AI APPLICATIONS FOR SECURITY, PEACE AND JUSTICE

Effective, secure and human centric

NL AI Coalition

INHOUDSOPGAVE

1. INTRODUCTION

Our society would now be inconceivable without artificial intelligence (AI), as it is already widely used. In an age of digitalization with an increasing level of threats, utilising AI presents opportunities for improving both our physical and our digital security. AI can help protect the public, ensure justice and improve general wellbeing. At the same time, using AI also engenders risks and a balance needs to be struck between using AI applications effectively and protecting people's rights.

The Netherlands AI Coalition's working group on Security, Peace and Justice (hereinafter also "VVR" – *Veiligheid, Vrede en Recht* in Dutch) brings relevant national players together so that effective, safe and responsible AI applications can be developed and used to benefit security, peace and a correctly functioning justice system in which humans are the focal point. The working group aims to make the Netherlands one of the leaders in human centric AI for security, peace and justice. In this publication, the working group shows what is needed to achieve this goal, the aspects that need to be focused on, and the current situation. Various solutions, initiatives and applications are also explained.

1.1 THE WORKING GROUP SECURITY, PEACE AND JUSTICE

The working group focuses on the use of AI within the security domain, in the broadest sense of the term. This could, for example, include pre-empting criminal acts, forensic investigations, physical safety, crisis management, jurisprudence and public administration. The group aims for a human centric approach in which AI is used safely and responsibly from the moment of development. The working group's focus is on privacy-proof data sharing, decision support applications, cybersecurity and speech and language technology. The guiding principle is a human centric approach in which, accountability, transparency and inclusion are key. The working group is making efforts to make the Netherlands one of the leading lights in artificial intelligence for security, peace and justice at the strategic as well as the operational and tactical levels. To that end, the working group is aligned with the AI strategy of the Ministry of Justice and Security.

1.2 THE TRIPLE HELIX

The working group consists of 110 members representing what is referred to as the 'triple helix' for every region of the Netherlands: governmental and semi-governmental (both policymaking and executive organisations), centres of expertise (academic and applied research, educational institutions), the commercial sector (ranging from start-ups and scale-ups to large companies) and social organisations and NGOs. Current and upcoming AI projects are discussed, as are existing examples from practice in which AI is being deployed responsibly (use cases). Project plans are set up and discussed and contacts between the projects are arranged. Consortia are assembled around the projects, consisting of stakeholders backing them and developing them further with the help of funding sources such as the National Growth Fund and the mission-driven innovation programme called 'Data and Intelligence'. Barriers standing in the way of responsible AI-use will be tackled together in the working group by all the parties from the triple helix.

The Netherlands AI Coalition is *the* collaborative umbrella for stakeholders involved in AI in the Netherlands. It has a public-private structure in which governmental authorities, the commercial sector, educational and research institutions and social organizations work together to accelerate AI developments and link AI initiatives together. The NL AIC consists of working groups that focus on various building blocks and focal themes that are crucial for the development and application of AI in the Netherlands and beyond.

1.3 STRENGTHENING THE INITIATIVES

The strategic lines within the working group are aligned with the 2020-2023 Knowledge and Innovation Agenda (KIA) for Security. That agenda is derived from the Dutch government's innovation policy, which concentrates on mission-driven innovation policy for four themes within Dutch society. Security is one of those societal themes. Various missions have been formulated within that theme, such as the 'data and intelligence' mission: ***by 2030, security organisations will be gathering new data and better data, the correct interventions will be carried out using smarter analyses and there will be no surprises.*** A long-term, mission-driven innovation programme (hereinafter also "MMIP") has been drawn up for the purpose, paying attention to privacy-proof information sharing and decision support. Connecting the working group to the KIA will prevent duplication of work and will allow the two initiatives to reinforce one another.

1.4 THE LINKING FACTOR

The working group had built an ecosystem by bringing all the relevant stakeholders together. It is creating crosslinks between all the relevant stakeholders in terms of security and AI, and it supports the security sector across its entire breadth with responsible development and implementation of AI-related technology. The idea of having a specific working group for this theme came from the Data Science Initiative, which was set up by the municipality of The Hague to utilise the value created by data science and artificial intelligence for security, peace and justice. This initiative has now been absorbed into the working group and further embedded in the South Holland AI hub.

The working group ensures that all the relevant players collaborate and make agreements with one another about the content and the results. Security Delta HSD, on behalf of the Ministry of Justice and Security, has taken on the roles of chair and coordinator for both the working group and the 'Data and Intelligence' MMIP. HSD promotes and facilitates knowledge-sharing and linking parties together according to the HSD cooperation model: demand-driven, goal-oriented and based on trust. Linking the HSD network to the working group and the Security KIA allows initiatives to be connected, reinforced and scaled up.

Funding

Suitable funding or co-funding is sought, depending on the content goals of the projects and consortiums. This may involve calls on the National Growth Fund and Netherlands Organisation for Scientific Research (NWO), or on the use of instruments available to the Netherlands Enterprise Agency (RVO) and regional, European or other international programme funding.

To strengthen the position of the Netherlands and to make the most of the opportunities, a long-term programme called AiNed has been drawn up by the Netherlands AI Coalition. The programme accelerates the development and application of AI so that the Netherlands can reap the economic and social rewards of AI and keep pace with other leading countries. The long-term AiNed programme requires an investment of 1.05 billion euros from the National Growth Fund over the period 2021-2027, with an equivalent sum also being invested jointly by private parties (companies) and governmental bodies. The overall scope of the program is 2.1 billion euros. Recognising the importance of AI and the strength of the NL AIC and those involved in it, the Dutch government allocated €276 million in April 2021 from the National Growth Fund to the first phase of the long-term AiNed programme.

The working group is affiliated to the Security Knowledge and Innovation Agenda (KIA). The NWO KIC 'Data and Intelligence' resources contribute the overall goal of this knowledge and innovation agenda and the underlying multi-year 'Mission-Driven Innovation' programmes. This funding makes it possible for interdisciplinary consortia of knowledge institutes, public and private partners to submit proposals with a total scope of at least €750,000 but no more than €3,000,000 per project or consortium. The aim is to get humanities, technical and social scientists to help create better and more usable intelligence products that meet the needs of intelligence and security professionals for operational, tactical and strategic tasks.

1.5 NORMS AND VALUES

The development of AI for security, peace and justice is an interplay between the deployment of effective technology and the norms and values that belong in a democratic, digital constitutional state. Responsible use of AI puts the demand at the centre, rather than the technology. In the case of the working group, it is then about the demands and needs of security professionals, while paying attention to the interests of members of the public. The working group wants to help provide better and useful intelligence products based on artificial intelligence that complement and add value to operational, tactical and strategic tasks. Certain norms and values are deemed important in the Netherlands. Allowances therefore need to be made for them when AI is being used. This is only possible if all the stakeholders involved assess the norms, values and interests together in their appropriate contexts. The normative frameworks that are relevant here are based on a human centric approach to AI (technology should serve the people), looking at e.g. privacy, social acceptance, sovereignty, inclusivity, transparency and explainability.

1.6 RELIABLE AND ACCOUNTABLE

The strength of AI is that it can use smart algorithms to carry out functions that are normally associated with the human brain, such as solving problems and recognising patterns. To a certain extent AI can automate human brainpower. It works highly efficiently and rapidly, 'seeing' things that people do not notice. In forensic investigations and detective work, for instance, this makes AI a welcome addition. Using this technology is not without risks, however. AI uses data, for example, and data can be manipulated. Datasets can also be 'contaminated' with incorrect or incomplete information, so that AI learns the wrong things and generates incorrect outputs. This can also be the result of the programmers' assumptions and biases being built into the AI system unintentionally. That is why it is extremely important to apply AI transparently and make clear what data and algorithms are being used. The algorithms and the concrete outputs must therefore be reliable and checkable. This takes us, in technological terms, not only into the realm of explainable AI but also into the field of ethics. The working group is therefore squarely behind the appropriate and transparent use of AI for improving welfare and prosperity in the Netherlands.

1.7 THE BUILDING BLOCKS OF THE NETHERLANDS AI COALITION

The Netherlands AI Coalition (NL AIC) has working groups for what are referred to as 'building blocks': preconditions that support the working group's focal themes. The building blocks help make projects successful and may sometimes even be essential. In these working groups knowledge, expertise and solutions for the challenges are accumulated jointly. As these challenges can be found not only in the Security, Peace and Justice working group's focal areas but also in those of other sector-based working groups such as Culture and Media, Defence, Energy and Sustainability, Health and Care, and Technical Industry. These building blocks are described briefly below and will be explained further for the VVR domain in Section 2. As data sharing is partly about developing privacy-proof technologies, we have chosen to make that a focal theme of this position paper.

1 Human Capital

AI is inevitably bringing about changes in the labour market. It is replacing tasks and making work more pleasant and more efficient. There is therefore a great need for talented individuals who can program AI and handle it in practice. That requires training and education, which are in very short supply at the moment. This is an opportunity for the Netherlands. If the governmental, commercial and education sectors and representatives of society cooperate well in planning and implementation, they can develop high-quality and innovative education and training programmes that will draw attention (internationally too) and attract people. The intention is that some of these talented foreigners will stay and work in the Netherlands and that Dutch employees will be well prepared for a future that involves AI in their work.

2 Human centric AI

Our society is facing a major change. Applying artificial intelligence is being called a 'systemic change' – a paradigm shift that will alter society radically in the same way that the Industrial Revolution, the invention of the computer or the rise of the Internet did. AI will have a huge effect on how we interact with digital systems and services. What should society's attitude be to this digital transformation? How can we safeguard public values, fundamental rights and democratic freedoms? And how do we make sure that everyone can benefit from this progress? AI serves humans. That is the concept underpinning this building block.

3 Data Sharing

Artificial intelligence depends on data. The more relevant data is available and the higher its quality, the better the AI system can learn. In the Netherlands, data is often kept separate and not made accessible for others. This is generally done for legal or commercial reasons. To break down those barriers, data sharing needs to be organised properly and responsibly, much better and more rapidly than what we're accustomed to right now. Privacy-enhancing technologies are resources that make data sharing possible. This makes it possible for AI systems to learn from data, so that services can be provided better, more accurately and more carefully. Values such as trust, knowledge, privacy protection, data protection and democratic principles are the cornerstones of this.

4 Research and innovation

A combined approach to fundamental and applied research and innovations throughout the value chain will strengthen and accelerate AI research, AI developments and AI applications. To do that, an AI network of partners will be built up that expands upon the existing situation and offers scope for new initiatives. The Innovation Center for Artificial Intelligence (ICAI) has for instance been involved since the founding of the NL AIC as a leading national network of collaborative scientific ventures in AI.

5 Start-ups and Scale-ups

The working group is aware that its members are not the fount of all wisdom and that some of the solutions can be found in start-ups and scale-ups. The group therefore casts its net wider, looking at companies, start-ups and scale-ups from other domains that could be significant for responsible and human centric use of AI. These parties are placed centre stage at the working group's meetings. They get the opportunity to present themselves, act as sparring partners for various issues and make contacts with the coalition's members.

Definition of AI

AI has become a widely discussed and much-studied topic over recent years. The wide range of applications that AI can be used for and the risks and opportunities associated with such use mean that it is a very interesting topic for many different parties. A diverse set of definitions and descriptions of AI are used in practice. This publication is based on the definition of AI as drawn up by the European Commission – one that is broad, well-founded and that offers a good basis for going more deeply into the specific characteristics of AI.

The European definition of AI is: "Artificial intelligence refers to human-designed systems that, when given a complex goal, take action in either the physical or digital world by perceiving their surroundings, interpreting the structured or unstructured data gathered, reasoning on the basis of the knowledge derived from that data and deciding on the actions that are best (according to predefined parameters) to take in order to achieve the given goal. AI systems can also be designed so that they learn to adapt their behaviour by analysing how the situation was affected by their previous actions."

1.8 FOCAL THEMES

The working group aims to make the Netherlands a leading light in responsible deployment of artificial intelligence in the security domain. Setting up a long-term programme and a multi-year strategy makes it possible to set priorities in research and innovation, make investments in innovation over a longer period and agree procedures together. The working group concentrates on four focal themes:

1 *Realisation of solutions for privacy-proof information sharing*

There is a huge amount of information about people from a wide variety of sources that is not currently analysed cohesively, for instance because of privacy concerns. However, that is also standing in the way of applications that would be socially desirable. The working group is looking for technological solutions that allow data to be analysed cohesively while preserving the confidentiality of the data.

2 *Solutions for taking the correct decisions based on data and intelligence*

To keep security professionals and managers prepared at all times for the decisions that they have to take during an intervention, it is crucial to provide as accurate a picture as possible of the situation that demands operational, tactical or strategic actions from them. This could for instance be a situation or emergency involving a first responder's actions, or equally the analysis of documents for a legal case that a judge bases their ruling on.

3 *Solutions for improving cybersecurity technologies and cybersecure AI systems*

Digital threats and the attack methods used are developing rapidly. This focal theme is where the working group concentrates on improving defensive cybertechnology. The emphasis here is on the one hand on designing secure, automated and privacy-friendly systems and products, and on the other on developing techniques for making systems and products more resilient and keeping them that way.

4 *The use of language and speech technology for security, peace and justice*

The majority of algorithms for converting speech into text have been developed for the world's main languages and are owned by large tech companies. Although standard Dutch is supported to some extent, the numerous variants, dialects, accents and atypical speech patterns cover sales markets that are too small to be interesting for them. The technology that exists within the Dutch-speaking regions has been developed by specific parties to meet their own requirements. As a result, these systems are limited and not easy to use in situations where standard Dutch is not involved. To avoid dependency on foreign big tech companies, there is a need for sovereign and diverse Dutch linguistic and speech algorithms that are accessible for everyone.

2. THE BUILDING BLOCKS FOR SECURITY, PEACE AND JUSTICE

The results obtained with AI are improved if certain key preconditions have been set up properly. Development of AI training courses and education, attracting talented individuals nationally and internationally, transparency about how AI works, the ethical and legal contexts that AI systems work in, international collaboration and privacy-proof information sharing are essential if AI is to be used correctly and responsibly.

2.1 BUILDING BLOCK: HUMAN CAPITAL

Development and responsible application of AI require a firm underpinning of education, schooling and talent development. That base is only present to a limited extent. The working group wants to help resolve the shortage of training courses and new talent in the Netherlands. It wants to encourage educational institutions to prioritise including AI in the curriculum. It is also providing input for the nationwide human capital agenda for AI, training and lifelong learning.

The development of education is aimed initially at various minors with themes such as 'AI in society', 'AI in research' and 'AI in engineering'. These are half-year programmes that 1,500 students can take part in. The way AI can be embedded in existing subjects is also being examined, as is the development of new subjects in existing courses such as cybersecurity, forensics and law. AI is widely applicable and therefore relevant to numerous sectors and their associated training courses. On top of that, educational institutions are developing specific AI training.

There are currently not many people working on AI development or thinking about using it responsibly. Talented individuals are often tempted away to other countries or major companies. Given that AI is relevant for many sectors of our society, talent is in high demand. Efforts are therefore being made to reduce the shortfall by developing new training courses for AI, including minors, masters and new curriculums. Several International Talent Programmes are also being started up, aimed at attracting talented people internationally (to live and work) and then to help employers keep hold of that talent.

Another path along which solutions can be found is giving more people AI skills through lifelong learning (training courses, online education, AI tools for personal development) or making efforts to raise the profile of AI and digitalisation in general at secondary schools. The 'National AI Course' is now available on the specially designed platform ai-cursus.nl, for example. The course is free, accessible to all and aims to prepare as many Dutch people as possible for a future with AI. A number of courses are available, including a basic course, a children's course and courses for various professional sectors such as health and care or agriculture and food. On top of that, there are courses for professionals that can be taken at the Innovation Center for Artificial Intelligence academy (icai.ai/academy). Efforts are also being made to improve students' mobility. Curricula will be aligned so that students can follow multiple educational pathways, for example, from lower to higher vocational and on to academic studies.

Job vacancies, education and training through Security Talent and the NL AIC Training Platform

Security Talent is where supply meets demand in the security domain. Over recent years, thousands of job vacancies and work placements have been published and over two hundred employers have used the site. There is currently broad demand for AI experts in research and the governmental and commercial sectors. More than five hundred educational and training courses are also listed there, such as various bachelor's and master's programmes in artificial intelligence. Examples include the National AI Course and Elements of AI. The NL AIC Training Platform has been developed by the NL AIC's Human Capital working group. This page focuses on retraining, further training and re-education, offering the range of AI training available for and from the participants. This also includes the AI for Business & Government certification, a standard for professionals at the higher vocational level or above who want to make use of AI in their working environment. Professionals who have acquired 'AI for Business' certification have broadly-based knowledge of how AI can be applied in an organisation and how an organisation can be set up so that AI can be used.

Professionals, employers and organisations in the security domain will have to prepare through education and training for a time in which security professionals are backed up by powerful digital technology such as facilitating and self-learning AI systems like speech-to-text for Dutch, chatbots and virtual agents that can recognise emotions. The long-term, mission-driven innovation programme (MMIP) called "The Security Professional" is developing innovative and technology-driven forms of education and training, as well as a twenty-first-century skill set for security professionals.

2.2 BUILDING BLOCK: HUMAN CENTRIC AI

Residents of the Netherlands need to be able to make the right choices when using AI in their daily lives and should (if possible) be involved in developing new AI services. That is why we are looking for ways to learn together and discover the best and most desirable AI solutions. The working group has adopted a human-centred approach to AI in which AI systems are developed in an ethically responsible and socially meaningful way. VVR is a field where AI applications can have a major impact on the public at large. It is therefore essential to highlight the general public when developing systems and to include them in the development process from the start. The relevant normative frameworks are discussed in the working group, along with what they mean for its members and how those members can then best set up their systems to suit.

Additionally, the working group encourages the human centric approach with the initiatives and helps those parties put these principles into practice. There are for instance several Ethical, Legal and Societal Aspects (ELSA) labs under development for VVR. This allows a contribution to be made to existing and upcoming issues such as meaningful human control. This will be explained in more detail in the next section.

2.2.1 EU HUMAN CENTRIC APPROACH

The European Union (EU) is fully on board and wants to stand out from the crowd with responsible, human centric AI. They expect that adopting a human centric approach will allow them to improve the lives of people living in the EU considerably. Additionally, it will yield major benefits for society and for the economy. To reinforce this approach and limit the risks associated with AI products, the European Commission has drawn up regulations. If the rules are not observed, fines can be imposed that can reach six per cent of an organisation's annual turnover. Taking the upcoming regulations into account in AI experiments already allows us to make sure systems are feasible and better prepared for when these rules come into force. Above all, this will let us avoid problems that might otherwise only come to light too late.

The regulations focus on three types of AI: banned AI systems, high-risk systems and others. AI systems that affect people subconsciously in harmful ways are banned, as are AI systems that can exploit people's vulnerabilities. AI systems that are used for social scoring (estimating and assessing reliability and desirable behaviour through social credit systems) and real-time biometric systems for detection and enforcement are also forbidden. The bulk of the regulations are about high-risk systems. These include AI systems that are used in toys, aviation, education and medical equipment, as well as ones for applications in jurisprudence and detective work. The most prominent requirement for high-risk systems is the obligation to have a conformity assessment carried out before the product hits the markets. The rule for other AI systems is that people must be informed about the use of AI and how the AI system is being deployed.

2.2.2 ETHICALLY RESPONSIBLE INNOVATION TOOLBOX FOR GOVERNMENTAL AUTHORITIES

In addition to the EU legislation, the authorities are also working through other human centric AI tools for guaranteeing public values and human rights when AI is used. These tools are to be gathered in the online toolbox called 'Ethically responsible innovation for governmental authorities'. This refers to the following three tools (in Dutch – names translated here): the 'AI system principles manual for non-discrimination'. This tool was completed in January 2021 and it offers a practically applicable design framework that helps developers to identify, prevent and tackle discriminatory actions in data as far as possible from early on in the development phase of an AI system. Additionally, that summer saw the publication of the Impact Assessment for Human Rights and Algorithms (IAMA), which can be used when choosing whether or not to develop an AI application and then to help the development and implementation proceed responsibly. The third instrument developed is the Code for Good Digital Public Administration. This code pays attention to the consequences of digitalisation for public administration, based on principles and values for democracy, a constitutional state and administrative power. An overview of the relevant frameworks for the use of AI can be found in the appendix.

2.2.3 PUBLIC CONTROL OF ALGORITHMS

The four biggest cities of the Netherlands, the provinces, police and Directorate-General for Public Works and Water Management (*Rijkswaterstaat*, RWS) have joined forces to develop policy instruments in the project called 'Public control of algorithms'. Algorithms are increasingly being used for supporting human tasks or even taking them over, within the authorities and elsewhere. Tools are therefore needed for ensuring that algorithms are checked and protecting the public against (for instance) incorrect results. This latter aspect also plays a role when applying the 2019 guidelines about protection against the risks inherent in data analyses. The key aim of the modified guidelines is creating transparency and meeting the conditions for counteracting the potential risks of the use of AI algorithms, e.g. explainability (justifications for using AI) and auditability (control mechanisms). These modifications have been aligned with a questionnaire from September 2021 showing that three quarters of over a thousand Dutch people surveyed deem themes such as privacy, human control of the algorithm and the reasons why the algorithm is being used to be important. They would like to be given more information about these topics.

2.2.4 THE AI PEACE PALACE

Under the name *AI Peace Palace*, the working group is working on implementing an international forum for responsible AI in the security, peace and justice domains. The forum's mission is to assist Europe in becoming the first intelligent, fair and secure AI society in the world. An innovative, international ecosystem is being created to that end consisting of governmental authorities, companies, centres of expertise and social organisations. Using the mission of the Data Science Initiative as its foundation, the forum is supporting innovative projects that are related to data and AI and that focus on security, peace and justice. In parallel with this, the AI Peace Palace is bringing parties together to define the governance structures and anchor them in the legal system. The forum also wants to encourage debate within society about AI.

The Hague Conference on Responsible AI for Peace, Justice and Security.

This conference, held in the Peace Palace in The Hague, is helping Europe develop to become the first intelligent, fair and secure AI society in the world. The Ministry of Justice and Security, Ministry of Foreign Affairs and the municipality of The Hague are organising this high-level meeting in 2022 at which the following themes will be discussed: 'International Policy & Law', 'Technology & Application', 'AI & Cyber' and 'Future of the City & AI'.

Hackathon for Good.

The mission of Hackathon for Good is to use hackers and innovators to assist in the innovation programmes that are being started up by governmental authorities, companies and centres of expertise and that are related to a social issue. They are assisted in making and testing prototype products that utilise AI and data. Attention is paid, for instance, to using blockchain to counter disinformation and AI techniques for recognising deep fakes. Hackers and innovators are also working on e.g. counteracting food waste and flooding using technology. At the annual event, more than a hundred participants from twenty countries worked on use cases from the Public Prosecution Service and the Ministry of Defence, among others.

2.3 BUILDING BLOCK: RESEARCH AND INNOVATION

A network of partners ensures that fundamental and applied research into AI can be done. Governmental authorities, companies, centres of expertise and social organisations are collaborating in research and in developing AI innovations. Research and practical applications are kept close to each other so that they can affect each other and reinforce the results.

2.3.1 COMBING FORCES

Getting fields of expertise from various disciplines to work together and enhance one another adds a great deal of value in research and innovation in AI.

CLAIRE and ELLIS

The European research network CLAIRE (Confederation of Laboratories for Artificial Intelligence Research in Europe) is aiming to set up a pan-European network of Centres of Excellence in AI for encouraging existing talent and providing a central point for researchers to interact and exchange ideas. Furthermore, the European ELLIS network unites excellence in AI research, particularly in machine learning, into three units in the Netherlands, namely the University of Amsterdam, Radboud University and Delft Technical University. This will enhance the flow of knowledge between European researchers and their institutes.

Collaboration between the universities of Leiden, Delft and Rotterdam

An example at the local level is the collaboration between Leiden University, Delft Technical University and Erasmus University Rotterdam. These universities have joined forces for *Convergence in AI, Data and Digitalisation*. Working together and combining knowledge about law and business with technical knowledge and expertise in Administration, policy and ethics is a valuable way of letting the fields of expertise of the research groups complement each other. Together, the researchers are carrying out pioneering research focusing on the social challenges of applying AI in the peace, justice and security domains, such as finding the correct balance between new and optimised functionalities while at the same time protecting the public and the institutional principles of our society against vulnerabilities and dependencies on AI.

There are opportunities for multidisciplinary research into both applied technology-related questions (IN AI) such as e.g. NLP for forensic investigations, in applications such as security, cybersecurity and jurisprudence (WITH AI), and regulatory issues, ethics, the dynamics of accountability, policy implementation, transparency, explainability and AI applications that will benefit public safety. The three universities have close links with start-ups, policy bodies, executive bodies and the companies that work in or for this domain. They are working with those parties to construct a leading network for AI that can tackle the opportunities and challenges of human centric AI in a multidisciplinary and integral way.

Ownership and support

In innovative projects, it is important to involve the problem owner, budget holder, recruiter and end user through public-private partnerships. It is sensible to take the time to explain the technology at the management level and underline its significance. Direct cooperation with legal departments is needed. If all the stakeholders are involved in the project development right from the start, it prevents unnecessary and unforeseen obstacles appearing that affect the implementation.

The deployment of AI is all about social acceptance. Collaboration with the public and with social organisations is needed in both the development and the use of AI to bridge the gap between scientific knowledge and society.

2.3.2 ETHICAL, LEGAL AND SOCIETAL ASPECTS LABS (ELSA)

ELSA covers the 'Ethical, Legal and Societal Aspects' that play a role in the development and implementation of technological innovations in society. In the labs (some of which are virtual), governmental bodies, companies, centres of expertise and social organisations carry out research together. Human centric AI solutions are developed, tested and implemented through co-creation. For the security domain, various labs will be set up over the coming period, one of which will focus specifically on the defence domain (led by TNO) and another that will focus on public AI systems within the judicial chain (led by the universities of Leiden, Delft and Rotterdam). A network will be set up around that to ensure that ideas and results are shared among them and can be scaled up and ultimately end up with the people who will work with them on a day-to-day basis.

Defence

The introduction of AI systems at the Ministry of Defence raises numerous ethical, legal and social questions such as that of how the systems can be kept under control by humans. Another question is how human power, dignity and responsibility can be retained when autonomy is given to machines. Who is responsible for the decisions made by those machines? The Defence ELSA Lab is going to set up a future-proof and independent ecosystem with experts and companies who can be consulted and utilised for responsible use of AI within the defence sector. A methodology will also be designed that ensures that the ELSA aspects will always be included in AI development within Defence.

Justice and Security

AI systems are being used increasingly often in the public domain. AI's share in democratic processes and public services is increasing, which can have major effects for the public at large. This ELSA lab carries out research into the options for retaining human control over the AI systems that are used by government and semi-government within the security domain. The research plays a role in the fair and just treatment of members of the public by enabling fair decisions through AI. The AI systems will be corrected by an effective approach (yet to be developed) so that the public can be confident that their rights will be protected. Implementing various ELSA checks and balances will let this will add to ensure that AI systems in the public domain help strengthen the rule of law.

2.3.3 NATIONAL ARTIFICIAL INTELLIGENCE POLICE LAB (NPAI)

The National Artificial Intelligence Police Lab (NPAI) is a collaborative initiative by the Dutch police forces, the University of Utrecht, the University of Amsterdam and Delft Technical University. The police lab is part of the national Innovation Center for Artificial Intelligence (ICAI). The partners in the collaboration are aiming to develop state-of-the-art AI techniques to support the police force in its operational processes. Through the NPAI, the partners want to improve security in the Netherlands in a socially, legally and ethically responsible way. The lab is working on techniques that cover the whole spectrum of AI: machine learning for getting the right information out of multimodal sources (photos, text and videos), algorithms for reasoning using information in e.g. legal or other documents and reconstructed crime scenes, simulations of complex criminal systems and robotics. Aspects such as transparency, privacy and explainability are every bit as important here as accuracy, computability and efficiency. The collaboration between police and scientists ensures that the police can use the latest techniques, whereas science gets interesting problems to investigate from everyday practice. In addition to scientific articles, the lab has also provided several applications for the police already, such as the smart selection tool for Internet fraud and an 'explainable AI toolbox' for data scientists working in the police.

2.3.4 THE AI AND LEGAL TECH LAB (AILT)

The AI and Legal Tech Lab of the Haagse Hogeschool has received funding from the municipality of The Hague to help support the position of the Netherlands as an international hub in this up-and-coming area. The AILT lab carries out legal risk and social impact assessments for new technologies to assess the consequences for society and the potential controversies through the lens of legal analysis, assisted by research carried out by the students. The AILT functions in the same way as a legal clinic and it has top-class international partners, including Stanford University. The lab provides a practice field for business administration, law and safety management students. Exposing them to genuine AI implementation problems let the lab prepare students as well as possible for finding multidisciplinary solutions to complex AI governance issues in the future. Previous study projects at the AILT lab included cooperation with data scientists, computer scientists and lawyers from respected organisations such as CWI, Deloitte and TNO, as well as various national and international start-ups.

2.4 BUILDING BLOCK: START-UPS AND SCALE-UPS

The active members of the working group include a large number of businesses that work with both the government and centres of expertise on a per-project basis to continuously improve their products and make a contribution to security, peace and justice. Many of them are involved in one of the projects listed in this publication. In addition to the numerous projects, the working group also provides a stage for start-ups and scale-ups. They get the chance to present their innovative solutions to the members of the working group. This yields interesting connections to the group's members and to its projects. The start-ups and scale-ups also get an opportunity to present questions and obstacles they are encountering to the group members and discuss them. We will list a few examples of organisations that presented their work during the working group's meetings.

2.4.1 LEGALAIR

LegalAIR arose from Gimix and BG Legal. This project aims to make knowledge about AI and its application accessible to the public, authorities, commercial sector and centres of expertise. Everyone who is affected by AI must be able to find answers to legal and ethical questions easily. To date, many AI projects are delayed or do not get off the ground because the knowledge shortfall is too great and too much is unclear. Nevertheless, answers can be given to many questions. That is the aim of this knowledge platform. Additionally, this platform provides template documents such as contracts and intellectual property clauses. Experts are also affiliated with the platform so that anyone with a specific legal or ethical question can speak to them. Ultimately, LegalAIR wants to create clarity about AI so that organisations will see and make use of the possibilities and opportunities.

Innovative SMEs for Security, Peace and Justice

Amberscript
Avantopy
awen collective
aXite Security Tools
AXVECO
Bitdefender
Centillien
Compumatica secure networks
Considerati
Cursor Asset Management
Cybermundus
Cybersprint breakthrough security
Data Science District
DataWorkz
dialogic innovatie • interactie
Duck Duck Goose
Dynaxion
EclecticIQ
Experience Data
FruitPunch AI
Go2the.cloud
InControl
Labelfuse
Oddity.ai
OPT/NET
orangehill
Pandora Intelligence
Parito
publicsonar
reprex
Researchable
robin radar systems
RosemanLabs privacy by design
SemLab
Squerist
Syntho
thrive dynamic
UbiOps
Xomnia
Zeta Alpha
ZIUZ visual intelligence

2.4.2 THE GLOBAL AI FOR GOOD COMMUNITY

The worldwide 'Global AI for Good' community, which is called FruitPunch AI, is using AI for major challenges that humanity is facing, such as sustainability issues. FruitPunch AI has, for instance, already helped in the protection and preservation of the environment and wild animals. In future, it wants to develop an unmanned aircraft that will initiate a revolution in the way that nature reserves are monitored and inspected. The community organises various events and experts from FruitPunch AI offer their assistance in tackling social challenges. The muscle of the FruitPunch AI community is being linked to the working group so that it can also help resolve security issues in the future.

2.4.3 COUNTERACTING BIAS AND DISCRIMINATION IN AI MODELS

One of the worries when AI is being trained is that bias can arise, caused by the source data that is used. Bias is deemed to have occurred if external factors have a negative result on the outcomes. The source data may, for instance, record the fact that a certain population group commits tax fraud. This does not have to be true in general, though, and may be the result of investigations in a particular period. An AI system that is given this data to process or use for training may potentially exacerbate the bias in the data, for instance by making more errors where a particular population group is involved or making more negative assessments about that population group.

Counteracting bias and discrimination are demonstrably possible, even if the source data contains preconceived views and opinions. Centilien, a computer vision platform, has trained AI systems and shown that counteracting bias and discrimination is perfectly possible. It cannot be prevented entirely – but neither is that possible with humans. What the system can do is make clear what the basis was for the result or recommendation. That transparency is essential if AI applications are ultimately going to be used.

3. FOCAL THEME: PRIVACY-PROOF INFORMATION SHARING

One of the focal themes of the working group is producing solutions for privacy-proof data sharing between various parties. A vast amount of data is in fact available about people, derived from all kinds of sources, that is not allowed to be analysed because of privacy considerations. If we are to use AI responsibly, it is essential that data can be shared and used in a way that respects privacy – particularly in the security domain, where the privacy of members of the public is an even more sensitive topic than in some other domains. That is why privacy-proof information sharing also acts as a building block for other focal themes.

Data in the security domain is hugely fragmented and located with numerous different parties. Municipalities, the Public Prosecution Service, the police and the security regions each have data that is relevant for preventing, detecting and prosecuting criminality, for instance. If AI is to be trained for security issues, that data is needed. However, the General Data Protection Regulation (GDPR) and the risk of data being made public by others mean that the parties concerned are very cautious about sharing data with each other.

The working group is looking for solutions to this issue so that data can be analysed and used within the security chain but without undermining individuals' privacy while doing so. This has to take account of explainability (justifications): when AI is used in the public domain, it must be possible to explain how it works and how it gets to the outcomes. It must, for example, be clear what the AI system's recommendations are based on.

Sharing data is one of the building blocks that are crucial if AI is to be used well and responsibly. In the context of the focal theme of 'privacy-proof information sharing', the working group gratefully uses the results produced by the NL AIC's working group specifically for Data Sharing. Conversely, the working group's initiatives and results also let it contribute to the Data Sharing working group.

3.1 PRIVACY-ENHANCING TECHNOLOGIES

One of the solutions for the data challenges is privacy-enhancing technologies (PET). These are cryptographic techniques for processing privacy-sensitive data (such as personal data) in a privacy-friendly way. They comprise a collection of techniques that are used within an information system to ensure that the personal lives of people are protected. This is done by preventing unnecessary or undesirable processing of personal data. PET makes it possible to use data in such a way that only the requisite minimum of information is made known to those involved. This technique can ensure that there is no longer a contradiction between privacy on the one hand and security on the other. The societal task of achieving a secure society then goes hand in hand with the protection of personal lives.

3.1.1 MULTI-PARTY COMPUTATION

How can an organisation exchange data without infringing privacy? By taking privacy into account straight away during the design stages of systems and technology. This is called 'privacy by design'. One such technique is 'multi-party computation' (MPC). This is a smart way of analysing data jointly without having to reveal it. Cryptographic techniques make sure that multiple parties can analyse data together and draw conclusions without ever having to see each other's data. Calculations and analyses are therefore carried out on encrypted data and only the result is then decrypted. MPC therefore means that no data is revealed – only the conclusions based on that data. Moreover, this technique ensures that only pre-agreed analyses can be carried out. This prevents inappropriate use of personal data. This is one of the things that are being worked on in the Techruption programme in the 'Multi-party Computation' field lab on the Brightlands innovation campus.

3.1.2 FEDERATED LEARNING

Another technique that can be categorised as 'privacy by design' is federated learning (FL). In this technique, the analysis is taken to the data instead of the data being brought to the analysis. With FL, it is no longer necessary to collect lots of sensitive data at a single location. The data remains decentralised, with the data owner, although it is still possible to use it for analyses. Decentralised calculations are carried out at the sites where the data is located. The results of all these partial calculations are shared with one or more parties who can draw overall conclusions from them. Like MPC, FL shows that sharing data is not necessary for nevertheless acquiring useful insights from several distributed data sources, while still ensuring that privacy and confidentiality are observed.

3.1.3 MA3TCH

Ma3tch is another privacy-by-design technique. It is used by parties that might want to use each other's data, in line with the requirements of the General Data Protection Regulation (GDPR), which states that parties must exchange the minimum amount of data to achieve the specific objective, that they may only exchange data for that one purpose ('purpose limitation') and that they must be transparent about it. Ma3tch is a technology that ensures parties comply with the GDPR when processing personal data. Thanks to Ma3tch, it is possible to 'see' if the other party has information that is relevant without having to exchange datasets. That is shown by an autonomous anonymous analysis (a3), after which targeted exchange of data can take place. The targeted retrieval of information, in contrast to querying large data files, also makes international cooperation between countries easier, meaning that the Dutch authorities and organisations will also be more inclined to exchange data this way.

3.1.4 SYNTHETIC DATA

The techniques and technology discussed above are all classed as privacy-enhancing technologies (PET). A fourth PET technique involves generating synthetic data. Privacy-sensitive information is replaced by completely new and artificial data. Using AI, it is possible to get a computer system to learn all the features, relationships and statistical patterns of an original dataset and then use that knowledge to 'invent' completely new data with the same features, relationships and statistical patterns. The data can then be used for all kinds of purposes without people's privacy being infringed. In collaboration with the NL AIC, Syntho and SAS have carried out research into the added values of synthetic data and the best opportunities for its use.

3.2 INITIATIVES AND APPLICATIONS

In addition to the techniques being developed, there are various initiatives that are relevant for the focal theme of 'privacy-proof data sharing'. We will list a few examples.

3.2.1 FIRE

In investigations of serious crime, the police often confiscate data carriers such as mobile phones, computers and hard drives to look for incriminating evidence. The police seize so many data carriers that it is impossible to search all the data items manually, one by one. Data scientists at the NFI have therefore developed a self-learning algorithm that can quickly extract specific types of images from all the data. It can see whether there are weapons or drugs on a photo, as well as recognise textual elements on photographs such as car registration plates or the bank account numbers on stolen passes. This AI technology has been included in the Hansken search engine.

3.2.2 A TRUSTED DIGITAL INFRASTRUCTURE FOR REPORTING

Under the name '*Informatiehuis Meldingen*' (Reporting Information House), work is being done on setting up a trusted digital infrastructure. This is a combination of a data warehouse and knowledge centre, aimed specifically at a safe living environment. The partners involved (NVWA, the police, the Public Prosecution Service, Oost NL, Saxion, Samen Veilig, Delft Technical University, University of Utrecht) are aiming to accelerate cooperation in the chain, avoid risks, infraction and incidents, and detect them more rapidly. The data that is collected is stored per data owner (holder of the source) and knowledge is extracted from it using inter alia privacy-enhancing technologies. Various data sources are consulted in order to acquire new knowledge. An example of this is combining satellite and drone data of the surroundings with public data to detect the locations of illegal drugs laboratories. Critical preconditions for setting up an information house include unambiguous goals and definitions, the absence of commercial interests, scientific calibration and utilisation of technological innovations, plus compliance with the conditions set for privacy and ethics.

A trusted digital infrastructure can serve various purposes. Investigations are for instance being carried out to see whether companies and third parties sharing sensor data and other relevant data through such an infrastructure can lead to better, data-driven security of industrial sites. There are also studies into whether data sharing through a trusted digital infrastructure gives valuable understandings that will help set up disaster management better and prevent incidents. Additionally, studies are looking at whether real-time sensor data and other data from buildings, residents and organisations can be used as the basis for a more effective approach to disasters: emergency service staff can for example do their work better and more safely if they have up-to-date details of people, buildings and systems.

3.2.3 SUSTAINABLE RESCUE

Human trafficking is a hidden crime. It is often a concealed component of criminal organisations. A major lack of communication between all kinds of bodies involved with the topic – not sharing data and information – means that there is only a poor picture of the problems and little progress is being made. The problem is in the data sharing. It is available from all kinds of bodies but often only shared when it is already too late. Sustainable Rescue has therefore applied the FAIR (Findability, Accessibility, Interoperability and Reuse) guidelines to this data. This means that the data will be linked (while complying with all legislation and regulations) to what are known as 'data stations' or 'data hotels' in which all those bodies are connected. Together with Roseman Labs and the Data Sharing Coalition, a prototype has then been created that allows data from informers to be shared securely. This secured information is used by Deloitte to develop human trafficking crime scripts that can be used by police forces. Because human trafficking does not stop at international borders, these crime scripts can be a powerful tool for disseminating knowledge and understandings at the European level too.

4. FOCAL THEME: AI, DATA AND INTELLIGENCE FOR DECISION SUPPORT

A first responder such as a police officer on the beat has to take lots of decisions every day. They therefore need an accurate picture of the situation to base those decisions on. Supporting them with analyses of real-time data and giving them a picture of what is happening can help them choose the right way to intervene. The AI system that is used for this could even tell the first responder which intervention would be the first choice and what an alternative could be. Another example of decision support is in legal cases or a control room. Spoken words on sound or image carriers can be recognised and converted into written text that is then analysed. All the data is combined and the analysis then gives a recommendation to the judge or triggers an action taken by the control room. Speech-to-text can also be useful in courtrooms, where everything that is spoken can be converted into a written report of what has been said. Transparency, accountability, explainability, verifiability, traceability and non-deniability of the recommendations made by such initiatives are essential if AI is to be used responsibly in this field.

4.1 INITIATIVES AND APPLICATIONS

Several initiatives and applications are currently ongoing that are relevant for the theme of 'AI, data and intelligence for decision support'. We will list a few examples.

4.1.1 KANSRIJKE KOPPELING (PROMISING CONNECTIONS)

The Central Agency for the Reception of Asylum Seekers (COA) matches residence permit holders (e.g. refugees) to a municipality. This includes looking at where they will have the best chance of building up a new life and contributing to society. Based on a digital profile that looks *inter alia* at education, work experience and ambitions, the residence permit holders are matched to municipalities, where they are then given accommodation. The COA has set up a project in which they are investigating how AI can be used to improve the integration outcomes (including labour market participation and study uptake).

4.1.2 QUIN

To reduce the burden of what analysts need to read in liquidation cases and track down untraceable criminals, TNO used artificial intelligence to develop an application for predicting the behaviour of fleeing criminals. In the *'Onvindbare Veroordeelden'* programme (Untraceable Criminals), the police, Public Prosecution Service (OM), Central Fine Collection Agency (CJIB) and the Justice Information Service (Justid) are using QUIN (Question & Investigate). This software application has been trained to make is smart enough to predict where criminals will flee to. QUIN is a useful tool for solving cases more quickly when time is of the essence, but it cannot replace the human component. Access to sensitive data in ongoing cases is often awkward, so this technology was also tested in the television programme *Hunted*. In addition, Pandora Intelligence has implemented QUIN in its scenario software to allow probable escape scenarios to be predicted, for example after a ram-raid or terror threat. Numerous other applications are possible too and this innovation is therefore being seen as a breakthrough in the domains of detection, information and combating terrorism.

4.1.3 LEGALLY AND SOCIALLY ACCEPTED AI IN THE COURTROOM

The project for legally and socially accepted AI in the courtroom will be investigating how the data that has been collected can be turned effectively into legal evidence by an AI-controlled toolkit, in a way that is acceptable to society. The requisite social acceptability is assessed in test cases. The AI system used produces visualisations of the data found in order to assist investigators who are looking for relevant evidence. The interplay between humans and machines leads to well-grounded decisions. The project is creating both AI software and visualisation software, as well as a standard guideline for digital evidence and procedures for automated production of legal evidence. This is being done at the National Artificial Intelligence Police Lab, in collaboration with police detectives and judges.

4.1.4 HANSKEN

The volumes of data and data sources that need to be investigated in criminal cases are increasing rapidly, especially in fraud, murder and cases of child sexual abuse materials. To increase the effectiveness and speed of these investigations, the Netherlands Forensic Institute (NFI) has developed Hansken, a forensic search engine. Hansken enables quick and efficient search through large numbers of confiscated data carriers such as computers and mobile phones. Hansken makes this possible by automatically structuring and indexing the data. Anything that might be relevant can be searched for, e.g. words and names or characteristics of traces such as e-mails, chat messages or photos (taken with a specific camera or in general). Thanks to AI, thousands of photos can then be searched through specifically, looking for a drug container for instance, because the programme has learned what a container looks like from training data. AI also makes it possible to analyse e.g. thousands of text messages and only show the relevant ones.

Evidence found with Hansken is being used as evidence in more and more criminal cases. This requires substantiated forensic evidence: the material must be processed transparently and the evidence reported must be traceable to the source (e.g. the confiscated phone or computer). The principles of transparency and explainability are massively important here. The Netherlands Forensic Institute (NFI) and Hogeschool Leiden (HS Leiden) have agreed that information sciences students will learn how to use Hansken. To that end, the search engine will be installed in the college's IoT Forensic Lab at the HSD Campus. Besides gaining an understanding of how the search engine works and the AI algorithms behind it, the students also learn how to update the search engine for new digital developments. The open nature of the software means that students also learn how to build new tools for the search engine themselves and thereby extend its capabilities.

4.1.5 CHILD SEXUAL ABUSE MATERIALS

Hundreds of millions of images and videos of possible child abuse are passed on to the police worldwide. It is a gigantic amount and investigating everything is hugely labour-intensive. As well as images, there are chat logs, geographical data, usernames, e-mail addresses, IP addresses and audio recordings. To avoid important items remaining unnoticed in that huge mountain, AviaTor has been developed by ZiuZ Visual Intelligence and Web-IQ within a European consortium that includes the Dutch and Belgian police forces. This system improves the information position of investigators so that they can prioritise the most significant cases and deal with them first. The AviaTor AI system maps out the contents of large datasets of child sexual abuse materials (CSAM) and adds unique hash codes. Images can be retrieved and examined using those hash codes and then used as evidence.

4.1.6 IMPACT COALITION SAFETY AND SECURITY – SAFE AND SECURE SMART CITIES

The Impact Coalition Safety & Security is a collaboration between municipalities, the police, knowledge institutes, companies, the Association of Netherlands Municipalities and HSD Office. The coalition is part of the broader Smart Society movement. The coalition was founded in April 2020 through a declaration of intent entitled the 'Impact Coalition Safety and Security for Smart Society'. Within that coalition, the municipalities of The Hague, Amsterdam, Almere, Eindhoven and Apeldoorn are in the lead. The efforts of municipalities and the police for public safety are reinforced by adding to the innovative strength of those organisations. Two application areas are being focused on in particular: innovations for crowd management and area security. The mobility flows of the visitors receive special attention in that regard.

4.1.7 POLICY IMPLEMENTATION

AI support for risk identification, decision-making and policy-making can significantly enhance public and national safety and cybersecurity. There are however still some important questions about how to use AI effectively and responsibly, which is why Leiden University is focusing on the question of how AI-driven analyses can be implemented in organisations' primary processes within the policy context, what this requires from organisations, and how it can then lead to innovations. A second question concentrates on bringing together two worlds within organisations: demand for AI-supported insights and assistance for decision-making on the one hand and its actual use on the other. A third question is about management-level accountability for AI and its use of data. The following issue is all about making clear what the algorithms do and what effects the outcomes have on the successive people or bodies involved in a chain. It also addresses the question of how norms, values and interests are weighed up in the responsible use of AI – and by whom.

5. FOCAL THEME: AI FOR CYBERSECURITY

Despite increasing investment in cybersecurity, most organisations are scarcely able to keep pace with the speed of developments and digital threats. A medium-sized company gets hundreds of thousands of alarms per day. Automating cybersecurity work is therefore a key solution for becoming and remaining resistant while only using limited human and financial resources. That is why the emphasis in this focal theme is on designing secure, automated and privacy-friendly systems and products, as well as on developing techniques for making systems and products more resilient and keeping them that way.

An essential element of designing such products is investigating the interactions between humans and automated AI security tools. Various knowledge and innovation issues are relevant for this, in particular the development of ways of automatically signalling vulnerabilities in source code, automating commonly occurring steps in the penetration testing process, automated patching at the application level, automation of the operational tasks in security incident and threat analyses, and automated detection of whether or not an organisation is compliant with respect to legislation, standards and its own policies.

New and innovative techniques such as applying AI and machine learning to detection and response processes offer solutions for making organisations resilient and keeping them that way. In the longer term, AI and machine learning in detection and response processes can replace human efforts, so that cyberprofessionals – a scarce resource – can be deployed in other, supplementary ways. AI has in the meantime become integrated into various processes such as detecting botnets, machine learning for software vulnerabilities and automated detection of hacks into industrial environments (e.g. in industrial ICS-SCADA systems).

5.1 CHANCES AND OPPORTUNITIES

As the number of data leaks and cybersecurity incidents increases, AI is being praised more and more as the new way of automatically detecting malware on networks, supervising the response to incidents and detecting crimes before they occur. This positive image of what AI can do is not entirely incorrect, but the expectations of what the next generation of AI techniques will be capable of do need to be tempered somewhat. It will still take years before AI is capable of producing analyses without human input. Current developments offer opportunities to reduce the load on overburdened cyberanalysts and counteract the shortage of cyberspecialists. Using AI as a defence against cybercrime also appears to be significantly more cost-effective, compared to the situation in which it is not used. It is expected that cyberattacks will become more complex, larger in scale and more dynamic. Humans’ abilities will no longer be sufficient for detecting and anticipating such attacks. The development of new defensive AI systems that respond adaptively using machine learning will help cyberanalysts defend digital systems.

5.2 AUTOMATED SECURITY OPERATIONS

Security Operation Centres (SOCs) are seen as crucial in detecting attacks and are the core of most cybersecurity strategies. At the same time, the number of attacks is increasing, personnel is thin on the ground and an ever-increasing number of cyberattacks are automated. This means that upgrades of existing SOC platforms are needed. One solution for this is automating security operations. In 2020, with backing from the Ministry of Economic Affairs and Climate Policy, TNO founded a consortium called Automated Security Operations (ASOP). The consortium wants to develop an automated security platform (in a public-private partnership) that will allow organisations to detect and respond to automated cyberattacks more quickly. This is intended to make it easier for the entire chain of end users, system integrators and developers to resist complex cyberattacks, both proactively and reactively.

6. FOCAL THEME: USE OF LANGUAGE AND SPEECH TECHNOLOGY

Organisations are having to deal with ever-increasing volumes of unstructured data in textual form, as well as time-consuming processes such as taking minutes of meetings, questioning and other verbal interactions. Language is everywhere – and fortunately, artificial intelligence is providing more and more opportunities for analysing text and speech automatically. It is expected that natural language processing (NLP) and automatic speech recognition (ASR) will be able to make our lives better and simpler over the coming years.

Potential applications include finding relevant information in large amounts of text, automatically transcribing conversations between patients and doctors, detecting emotions in a person's voice, detecting various medical conditions such as early-onset parkinsonism, dementia or Covid-19, and controlling devices through automatic speech recognition. All these capabilities of NLP and ASR have applications in various sectors such as security, government, education, new media and healthcare.

There have been major developments in language and speech technology in recent years. However, it seems that the technology does not yet offer good enough results for Dutch text or speech. First of all, the quality of the generic models for Dutch is not good enough for them to be used directly in organisations, and there seems to be a great need for further development of specific language models that focus on domain-specific terminology, dialects, street slang or foreign accents. Developing such models is, however, a costly process. Although Big Tech such as Google do offer models for Dutch, the sales market for specific language models is too small to interest those parties in supplying customised models. Moreover, it is undesirable or impossible for many organisations to let their data and processes go through a non-Dutch party and be dependent on a large foreign tech company. Moreover, guarantees cannot be given that these language models (where available) are inclusive, transparent and bias-free.

6.1 LANGUAGE AND SPEECH TECHNOLOGY IN DUTCH FOR THE SECURITY DOMAIN

For investigative work and all kinds of other tasks in the criminal law chain, it is important to have good algorithms for Dutch that can also handle dialects, slang, accents, children's language and speech disorders. This language and speech technology will analyse spoken and written Dutch, establish relationships between spoken and written language, and convert the spoken word into text. Applications such as automated analysis and processing of phone taps, reports, questioning, reports, emergency calls and much more are all possible.

Moreover, speech technology is seen as promising in the security domain, as it can be used to identify and verify people. In identification, a person's identity is determined by comparing their voice against a broad dataset. Verification works differently. That is the case where someone claims a certain identity and the software uses an audio recording to determine whether a voice is authentic and indeed belongs to that specific person (identity). All this requires is a recording of each person stored in advance.

Verification by speech (voice authentication) can be an additional safety factor for combating online fraud or securing buildings. In some cases, it can be used as a signature. In the United States, voice recordings of the words "I agree" can be used to sign a digital contract, for example. Voice authentication can also be used for tracking down criminals; in that case, it is about identification. A sound fragment recorded during a shop robbery, for instance, can be listened to and the police can then identify the perpetrators on the basis of a broad dataset of voice IDs.

Conversely, speech technology can also impinge upon people's security. Voice data can be stolen and misused, for instance to commit identity fraud. Despite the improvements, speech technology is not error-free and mistakes can happen. Before speech technology can be used for critical applications in healthcare, defence, the security domain or manufacturing industry, the reliability of the techniques will have to be unassailable and investments will have to be made in techniques that counteract that misuse. Consider the example of deepfake videos, in which someone's appearance and voice are simulated ('cloned') with the aim of misleading people and undermining the public debate.

6.2 NEDERLANDSE AI VOOR HET NEDERLANDS (NAIN)

Several public and private parties in the Netherlands see opportunities for Dutch speech models and AI applications based on them. They want to get down to work but are unable to improve the language models because – as individual organisations – they do not have enough knowledge, training data or specialised hardware, or because the algorithms still need further development. Moreover, generic language models have also not proved accurate enough to be used in specific cases, requiring a great deal of customisation. It is not feasible for individual public parties to develop this customisation themselves, and not profitable for private parties to improve both the generic and the specific language models for smaller parties. This is referred to as 'market failure': the technologies required cannot be profitably supplied by the Dutch market, whereas being dependent on the large overseas parties is deemed undesirable.

That is why the consortium Nederlandse AI voor het Nederlands (*Dutch AI for the Dutch language*, NAIN) was set up. NAIN takes speech technology in Dutch to the next level, where the models and techniques developed are sovereign, inclusive, diverse and transparent. The following points are being tackled to realise this: further development of privacy-enhancing technologies to allow privacy-sensitive language and speech data to be shared; establishing and tightening up the legal and ethical frameworks for using NLP and ASR; joining in with or developing infrastructures for unlocking the next generation of Dutch language and speech technology.

NAIN solves the problem of the low quality of language and speech technology for Dutch, which cannot be used properly (if at all) in its current state in many sectors. The consortium is unique because of its size in terms of the many sectors involved and the types of institutions (public, private, start-ups, centres of expertise and educational institutions, and public and private organisations in Flanders). Sectors such as health, new media, education, the commercial sector and security are represented in NAIN. Alignment has also been sought with the Spraak Coalition. Joint efforts are being made to bring together current initiatives (avoiding duplication of work, using project funds efficiently), building up knowledge about language and speech technology, sovereignty (more control and independence from large foreign commercial parties) and inclusivity (language and speech technology that performs better with dialects, accents and street slang). Ultimately, this should lead to high-quality models of both generic Dutch and several specific variants, improved algorithms (incorporated in software), licensing models or other forms of dissemination of the language models developed, and legal and ethical frameworks for using language and speech technology in the Netherlands.

Applications such as automated processing of phone taps, reports, questioning, reports, emergency calls and much more are now possible. One prospective application in the media sector, for example, is that speech technology can be used for automated subtitling of films. In the healthcare sector, the possibility is emerging of care providers verbally reporting what they do as they work, with that then being automatically processed and stored for administrative reporting and accountability. Within the NL AIC, there is close cooperation with the other sector-specific working groups on cross-sectoral applications.

In the landscaping of November 2021 and with the support of the South Holland AI hub, the NAIN consortium presented the current state of language and speech technology in the Netherlands and Flanders. Based on that landscape map, work can continue over the next five years on developing state-of-the-art sovereign language and speech technology for Dutch that is inclusive, diverse, transparent and explainable, and to which domain-specific extensions can be coupled. The end results of this project will be usable throughout Dutch society.

CONTACT US

From our role as a working group, we ensure a human centric approach where AI applications are developed and deployed effectively, securely and responsibly. We are looking ahead. We make connections based on the content, make sure the lessons learned are recorded and ensure upscaling where possible. We drive investment, support promising projects and make sure they have end users. We put the general public centre stage, create support for innovative solutions and keep security professionals involved at both the administrative and operational levels.

Your contribution is much needed. Which of the themes described here are you already working on? Which new themes have piqued your curiosity? Where would you like to get involved? What projects would you like to get started that will help achieve the aims and goals in this document, and who would you like to work with?

We want to hear what you have to say, so please contact us! We will be happy to help you think things through, such as project development, building a consortium or options for funding.

E-mail: vredeveiligheidenrecht@nlaic.com

PARTICIPANTS

More than a hundred representatives of the commercial sector, centres of expertise, social organisations and governmental authorities have worked on this document. Our thanks to the core group in particular.

CORE GROUP

Avans Hogeschool
Ben Kokkeler

Centraal Orgaan Opvang Asielzoekers
Sjef van Grinsven

Centilien
Gerard Kanters

Data Science District
Kai Lemkes

Dienst Justitiele Inrichtingen
Ramona Apostel

Dynaxion
Cor Datema

Erasmus Universiteit
Klaus Heine

Faculty of Impact
Frans Nauta

Haagse Hogeschool
Liduine Bremer
Elif Kiesow Cortez

Heijnen Consulting/Samen Veiliger
Alexander Heijnen

Hogeschool Leiden
Jos Griffioen
Hans Henseler

Jheronimus Academy of Data Science (JADS)
Peter de Kock (Pandora Intelligence)
Liesbeth Leijssen

Justitiële Informatiedienst
Tom Schepers

Ministerie van Justitie en Veiligheid
Ron Hanoeman
Bas ter Luun
Michel van Leeuwen
Caspar Heetman
Jitske Wuite

NFI
Lisanne van Dijk
Erwin van Eijk

Openbaar Ministerie (OM)
Tjistke Visser

Politie
Theo van der Plas
Bas Testerink

Radboud Universiteit
Frederik Zuiderveen Borgesius

Reclassering Nederland
René Poort

Researchable
Eduard van Pagee

Rijksuniversiteit Groningen
Bart Verheij

Saxion Hogeschool
Remco Spithoven

Security Delta HSD
Marlou Snelders
Joris den Bruinen

Sustainable Rescue
Paul Fockens

TNO
Saskia Lensink
Joachim de Greeff
Eelko Steenhuis

TU Delft
Marlou Smulders
Eveline Vreede

Universiteit van Amsterdam
Marcel Worring (Nationaal Politielab AI/ICAI)

Universiteit Leiden
Bram Klievink

Universiteit Utrecht
Floris Bex (Nationaal Politielab AI/ICAI)

Zuyd Hogeschool
Mark Liedekerken

ZiuZ
Jos Flury

AN INITIATIVE OF

HSD securitydelta.nl

Ministerie van Justitie en Veiligheid

NL AI Coalition

dutch digital delta connect and create

The appendix contains relevant framework information about using AI from the parliamentary letter on the progress of algorithms and artificial intelligence, dated June 2021. See other document.

Editing
Marlou Snelders (Security Delta HSD, Netherlands AI Coalition)
Martin Bobeldijk (Turnaround Communicatie)

Contact:
Email — vredeveiligheidenrecht@nlaic.com
Website — nlaic.com

January 2022