specific processing team is a non-trivial task, requiring in-depth knowledge of the specific competencies and existing IT infrastructure. These challenges in automated ticket processing can be solved by incorporating domain knowledge into the hierarchical multi-label classification system[68]. Besides automated routing, ticket classification systems can also be used to generate solution recommendations based on historical ticket data. Automated generation of resolution recommendations can substantially shorten the time from issue detection and resolution, thus leading to higher levels of availability and also better service for the end user.

Most benefits can be reaped by combining automated ticket processing systems with conversational agents. This enables a significant proportion of simple, frequently occurring requests to IT service desks to be addressed without the need to engage service desk staff, thus reducing routine work and allowing service desk staff to focus on more complex requests. Aside from increasing the efficiency of the service desk, this may also improve the overall quality of service.

4.2. Conversational agents and virtual assistants

As mentioned earlier, conversational agents, virtual assistants and chatbots are particularly relevant to eu-LISA, especially for the introduction of new systems, but also for managing existing systems. Conversational agents can be used to train new users of the systems, in particular the EES and ETIAS. Passenger carriers will need to check the EES to determine if a third-country national travelling to the EU with a short-stay single or multiple-

[66] Li et al. (2017) Anomaly detection for virtualized data center via outlier analysis. Available online: https://ieeexplore.ieee.org/document/8000085
[67] Kubiak and Rass (2018) An Overview of Data-Driven Techniques for IT-Service-Management. Available online: https://ieeexplore.ieee.org/document/8491279
[68] Zeng et al. (2017) Knowledge Guided Hierarchical Multi-Label Classification Over Ticket Data. Available online: https://ieeexplore.ieee.org/abstract/document/7852503

entry visa has used up their number of authorised entries, and carriers will need to check ETIAS to ensure that visa-exempt passengers using their services possess a valid travel authorisation. Ensuring fast adoption and correct use of the systems by the passenger carriers will require training of staff. One possibility is to provide such training through the integration of a chatbot-facilitated training module, providing step-by-step guidance and responding to possible questions. Similar chatbot functionality can be implemented in the ETIAS web-based system (or app), providing support to end users of the systems applying for travel authorisation. As indicated in the relevant regulatory acts, ETIAS will be available in multiple languages, therefore, support functionality will also need to be available in multiple languages. This can be ensured by integrating automated translation tools or multi-lingual chatbots.

Conversational agents are also relevant in IT service management. In addition to facilitating ticket management and response within service desks, conversational agents and virtual assistants could be used for initial customer contact. In the context of an IT service desk, it is reasonable to assume that the best quality can be achieved by using a closed-domain conversational agent, since IT service desks face a large number of recurrent requests which can be resolved with reasonably standardised solutions. These agents can be deployed in both next and speech modes, thus covering both requests received in writing and via telephone. The conversational agents can be used as the first contact layer with customers, responding to the most common standard requests, for which solutions are readily available, and routing more complex request to the relevant teams. This will at least partially reduce the workload of service desk staff and allow them to focus on more complex and creative tasks. The speed and quality of IT service desks and also the quality of the services provided can thereby be improved.

Conversational AI systems are widely available on the market as both open source code frameworks and proprietary technologies managed by vendors. The performance of these systems will vary depending on the approach used. Hence, intensive testing including static and interactive domain-specific conversations needs to be performed to ensure that the system

Some lessons from chatbot implementation (Lacity et al, 2017)

Although there are certain differences between private and public sector organisations when it comes to implementing new technologies such as chatbots, some of the lessons learnt from their implementation are worth considering.

1. **Treat the first chatbot implementation as a learning project.** In most organisations, IT projects are implemented either as business projects or innovation/R&D projects, depending on the readiness of technology, and with project ownership by business operations or innovation/R&D. Chatbot implementation is a borderline case, as there is a clear business case, but the technology requires substantial effort in terms of testing and implementation. This allows for a more flexible approach and space for exploration and error.
2. **Define criteria for use cases with potentially high impact.** Conversational agents are suitable for services with a large volume of routine tasks relying on vast amounts of unstructured data in a relatively narrow domain of expertise (for now). It is important to focus on specific pain points for staff and customers, especially the frequently recurrent issues that can be resolved end-to-end.
3. **Decide on a reasonable threshold from which to engage human agents.**
4. **Decide on the approach and technology and a corresponding approach to training.** Depending on whether to engage a closed or open-domain chatbot, and depending on the machine learning approach applied (supervised, unsupervised, etc.), an appropriate team should be assembled and a training framework and data sets defined.
5. **Gain management support without overselling.** Outline clear benefits to the organisation and relevant stakeholders, ensuring support and continuous funding for the initiative.
6. **Engage clients (first internal, then external) to try the chatbot, but leave other channels open.** It is important to promote the new approach to ensure user engagement, gradually making the chatbot a first point of contact. However, it is equally important to keep other channels open to reduce frustration.

meets user expectations[69]. Furthermore, certain thresholds must be defined, beyond which human agents enter the scene in order to ensure a high level of service. Other strategies, such as intelligent switching of domain-specific conversational agents, can be deployed to ensure high quality and user satisfaction.

As with any new technology, its implementation should be done preferably in an iterative way in a number of cycles. First, conversational agents can be implemented as a proof of concept (PoC). At the PoC stage, a basic system prototype is developed which includes the analysis of system and data requirements, collection of data and preparation of training data, and development and testing in a contained environment. If the PoC is successful, the tool is developed into a minimum viable product and piloted in a real environment. At this stage, the product can be rolled out with a limited scope (e.g. internal users or a test group of users) and integrated with the databases in place, and its functionality extended, incorporating a broader scope of requests, for example. At this stage, the tool can also be tested further to include collection of user feedback through questionnaires. Following pilot implementation and taking into account the results of the pilot, the product can be scaled up, with further integration into the existing systems and overall IT service management system of the organisation.

There are a number of benefits to using conversational agents in the operation of large-scale IT systems. First and foremost, it increases the speed with which customer service requests are addressed, thus improving service quality. It also reduces the workload of the IT service desk and focuses their attention on more complex issues, thus improving availability and overall quality of services. In the case of ETIAS, virtual assistants can support applicants in submitting application, thereby improving user experience and ensuring data quality.

4.3. Automated fault detection and prediction as an integral part of the fault management system

Automated fault detection and prediction is the next logical step in ensuring high quality operational management of the IT systems entrusted to the Agency, which is now ensured through constant monitoring, incident management and interaction with the Member States. eu-LISA has been working on identifying and exploiting the synergies in infrastructure and services, in particular with the implementation of an integrated solution for monitoring tools for the existing systems. Integrated monitoring in combination with automation in fault detection, as well as predictive analytics, can significantly enhance the efficiency of the operational management of the core business systems.

In most IT and network environments fault detection is done in three ways. First, by relying on user reports. This is the simplest method that can only be used in non mission-critical environments, since users reporting a failure means that the requested service is not available. Second is the use of specialised testing suites, which check the availability and performance of systems or applications against the set parameters. Although this approach is highly accurate, it is very cost and resource-intensive, consuming infrastructure resources in order to run, which means that the test cannot be run too frequently. Third, by using relatively simple rules-based systems that trigger an alarm on the basis of a few metrics being monitored. Rules-based methods are reasonably easy to implement and do not impose a significant load on infrastructure. However, both the second and third methods require significant human effort in manual configuration taking into account metrics from the full infrastructure stack. Considering the frequent changes in software, rules-based systems need to be frequently updated, requiring expertise on all system levels and a team effort[70].

Automated fault detection and prediction systems based on machine learning address at least some of the

[69] Methods for chatbot evaluation similar to those proposed in the following paper: https://arxiv.org/pdf/2001.09977.pdf
[70] https://www.ericsson.com/en/blog/2019/3/automating-fault-detection-for-management-systems-using-ml

issues of currently deployed fault detection described above. However, they do not come without a cost. Successful deployment of fault detection and prediction tools based on machine learning requires large amounts of data for training ML models. Therefore, the very first step in automated fault detection and prediction is data collection. Although historical data already exists in most organisations, the metrics that are stored do not include all possible metrics and are selected by human experts. In the case of machine learning systems, the best approach is to collect as many metrics on the performance of the systems as frequently as possible. Different solutions for the storage of multi-dimensional time-series data that can be used for collecting monitoring data are available on the market[71], deployed with additional components, such as Node exporter[72] to allow for exporting hardware and software performance metrics.

The second step is data pre-processing, the nature of which depends on the data collected, as well as on the requirements of machine models to be trained. Data pre-processing includes the following steps:

- **Data synchronisation:** time-alignment of data collected from different agents;
- **Data cleansing:** removing data that cannot be used in time series (e.g. non-numeric) and generating missing data whenever necessary through interpolation or other means;
- **Normalisation:** re-scaling metrics so that they are in comparable magnitude (e.g. via min-max normalisation or standardisation);
- **Feature selection:** identifying relevant metrics for ML model training.

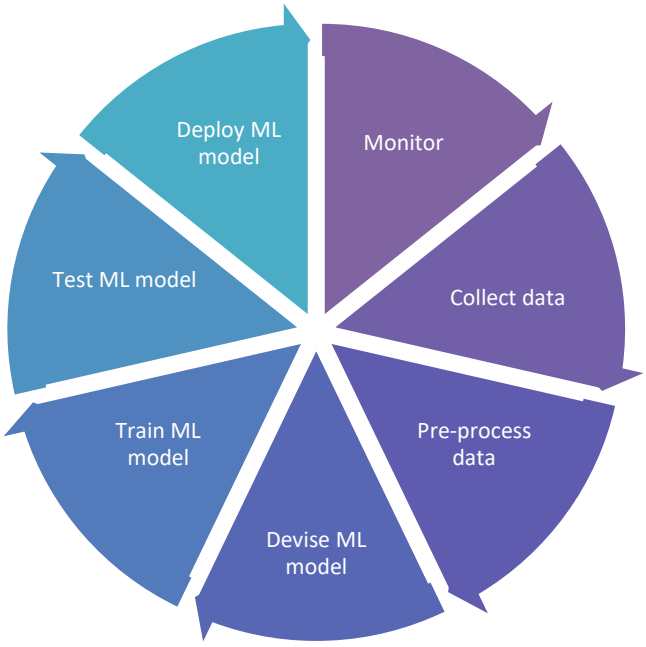

Figure 5: Deployment of machine learning in fault detection (Source: author's creation based on https://www.ericsson.com/en/blog/2019/6/automated-fault-management-machine-learning)

[71] For example Prometheus of Influxdb.
[72] https://github.com/prometheus/node_exporter

In particular during the initial deployment, this process can be rather resource-intensive; however, the resources required for data pre-processing decline after the initial set-up. Nevertheless, there will be a constant need for data management due to the development of infrastructure, new types of data and the resulting need to update or retrain the existing ML models. In the case of a stable system, model training takes place only once, whereas in systems with changing components or characteristics, models need to be retrained in order to ensure adequate performance. Novel approaches for continual learning are being developed, which will resolve the issue of model retraining, but will not be discussed here[73].

As is the case with the implementation of any AI technologies in environments characterised by required high levels of robustness and availability, implementation of automation in fault detection should be done iteratively, ensuring that the devised system performs in a highly precise manner within a specific setting. Therefore, prior to deployment in a production environment, the system should be validated either using real-time data or on a separate data set which was not part of the training data set. When the model has been validated, it can be deployed in the production environment.

Two broadly defined frameworks can be used in automated fault detection – supervised and unsupervised learning – for somewhat different objectives. As discussed above, supervised learning uses labelled training data sets; in the case of fault detection systems, the training data set would include data points containing metrics on the status of the system and the corresponding label (fault/no fault). A range of supervised learning methods including neural networks, support vector machines (SVM) and random forests (RF) can be used for fault identification, although due to their relative simplicity, SVM and RF are relatively more efficient and faster to train than neural networks. However, neural networks are more effective with data sets with a very large number of features but require more computational resources. Unsupervised learning, where models are trained with the same data samples but larger data sets and without labels, can be effectively used for anomaly detection in the performance of the system. Anomaly detection is normally based on pattern recognition techniques.

Automatic fault detection systems based on machine learning algorithms help identify faults earlier than is possible with conventional tools. However, systems for automated fault detection are not optimal for mission-critical components or systems, where a fault can mean unavailability of service. In such cases, automated fault prediction is essential for the successful performance of systems. However, for a number of reasons, fault prediction is a more difficult problem than fault detection. The first problem is having sufficiently large data sets for training. Since faults are exceptional states in the system, they occur rarely and an accumulation of a sufficiently long time series can therefore be a challenge. One workaround is to intentionally cause faults in the system in order to produce synthetic data. Faults also need to be predictable (i.e. they either result along with specific changes in the state of the system with the corresponding changes in the metrics) or they follow a certain pattern that can be identified by a machine learning algorithm. Different tools can be used for failure prediction depending on the availability and quality of data sets. In the case of smaller data sets, statistical methods, such as the autoregressive integrated moving average, can be applied for predicting values in the time series. Such approaches, however, have limited scalability and therefore cannot be successfully deployed in large systems with many parameters. In such cases, neural networks can be used, such as recurrent neural networks, long short-term memory or convolutional neural networks. These methods vary in how data and computation intensive they are[74].

[73] For an example of an approach to reinforced continual learning see http://papers.nips.cc/paper/7369-reinforced-continual-learning
[74] https://www.ericsson.com/en/blog/2019/6/automated-fault-management-machine-learning

4.4. AI for cyber security

The issue of protection of large-scale IT systems from modern security threats was addressed in detail in the dedicated report published by eu-LISA in 2016[75]. However, there have been some substantial developments in the field of cyber security since 2016, most importantly related to the development of cyberattack capabilities relying on artificial intelligence, which has necessitated the development of cyber defence technologies using machine learning[76]. Adversaries already deploy machine learning in a number of ways. Primarily, machine learning is used for automation. For example, automation of data analysis for identification of vulnerabilities substantially reduces the labour costs and time necessary to prepare an attack. Natural language processing and generation can help improve targeting, as well as automation of phishing attacks. Furthermore, machine learning algorithms may be used for probing cybersecurity systems used by organisations and adapting according to the target's potential response, thus avoiding potential detection[77]. Lastly, automated attacks using AI can be widely spread (by sharing the source code via the darknet, for example) and initiated by adversaries with much lower levels of expertise, and can therefore become more widespread.

Ensuring the security of the systems, as well as the safety of data stored within them, given the complex cyber threat landscape requires novel approaches. In this regard, novel cyber security capabilities utilising AI will become a necessity in an environment characterised by the ever growing sophistication of cyber attacks. AI, more broadly, and machine learning, more specifically, can be used in a number of applications for cyber security. At the most basic level, the protection of digital infrastructure relies on cyber threat intelligence in order to identify vulnerabilities and new attacks. The sheer amount of information and the pace of development of new attacks means it is physically impossible for organisations to manually keep track of possible vulnerabilities, new exploits or new attack strategies. Therefore, automating cyber threat intelligence using machine learning is an important first step. One of the approaches is based on natural language processing and identification algorithms in combination with classification algorithms. This approach was used for cyber threat discovery using data from the dark web hacker forum discussions[78], where the method proposed by the authors helped relevant and actionable evidence that could then be integrated with the standard security processes to be quickly extracted. Other authors used Twitter data to identify security-related information relevant to specific IT infrastructure assets. In this case, the researchers used a binary classifier based on a convolutional neural network for natural language processing, which was complemented by a named entity recognition model based on a bidirectional long short-term memory neural network. Based on the experiments using this approach, the authors were able to demonstrate that Twitter can be a source of valuable data for cybersecurity operations, thus improving the level of cyber threat awareness.[79]

The second level of application of AI in cyber security is in network analysis and intrusion detection. A survey article focusing on such applications defined three categories of attack detection methods using machine learning and deep learning: misuse-based, anomaly-based and hybrid. Misuse-based methods for attack detection aim to identify attack signatures on the basis of historical patterns. Such systems are relatively precise and generate low levels of false alerts, but they require frequent updates of reference databases of attack signatures, and are not able to detect zero-day attacks. Anomaly-based attacks monitor the performance of systems in normal conditions to identify anomalous behaviour. Such systems are able to detect zero-day attacks, but are prone to relatively high false alert rates. Hybrid systems are more advanced, as they rely on a

[75] eu-LISA (2016) Protecting large-scale IT systems developed and/or managed by eu-LISA from modern threats. A review of recent developments in IT security and associated technologies. Research and technology monitoring report. doi:10.2857/320307
[76] https://www.weforum.org/agenda/2019/04/how-ai-raises-the-threat-of-cyberattack-and-why-the-best-defence-is-more-ai-5eb78ba081/
[77] https://www.wilsoncenter.org/blog-post/the-real-challenges-artificial-intelligence-automating-cyber-attacks
[78] Deliu et al. (2018) Collecting Cyber Threat Intelligence from Hacker Forums via Two-Stage, Hybrid Process using Support Vector Machines and Latent Dirichlet Allocation. Available online: https://ieeexplore.ieee.org/abstract/document/8622469
[79] Dionisio et al. (2019) Cyberthreat Detection from Twitter using Deep Neural Networks. Available online: https://ieeexplore.ieee.org/abstract/document/8852475

combination of the two previous models and can therefore both identify novel attacks and perform with high levels of precision. One of the challenges of systems based on neural networks is a relatively high level of false alarms, which makes their application in operational environments challenging. The approaches based on decision trees appeared to be more precise[80].

The next level is the creation of autonomous cyber security systems capable of threat detection and response without the need for immediate human intervention. In order to stimulate the development of systems for automated vulnerability detection and patching, the US Defence Advanced Research Projects Agency (DARPA) organised a Cyber Grand Challenge in 2014-2016. The system that won the challenge, however (Mayhem by ForAllSecure) did not rely on artificial intelligence or machine learning, but instead relied on methods utilising symbolic execution of parts of target software in combination with a heuristic search strategy. Only one of the competing systems explicitly mentioned the application of machine learning. Nevertheless, the DARPA challenge proved not only that automated systems for discovery and patching of vulnerabilities were possible, but also that these systems were demonstrated in practice[81]. The follow-up programme launched by DARPA in 2018 and called Computers and Humans Exploring Software Security (CHESS) focuses on the development of capabilities for advanced discovery and patching of vulnerabilities and in particular supporting advanced human-computer interaction, which remains essential for cyber security[82].

In more recent years, network traffic analysis has become an important area for automated security. A number of vendors have recently entered the field of network traffic analysis, most of them using machine learning at the core of their technologies. It is worth noting, however, that at this point in time the core systems whose operation eu-LISA is responsible for run on a secure private network (TESTA-ng) and are therefore not susceptible to attacks common in systems connected to public internet. With the launch of the ETIAS system, however, this situation will change, requiring a more proactive approach to system security. Deploying AI-based cyber security technologies will become particularly relevant to ensuring robustness and availability of the systems.

4.5. AI-enabled services in the context of the new systems – ETIAS

ETIAS is one of the new systems whose development and operational management has been entrusted to eu-LISA. ETIAS is designed for the automated processing of most applications, with only those applications that receive a hit when checked against the relevant systems being processed manually. Therefore, an additional level of automation or analytics based on AI or machine learning could be introduced when dealing with any 'suspicious' applications. Such a system could support case officers responsible for evaluating applications with additional risk assessment based on the data stored in the relevant systems and the historical data on the individual submitting the application. The result of such a risk assessment could be a binary suggestion to either confirm or reject the application or a grading of risk, which would allow the case officer to then make a further evaluation of the application. Automation could also be introduced for triaging the applications, making sure that applications reach case officers with the necessary expertise or language skills, for example (in the case of issues with automated translation of documents). In the case of triaging, a relatively simple, rules-based system can be used, whereas risk assessment requires a more sophisticated system using one of the machine learning approaches (e.g. support vector machine algorithms, which have been successfully deployed in credit risk

[80] Xin et al. (2018) Machine Learning and Deep Learning Methods for Cybersecurity. IEEE Access, 6, pp. 35365-35381
[81] Loaiza et al. (2019) Utility of Artificial Intelligence and Machine Learning in Cybersecurity. Institute for Defence Analyses. Available online: https://www.ida.org/-/media/feature/publications/u/ut/utility-of-artificial-intelligence-and-machine-learning-in-cybersecurity/d-10694.ashx
[82] https://www.darpa.mil/program/computers-and-humans-exploring-software-security

analysis[83]). There are obvious ethical implications that need to be explored when applying AI in automated decision making as well as profiling individuals on the basis of their personal data, even if the data is available publicly. As discussed above, the principle of the 'human in the loop' should be applied and the ultimate decision should always be made by a human agent. This does not, however, preclude deploying AI – in its simpler and more sophisticated forms – in supporting the decision-making process by providing the necessary evidence.

4.6. AI for energy efficiency in data centres

Along with data analytics, prediction and virtual assistants, energy consumption is yet another challenge that can be addressed with machine learning. Data centres consume large amounts of energy. Data centre energy consumption consists of two major components: data centre IT infrastructure and the cooling systems that maintain the optimal temperature of the IT infrastructure. With the introduction of the new systems, eu-LISA will significantly extend its data centre infrastructure, so optimising the performance of data centre infrastructure will become increasingly important. This is particularly relevant considering the role public sector organisations play in reducing the environmental impact of their business, as well as leading the transformation towards more environmentally sustainable organisational models.

Companies and public sector organisations operating large-scale data centre infrastructures have thus far been using a range of methods to improve the energy performance of IT infrastructure, including techniques focusing on hot air containment, comprehensive monitoring and optimising temperature performance of infrastructure (i.e. operating IT infrastructure with tighter margins in terms of temperatures indicated by manufacturers as optimal)[84]. In order to improve data centre energy performance further, other methods are necessary. Machine learning algorithms applied in combination with the existing monitoring systems provide an opportunity to further significantly improve the efficiency of data centre operations.

Modern data centres operate in a very complex environment, with a large number of electrical and mechanical equipment operating in constant interaction (e.g. chillers, heat exchangers, cooling towers, water pumps and control systems), with each system potentially affecting another in non-linear ways that are not easily predictable with conventional methods. The growing overall complexity of data centres also makes it increasingly difficult to identify the most optimal settings for all components in order to reach maximum efficiency within the given performance requirements[85]. Deploying statistical methods and more complex machine learning frameworks, such as neural networks in developing energy efficiency models for data centres, has over recent years gained increasing attention in organisations operating large-scale IT infrastructure. For example, machine learning models piloted by Google in multiple facilities in 2016 reduced cooling energy consumption by 40% and overall energy overhead by 15%. These performance improvements were achieved on data centre infrastructure that was already efficient, consuming 50% less energy than the industry average[86][87]. Energy performance of data centres can also be improved by utilising machine learning models in data centre management automation with the aim of ensuring a balance between throughput, quality of service and power consumption[88].

Although invisible to the outside, the application of AI in improving data centre performance, in particular energy efficiency, is an important component in the overall improvement of the operational performance of an

[83] Harris (2013) Quantitative credit risk assessment using support vector machines: Broad versus Narrow default definitions. Available online: https://www.sciencedirect.com/science/article/pii/S0957417413000754
[84] https://static.googleusercontent.com/media/www.google.com/en//corporate/datacenter/dc-best-practices-google.pdf
[85] For example, a system which consists of 10 components, each of which with 10 different settings, will result in 10 billion possible configurations. Data centre systems consists of many more systems with many more possible parameter settings, resulting in an exponentially larger number of possible configurations.
[86] https://sustainability.google/projects/machine-learning/
[87] Gao, J. (2014) Machine Learning Applications for Data Centre Optimization. https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/42542.pdf

organisation, as well as an important step towards a more environmentally sustainable organisation.

5. Conclusions

Over the last decade we have observed a very rapid pace of advancement of AI in terms of both theoretical developments and applications – and it is unlikely to slow down any time soon. This rapid development as well as the pervasiveness of AI as a new technology also means that it has already profoundly affected organisations and societies, with implications for many of our spheres of life. Like many new technologies, regulation is lagging behind and therefore often cannot address many of the pressing concerns, including the ethical implications of AI, which are much cited in particular when AI is used for automated decision making. These concerns are valid and need to be appropriately addressed. However, they should be addressed in a way that does not hinder the development of AI capabilities in the EU. It is therefore paramount to adopt a strategic approach to regulating AI in the EU. This approach should allow for sufficient flexibility for the development of new technology and its application across the economy, while at the same time ensuring that whenever AI is deployed where it may have direct implications for human beings, it must be transparent, auditable and compliant with all applicable legal and ethical requirements.

For eu-LISA, as for any other organisation providing IT services, implementation of AI is not a question of 'if', but 'when' and 'to what extent'. The EES and ETIAS both foresee a certain level of artificial intelligence or automation and will therefore have an immediate effect on individuals. Considering that the precision of AI systems based on machine learning algorithms depends to a significant extent on data sets used for training, and the quality of data sets used for the training of biometric recognition systems to be used in the EES will to a significant extent determine the quality of the AI systems. With this in mind, a careful evaluation of the trade-off between privacy protection and system performance will need to be made, in particular focusing on the benefits and disadvantages of using real data sets for ML system training.

There are, however, many applications where AI can be deployed without any implications for human beings. This report has therefore largely focused on applications where ethical and legal considerations are either not relevant or are secondary; namely on those applications where AI can improve the performance of eu-LISA either by improving the quality of service (e.g. chatbots or virtual assistants) or the performance of the infrastructure (e.g. automated failure detection), or by reducing costs and improving performance by taking over some of the repetitive routine tasks (e.g. AI on IT service desks).

Although AI has great transformative potential, it does not come without challenges and risks. Therefore, organisations considering the deployment of AI in either internal operations or services should take the following factors in consideration:

- Development of machine-learning based systems will require large data sets in order to train the ML algorithm. In order to make sure that the AI system performs with a relatively low level of errors, the training data sets need to be of very high quality. Creation of high quality data sets will require significant initial investment.
- Development and deployment of AI based on machine learning (in particular deep learning neural networks) algorithms require substantial computational resources, which are normally not available within one organisation. Therefore, the development and deployment of AI external service providers, such as Amazon Web Services or Microsoft Azure, or shared infrastructure, such as the EuroHPC, will be necessary.
- AI systems – both rules-based and ML-based – are often not static, as they operate in a continuously changing environment. In order to take into account the changes in the environment, AI systems must be continuously updated to remain accurate. This will require either updating the code of the rules-based systems, or retraining the ML algorithms with new data sets or using the more recent continuous training approaches. This has direct implications for the procurement of such systems, as well as creating a

potential need for in-house staff to support the continuous maintenance of such systems.

- It is important to look beyond the headlines and consider where AI can have significant effects without substantial barriers to its implementation. Such 'low hanging fruits' can help legitimise AI within the organisation and gain management support. They are also a good way to acquaint staff with AI systems, which will help with gradually developing and deploying more sophisticated systems where necessary.

This report is the first step for eu-LISA in its exploration of the full potential of AI for enhancing the performance of the Agency. We will continue to monitor the developments in AI and look deeper into some of the use cases as part of our research and technology monitoring activities in the future.

Glossary

- AI – artificial intelligence
- CEF – Connecting Europe Facility
- DARPA – Defence Advanced Research Projects Agency
- ECRIS-TCN - European Criminal Records Information System – Third Country Nationals
- EES – Entry/Exit System
- ETIAS – European Travel Information and Authorisation System
- Eurodac – the system for the comparison of fingerprints for the effective application of the Dublin Convention
- EuroHPC - European High-Performance Computing Joint Undertaking
- FRT – facial recognition technologies
- FRVT – Facial Recognition Vendor Tests
- GDPR – The General Data Protection Regulation
- HLEG – High Level Expert Group
- ICT – information and communication technology
- IT – information technology
- ITSM – IT Service Management
- JHA – Justice and Home Affairs
- ML – Machine Learning
- NIST – National Institute of Standards and Technology
- NLP – natural language processing
- NLG – natural language generation
- PoC – proof of concept
- RF – random forest
- SIS II – second generation Schengen Information System
- SLA – service-level agreement
- SME – small and medium-sized enterprise
- SVM – support vector machine
- TESTA-ng - Trans European Services for Telematics between Administrations (next generation)
- VIS – Visa Information System

Annex I: Application of AI in the public sector

AI application	AI function and value proposition	Use cases in public sector environments
AI-based knowledge management (KM) software	• Generation and systematisation of knowledge-gathering, classifying, transforming, recording and sharing knowledge; • Expert systems supporting codification of knowledge in KM systems; • Neural networks for analysis of data and production and sharing of knowledge.	• Clinical documentation powered by AI in hospitals; • Case documentation for law enforcement; • Knowledge management, e.g. IT service desks.
AI process automation systems	• Automation of standardised tasks using either rules-based or machine learning based systems; • Automated workflow processing; IoT and intelligent sensor based technologies; case-based reasoning; • Robotic process automation, substituting humans in routine or dangerous processes.	• Automated clinical diagnostics (e.g. cancer diagnostics based on image recognition); • Automated data entry and classification; • Automated request processing and response.
Virtual agents	• AI-based agents for interaction with humans using NLP and generation via text or sound; • Closed-domain conversational agents as a first layer of contact; • Can be implemented with automated translation.	• Chatbots for IT service desk; • Systems for automated routing of requests with NLP capability; • Human-computer interaction in cases of repetitive tasks with defined outcomes (closed-domain systems with limited knowledge bases).
Predictive analytics	• Statistical analysis in cases of small data; • Big-data analytics using machine learning methods for decision making support; • Predictive analytics for automated decision making.	• Predictive analytics in policing (threat prediction and prevention); • Predictive modelling for weather emergencies and seismic activity; • Predictive modelling for infrastructure maintenance.
Identity analytics	• Using machine learning for advanced analytics of identity data in real time (including facial recognition for identification and authentication); • Risk-based identity checks using big data and machine learning.	• Facial recognition technology for crime prevention; • Facial recognition on borders for automated traveller processing; • Facial recognition for authentication in online service environments; • Fraud prevention in online service environments to secure government data.
Cognitive robotics & autonomous systems	• Systems with higher-level cognitive functions, involving knowledge representation and ability to learn and respond; • Sometimes in connection with affective computing to respond to human behaviour including emotions.	• Autonomous vehicles in public transit; • Autonomous vehicles in security and defence applications; • Robot-assisted surgery.
Recommendation systems	• Information classification systems supporting human decision making and action; • Prediction of user preferences based on historical data.	• Service personalisation in online service environments based on historical data on service consumption.
AI for cybersecurity and threat intelligence	• Data mining and big data analytics for threat landscape scanning, including NLP; • Cyber threat identification, analysis, and proactive response; • Vulnerability monitoring.	• Cyber threat analysis, vulnerability identification based on big data analytics and machine learning. • Automated response to potential cyber threats.

Table 1: Opportunities for application of AI in the public sector (adapted by the authors based on Wirtz et al., 2019)

EL-04-20-018-EN-N

ISBN 978-92-95217-67-6

DOI:10.2857/58386

Catalogue number: EL-04-20-018-EN-N