Shaping the Future of Austria with Robotics and Artificial Intelligence

White Paper by the Austrian Council on Robotics and Artificial Intelligence
Shaping the Future of Austria with Robotics and Artificial Intelligence

White Paper by the Austrian Council on Robotics and Artificial Intelligence

Vienna, November 2018
Preamble

The concept of the Austrian Council on Robotics and Artificial Intelligence was first presented at the European Forum Alpbach 2017 by the Austrian Federal Ministry of Transport, Innovation and Technology and was constituted on 24 October 2017. At the time the White Paper was adopted, the Council had nine members:

Sabine Theresia Köszegi, Chairwoman of the Council

- Professor of Labour Science and Organization at the Institute of Management Sciences of TU Wien
- Head of the MBA Program for Entrepreneurship and Innovation at TU Wien
- Head of the Doctoral College "Trust in Robots"
- Member of the High-Level Expert Group on Artificial Intelligence of the European Commission

Matthias Scheutz, Deputy Chairman of the Council

- Professor of Cognitive and Computer Science at the School of Engineering, Tufts University, Massachusetts, USA
- Director of the Human-Robot Interaction Laboratory at Tufts University, Tufts School of Engineering (Massachusetts, USA)
- Member of the AIWS Standards and Practice Committee of the Michael Dukakis Institutes for Leadership and Innovation of the Boston Global Forum
- Member of the Partnership for AI

Mark Coeckelbergh

- Professor of Media and Technology Philosophy at the Institute of Philosophy at the University of Vienna
- Professor of Technology and Social Responsibility at De Montfort University, Leicester, UK
- President of the International Society for Philosophy and Technology
- Member of the High-Level Expert Group on Artificial Intelligence of the European Commission

Corinna Engelhardt-Nowitzki

- Head of Department Industrial Engineering at the University of Applied Sciences Technikum Wien
Franz Höller
• Chief Technology Officer (CTO) KEBA AG, Linz
• Chief Technology Officer KEBA AG, Linz

Sylvia Kuba
• Head of "Process Digitalization" at the Chamber of Labour Vienna

Andreas Kugi
• Professor for complex dynamic systems at the Vienna University of Technology
• Board of the Institute of Automation and Control Technology (ACIN) and
• Head of the Christian Doppler Laboratory for Model-Based Process Control in the Steel Industry
• Co-Head of the Center for Vision, Automation & Control am Austrian Institute of Technology (AIT)
• Vice President of the Austrian Association for Electrical Engineering (OVE)
• Full member of the Austrian Academy of Sciences
• Member of the German Academy of Science and Engineering (acatech)
• Member of the Expert Panel of the Joint Science Conference (GWK) for the Excellence Strategy for the Promotion of Top-Level University Research in Germany

Martina Mara
• Professor for Robot Psychology at the Linz Institute of Technology (LIT) at Johannes Kepler University Linz

Erich Schweighofer
• Professor at the University of Vienna—teaches and conducts research in the fields of Legal Informatics, International and European Law
• Head of the Centre for Computers and Law
This White Paper is the Council’s first written opinion. It is based on the technical status quo of robotics and AI at the constitutional period and defines the main fields of action and framework conditions for Austria from the Council’s point of view.

In accordance with its mandate, the Council is an advisory body set up by the BMVIT whose recommendations are publicly available to all stakeholders in politics, society, business and science. The Council seeks cooperation with other relevant bodies such as the Bioethics Commission, the Council for Autonomous Driving and the Advisory Council of the Digitisation Agency.

Mandate of the Austrian Council on Robotics and Artificial Intelligence

1. The Austrian Council on Robotics and Artificial Intelligence identifies and discusses current and future chances, risks and challenges arising from the use of robots and autonomous systems (RAS) and artificial intelligence (AI) from a technological, economic, social and legal perspective.

2. The Austrian Council on Robotics and Artificial Intelligence thereby creates a strategic framework for BMVIT’s own RAS and AI activities and drafts strategic guidelines. These guiding principles include, inter alia, opinions and recommendations on
   I. creating economic framework conditions which promote innovation and technology and which ensure that the potential of RAS and AI are fully leveraged to ensure the competitiveness of Austrian industry;
   II. creating a legal framework to ensure safe use of RAS and AI for individuals and society as a whole, in compliance with the legal framework of the European Union;
   III. developing measures to identify, mitigate or prevent potential danger or harm to people and society caused by RAS and AI at an early stage; and
   IV. planning public activities to inform the public on RAS and AI and to responsibly address society’s fears and concerns.

3. The Austrian Council on Robotics and Artificial Intelligence regularly substantiates the strategic guidelines in the form of recommendations and implementation proposals to the Federal Minister for Transport, Innovation and Technology.
Executive Summary

The use of robotics and artificial intelligence (AI) in various areas of our lives will lead to fundamental changes in society. These new technologies have the potential to solve the great challenges of our times in many areas of application: dangerous, monotonous, unhealthy or strenuous activities can be performed by robotic systems; AI can contribute to better early diagnosis and treatment of diseases, and robots can support the autonomy and quality of life of the elderly and/or people dependent on care. Robotics and AI could help to sustain the international competitiveness of Austrian companies and thus create and secure jobs in the long term.

While discussing these chances we must, however, not forget about the ethical and social challenges that remain to be solved. These new technologies will not only entail changes and new requirements in the working environment: they will also raise ethical, legal and social questions that must be discussed and resolved by Austrian decision-makers from every area of society.

We have identified the following cornerstones for an Austrian robotics and AI strategy

1. Smart Governance
   All Austrians should benefit from robotics and AI technologies. We are convinced that a broad participation of all stakeholders—in particular of citizens themselves—in the process of defining this strategy is needed to increase the acceptance of new technologies. This can only be achieved if people's needs and fears are taken into account.

2. Smart Innovation
   We need a targeted research, development and investment policy to leverage the potential of robotics and AI technologies in all areas of application and thus tap into new markets and discover new fields of application. In doing so, the specific structure of Austrian economy, characterised by small and medium-sized companies, has to be taken into consideration. This will strengthen the economy and secure jobs and prosperity in Austria.

3. Smart Regulation
   A stable and secure framework is needed to ensure the trust of economic players and the positive development of markets. The use of robotics and AI must guarantee the safety of people and comply with ethical standards, fundamental human rights and European values. Wherever existing norms and standards are not sufficient to ensure this, new European standards and norms must be developed without stifling socially beneficial innovation. We need new and creative ways to resolve the trade-offs between innovation and regulation.
In this white paper, we identify—based on the current state of research—four fields of action which we consider the top priorities for developing a smart strategy for robotics and AI:

- **Smart Governance**
  - Due to the complexity and pace of technological developments, we recommend an incremental strategy process ("strategizing") with institutionalized learning processes and feedback loops.
  - The strategy process should include continuous research and technology monitoring.
  - The Council recommends the broad involvement of stakeholders from all affected areas (e.g. research, technology development and production, business, education and training, health, safety, housing, mobility) in this strategy process.
  - The Council recommends integration with the strategic activities currently taking place on the European level (European AI Alliance and AI High Level Expert Group).
  - The Council recommends the wide-scale inclusion of Austrian citizens into the development of the Austrian Robotics and AI Strategy in order to inform and educate the public and to increase public acceptance. The Montreal Declaration on Responsible AI, for example, can be used as a reference.
Smart Innovation

- The Council recommends significant targeted public investment and measures to promote innovation in robotics and AI technologies and to develop infrastructure for the use of these technologies.

- At the same time, the Council recommends to introduce systematic training measures to ensure the availability of skilled professionals. In addition, extensive training programmes must be provided for the retraining of existing workforce.

- As major technical challenges are involved and the impact of these new technologies cannot yet be fully predicted, potential industries and application areas ("use cases") which permit rapid implementation and leveraging of robotics and AI potentials and enable a fast learning process are to be identified as quickly as possible.

- When selecting use cases, the following two criteria should be taken into consideration: framework conditions must be controllable, and a clear differentiation is needed between unproblematic and sensitive application areas. “Sandboxes” and “testbeds” are to be set up for sensitive and/or high-risk application areas, providing rapid learning cycles and knowledge transfer to stakeholders from all areas (research and development, business and politics).

- The experience gained from such use cases has to be continuously reviewed, so that after multiple feedback loops it can be used to lay down effective guidelines for government support measures and for development and innovation policy.

Smart Regulation

- The Council recommends reviewing the existing ethical and legal framework with regard to expected changes and, where appropriate, introducing new regulation and standards to ensure the safe use of robotics and AI for individuals and society. This must be done in close coordination with the current regulatory efforts of the European Commission.

- The Council recommends the establishment of appropriate certification, audit and compliance tools for robotics and AI technologies.
# Table of contents

1. Introduction 11

2. Ethics as a Guiding Principle for the Council on Robotics and Artificial Intelligence 15

3. Robotics and Artificial Intelligence – Definition and latest Research 17
   3.1 Definition 17
   3.2 State of Research 19

4. Social and Ethical Potentials and Challenges 23
   4.1. Compliance with the European Union’s Ethical Values and Principles 23
   4.2. Safety and Security of Technologies 24
   4.3. Human Oversight 25
   4.4. Consideration of Social and Ecological Consequences 26
   4.5. Initiating a Wide-Ranging Public Debate on Fundamental Questions in Connection with Robotics and AI 26

5. Fields of Action 28
   5.1. Technology, R&D and Economy 28
   5.2. Workplace and Qualification 31
   5.3. Law and Society 35
   5.4. Awareness Raising, Communication and Public Relations 40

6. Strategy Process and Governance 43

End Notes 44

Glossary 45

List of References 46

Online Resources 47
1. Introduction

Robotics and Artificial Intelligence (AI) are strategically highly relevant technologies and involve the risk of disruptive change. It is therefore crucial that we take action now in Austria and shape the conditions that determine the world in which we want to live in the future.

With the existing knowledge about the course of technological change processes, we can prepare ourselves well for these changes. We will be able to shape innovation and technology in a way that not only makes our lives more comfortable, but also helps us meet the great challenges of today’s world.

The developments in the fields of robotics and AI, for example, have the potential to improve the quality of life of many people with prevention and precise diagnostics for all population groups; to increase the international competitiveness of Austria and to improve the quality of life of the population.

The aim is to strengthen the position of companies and to enable sustainable production and agriculture on the basis of more efficient, environmentally and climate-friendly use of resources.

At the same time, technological progress also entails major challenges and risks. As a society, we can and must not leave the development of these important technologies to chance. Rather, we have to play an active part in shaping them.

In order to utilize the full potential of robotics and AI and at the same time ensure that all people can benefit from these technologies, we need the joint efforts of politics, economy and society. In this context, innovation is not to be understood exclusively as technological progress, but also as the interaction of social, institutional and organisational innovation.

Austrian policymakers are called upon to invest in research and technological development, to create appropriate conditions and set incentives that on the one hand promote innovation and technical progress and on the other hand ensure that these developments do not lead to or exacerbate social inequality or jeopardise human rights.

Effective measures are to be taken in the coming years in the areas of research and technology promotion programmes, the definition of an appropriate legal framework and concepts for information, qualification and integration of the population.

These must, of course, also be coordinated with ongoing strategic activities of the European Union.
The Austrian Council on Robotics and Artificial Intelligence was founded as an advisory body to support the development and implementation of such a robotics and AI strategy. It analyses the potentials, challenges, risks and effects of the use of robots and artificially intelligent systems. Its aim is to draft statements and recommendations for Austrian policymakers—proactively, if needed—which are practical and operational in real-life conditions.

The complexity of the issues and challenges at hand requires an interdisciplinary approach. Experts from various disciplines were therefore appointed to the committee—from robotics and automation, artificial intelligence, machine learning and computer science to law, social sciences, economics, psychology and philosophy of technology. Due to the expected far-reaching changes for the economy and the working world, representatives of the Chamber of Labour and the Federation of Austrian Industries are also represented in the Council on Robotics and Artificial Intelligence.

In order to guarantee responsible technology design and to sustain Austria’s competitive position, the Council is committed to the diversity of relevant perspectives and a holistic approach and interdisciplinarity: Council members’ diverse professional background enables them to analyse complex problems with an interdisciplinary, multi-perspective approach and formulating practical recommendations, simultaneously using specific competences.

These recommendations are not limited to the Federal Ministry of Transport, Innovation and Technology’s more narrowly defined agenda on technology policy, but are addressed to all ministries concerned and all fields of action of the Austrian Federal Government.

In addition, the Council endeavours to create synergies with existing Austrian, European and international bodies and interest groups from the field of new technologies and digitisation; include their expertise and to conduct a productive dialogue.

On 25 April 2018, the European Commission published a Communication on "Artificial Intelligence for Europe" (European Commission, 2018), announcing the development of a European Strategy on Artificial Intelligence, which would put the EU at the forefront of robotics and AI. The document outlines a set of actions for the coming years.
The objectives of the European AI Initiative are

- strengthening the EU’s technological and industrial performance
- encouraging the adoption of AI in all areas of economy
- preparing for the socio-economic changes associated with AI
- establishing an appropriate ethical and legal framework

In the view of the Council, these objectives of the European AI Initiative can also be adopted as objectives for an Austrian robotics and AI strategy. The Austrian Council on Robotics and Artificial Intelligence considers it its task to actively shape the exchange of information and cooperation with EU expert committees for Artificial Intelligence and to support the Austrian government in proactively participating in European robotics and AI policy.

The Council regards its work as an essential contribution to shaping a sustainable future in which technology and humans are not played off against each other, but are understood as complementary factors. They are outlined in the following section. Section 3 contains a more detailed definition of robotics and AI, illustrating their relevance and problems. In addition, it includes a brief overview of the state of research and prognoses with regard to the further development of research in the field of robotics and AI. In section 4, fundamental potentials and challenges in connection with robotics and AI are presented. Section 5 identifies the key areas where the Council on Robotics and AI sees the need to act and which therefore should be at the focus of an Austrian Strategy for Robotics and AI. By setting such priorities, the Council intends to help define key initiatives for Austria.
2. Ethics as a Guiding Principle for the Council on Robotics and Artificial Intelligence

The Austrian Council on Robotics and Artificial Intelligence follows accepted ethical principles, and recommends their application to national and international decision-makers.

This includes in particular the following ethical principles, values and rights:

| Human rights: enshrined in national and international charters of fundamental rights, human rights treaties and declarations, they cover human freedom and dignity, economic, cultural and social rights and the protection of privacy; |
| Justice and fairness, inclusion and solidarity, including the protection of vulnerable members of society; |
| Democracy, social and political participation; |
| Principle of non-discrimination; |
| Personal, social and shared responsibility; |
| Additional ethical and social values in Austria and Europe |

Due to its mission, the Council also needs to consider additional standards when drawing up recommendations and statements: In addition to the general safety and development standards recommended by standardisation authorities and international expert groups (e.g. "Institute of Electrical and Electronics Engineers"—IEEE), the principles of "Responsible Innovation" and "Ethically Aligned Design" of the international IEEE initiative must also be observed.

These initiatives call for a proactive approach to technology development that takes ethical and social considerations into account from the early stages of innovation and design and involves relevant stakeholders in the development process at an early stage.5

This process has to be participatory to ensure that society as a whole can benefit from these technological developments. Only a participatory approach can ensure that the benefit of society as a whole is not outweighed by particular interests.

In addition, the Council relies on the general principles of research ethics, with a special focus on the role of interdisciplinary and transdisciplinary cooperation in research and development. Cooperation between various disciplines and stakeholders from business, society and research are to be specifically promoted.
3. Robotics and Artificial Intelligence – Definition and latest Research

3.1. Definition

Robots are machines that perform tasks—to some extent or fully—autonomously. In the light of the increased cognitive abilities that these machines have acquired with the use of intelligent software, they are now often referred to as "autonomous cognitive systems". The cognitive aspects of such systems include automatic data processing, creating environmental models and target-oriented action plans, the decisions made and action taken based on these models, and the automatic improvement of system functions using machine learning.

Autonomous cognitive systems are for example collaborative robots in industry and services, automated vehicles and other autonomous transport systems, active exoskeletons, social robots or medical robots as well as cognitive assistance systems and autonomous software agents.

An essential component of autonomous cognitive systems is "Artificial Intelligence". For this reason, the terms "AI systems" or "cognitive systems" are often used synonymously. The European Commission's definition of AI (2018) illustrates that a clear distinction is hardly possible:

- Artificial intelligence (AI) refers to systems with an "intelligent" behaviour which analyse their environment and act with a certain degree of autonomy to achieve certain goals.
- AI-based systems can be exclusively software-based, operating in a virtual environment (e.g. language assistants, image analysis software, search engines, speech and facial recognition systems), but can also be embedded in hardware systems (e.g. modern robots, autonomous cars, drones or "Internet of Things" applications).
- We use AI every day to translate texts, create subtitles in videos or block unwanted emails. Many AI applications need data to function optimally. Once they are functioning smoothly, they can improve and automate decisions.

Being aware of the rapid pace of technological change, the Council generally refers to "autonomous cognitive systems" to include both robotics and AI systems. This terminology is open to further clarification on a case-by-case basis.
AI research can be roughly divided into knowledge-based symbolic and data-based sub-symbolic approaches, although hybrid approaches have been gaining ground recently. Knowledge-based systems such as Classical Cognitive Systems have explicit symbolic representations of available knowledge, which they utilise to make decisions.

This knowledge is partly provided as default, but the system is also able to gain new knowledge from the tasks it is performing (e.g. through verbal instructions, by reading texts or by observing events in their environment).

Currently, data-based systems are mainly developed at the sub-symbolic level, such as deep neural networks. They acquire implicit knowledge in a long “training phase” using machine learning methods, before they are ready to perform their designated tasks. This means that such systems depend on the availability of high-quality data that can be used to train them. Hybrid systems use the advantages of explicit knowledge representation (e.g. for natural language interaction of robots with humans) and data-based methods (e.g. for image recognition of autonomous vehicles).

It is generally expected that the recent successes of AI will continue and intensify in the future. This raises the fundamental question whether it is possible to replicate or even surpass human intelligence and cognition in machines—with all the ethical consequences that such a development would entail. However, a long-term prognosis is difficult, as (disruptive) innovation in this field is happening at unprecedented speed.
Analyses of errors, findings and lessons learned from previous prognoses on AI made between 1950 and 2012 show that

There is no convergence and only little correlation between expert prognoses, there is no indication that prognoses made by experts are more accurate than those of laypersons, and that there has been a strong trend towards long-term forecasts (15 to 25 years from the forecast date) on AI development, and that these predictions have been far too optimistic.

These findings should warn us of making ill-considered predictions about the future that could unnecessarily fuel fears and lead to unwarranted scepticism towards new technologies.

Although significant progress has been made in AI development, adaptive autonomous systems are still limited to very specific abilities, even if they can sometimes surpass humans in this narrow framework. In general, people are referred to as "intelligent" because they can comprehend new situations and solve problems. This form of "general intelligence" is not yet achievable with currently known technologies and algorithms for AI systems.

3.2. State of Research

In recent years, research in the field of robotics and AI has been fostered by

- the widespread availability of affordable powerful computer hardware (computing power, data storage),
- intelligent and networked sensors, sensor data fusion,
- new actuator technologies and control algorithms,
- global networks with high bandwidths and a multitude of data sources (social networks, online platforms, Internet of Things), as well as
- the development of new powerful algorithms

and has already led to significant scientific breakthroughs and demonstration systems.
The active participation of major technology companies such as IBM, Microsoft, Google, Facebook has contributed significantly to the spread and refining of AI in recent years, as these companies have been able to provide both the financial resources and the large amounts of data required for training machine learning algorithms.

In 2011, for example, IBM’s Watson was able to defeat the best human Jeopardy player for the first time, thereby proving that intelligent systems can conduct effective question-and-answer dialogues in natural language. Shortly thereafter, Google introduced Alpha-Go, a program that defeated the best Go player in the world in 2016, as well as the follow-up program Alpha-Zero in December 2017, which despite minimal requirements could learn to play chess within a few hours at the level of a world champion.

AI research is subdivided into various fields, forming largely independent scientific communities in areas like machine image or language processing, machine learning, logical reasoning, planning and testing, dialogue systems (e.g. the new Google Duplex System, which can solve simple everyday tasks independently in natural language dialogues), game theory, decision theory, knowledge representation, cognitive systems and integrated control architectures.

The methods used can be roughly divided into symbolic and sub-symbolic approaches:

Symbolic methods include both logic-based and stochastic approaches; while the subsymbolic methods are usually represented in the form of neural networks and more recently by so-called deep neural networks. The latter have been able to achieve astonishing successes, in particular in areas like image processing and in learning simple control rules (by "deep Q-learning"). Some of these AI systems are already so advanced that they are clearly superior to human performance in certain tasks in terms of speed and result quality.

While logic-based methods are frequently used in language comprehension, formal verification, semantic networks and general problem solving; stochastic methods are mostly applied in robotics, where signals are typically noisy. Therefore, absolute values are not an option—distributions have to be used instead.

In addition to AI, vast progress has also been made in robotics development, for example in humanoid robots or so-called lightweight robots. For example, the research company Boston Dynamics managed to develop a humanoid robot which can autonomously pick up heavy objects from the floor and sort them into a shelf. What is more, after doing a backflip, the robot can stand on both legs without falling over.
A rapid development is also taking place in classical industrial robotics, complementing the previous generation of rigid, highly specialized machines by flexible and adaptive robot and tool systems with specialized cognitive capabilities. This development is driven, among other things, by the increasing individualisation of mass-marketed products and the associated increasing demand for flexibility in production.

There is already a large number of so-called lightweight robots available on the market that are limited in their payload, but have been designed to interact safely in direct contact with humans and can be operated easily and intuitively.

Also, research in recent years has made autonomous drones far more agile, with vastly improved manoeuvrability and countless additional functions. For example, state-of-the-art technology allows the targeted use of stalls to navigate drones through narrow openings.

In classical industrial robotics, a rapid development is taking place: the previous generation of rigid, highly specialized machines is now complemented by flexible and adaptive robot and tool systems with specialized cognitive capabilities.

The integration of robotics and AI can best be illustrated by the development of autonomous vehicles, which combine modern control technology with machine perception and advanced planning, simulation and decision algorithms in order to be able to navigate autonomously without collisions in a highly dynamic environment, such as innercity road traffic.

Other potential application areas range from real-time language translation to diagnostics based on global databases, that enables early detection of diseases; and from optimisation and customisation (e.g. in buildings), maintenance and resources to individual learning. AI and robot technology are already increasingly used in digitally networked production and the social sphere e.g. as assistance robots in hospitals or surveillance robots in shopping centres.

In addition, robots are also used in households: from vacuum cleaner robots to social robots such as Jibo, which can—similarly to other natural language systems like Amazon’s Alexa—complete smaller tasks (such as display photo albums, search online, schedule appointments etc.).

New and more powerful variants of these “intelligent cognitive assistants” are to be expected in the near future.
4. Social and Ethical Potentials and Challenges

Robotics and AI not only have the potential to make our lives more comfortable, but also to help us solve the great challenges of modern society. A few examples to illustrate key developments: robotics and AI deployed in the health sector will help to significantly improve early detection and treatment of illnesses. Nursing assistance systems could support high quality care and thereby bolster the autonomy and quality of life of people in need of care.

Robotics and AI technologies could take over dangerous, monotonous, repetitive, physically unhealthy or strenuous activities from humans (such as inspection work in mines or clean-up work in radioactively contaminated areas).

Deploying robotics and AI could secure the international competitiveness of companies and thus maintain jobs in a high-wage country like Austria in the long term. More efficient, environmentally and climate friendly use of resources—for example in agriculture and industry—are the key to securing sustainable prosperity.

While discussing these opportunities we must not, however, forget about the ethical and social challenges that remain to be solved. The Council recommends that political decision-makers involved in the development of a Robotics and AI Strategy for Austria pay particular attention to the following principles in addition to the technological and economic aspects:

4.1. Compliance with the European Union’s Ethical Values and Principles

In developing and using robots and AI systems, general ethical principles must be observed. Therefore, it must be ensured that the deployment of robots and AI systems safeguard the principles of human rights, including data protection and privacy, as well as the related principles of human dignity, human freedom and autonomy.

A clear recommendation of the Council is to ensure that AI systems—as far as technically possible and as far as this contributes to the protection of users and affected persons—are designed transparently and that decisions taken by AI systems made comprehensible (“explainable AI”). Experts are currently discussing the implementation of an “Ethical Black Box”\(^\text{10}\) and alternatively the “Counterfactual Explanations” approach\(^\text{11}\).
In addition, more far-reaching values and general principles such as justice and fairness, diversity and inclusion, solidarity and protecting vulnerable people must also be taken into account in the course of developing and deploying new technologies.

4.2. Safety and Security of Technologies

It must be ensured that robotics and AI technologies are safe to use, and comply with general development standards. Key questions in this context are whether there is a need to adapt existing standards and whether new standards are needed.

The safety and security of new technologies includes both machine safety and IT security, as well as the integration of these aspects. Attention should also be paid to enabling a high subjective sense of safety in using robots and AI systems.
4.3. Human Oversight

Although robots and AI systems are becoming more and more intelligent, their moral status is still unclear. In practical applications, the question arises whether robots and autonomous AI systems can be moral agents (in a practical, not philosophical sense) both able and authorised to make ethical decisions on their own. This aspect is discussed, for example, in connection with autonomous vehicles, as they might be involved in situations where they have to make ethical decisions. Such decisions could range from breaking laws to avoid an accident (e.g. ignoring a stop sign in order to avoid a rear end collision) to deciding which life is more worth protecting in the event of an unavoidable collision (e.g. that of the passengers or of other road users).

In view of the increasing autonomy of such systems and the associated delegation of decisions to technical systems, it is necessary to ensure human oversight in every situation and to define a clear moral and legal framework which also clearly outlines the responsibilities of all parties involved (e.g. developers, manufacturers, operators, users, customers, etc.).

Important cornerstones were set out by the European Parliament’s resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules in the Robotics Area (2015/2103(INL) (report compiled by: Mady Delvaux) and with the preparatory study, European Civil Law Rules on Robotics by Nathalie Nevejans, whereby an EU-wide regulation of liability issues and an ethical code of conduct were called for.

It is important to note here that this Council recommendation should not be regarded as a carte blanche for the development of autonomous machines without any ethical regulatory mechanisms. On the contrary: the Council recommends that algorithms and computational precautions for ethical behaviour be directly integrated into machine architecture and machine design in order to minimize possible ethical violations and amoral behaviour in autonomous machines.
4.4. Consideration of Social and Ecological Consequences

Robotics and AI technologies must be aligned with human needs; respect human behaviour and human diversity and must not enforce any inhuman or inhumane adaptation of humans to technology. What is more, it must be ruled out that technologies lead to biased or discriminatory treatment of individuals.

It is crucial to assess potential economic and employment impact of new technologies at an early stage, and to develop technologies that support sustainable economic activity that is both economically and ecologically beneficial.

4.5. Initiating a Wide-Ranging Public Debate on Fundamental Questions in Connection with Robotics and AI

It must be ensured that political decision-makers already today reflect fundamental social questions in connection with robotics and AI which might not be relevant immediately, but which we will be confronted with in the medium term. Experts from various disciplines, interest groups and the general public are to be involved on a meta-level in a comprehensive and interactive discourse on fundamental issues.

This includes, for example, a fundamental debate on how AI technologies can influence our democratic system and how we can uphold democracy and the rule of law in the long term. We also need to discuss how work—paid and unpaid—is to be distributed between humans and machines in future. This also requires a new understanding of the concepts of work and labour. The latter is closely connected to challenges associated with people in need of care. Robotics and AI will undoubtedly play an important role in addressing these issues. However, we have to clarify first which roles we want to assign to these technologies.
Shaping the Future of Austria with Robotics and Artificial Intelligence
5. Fields of Action

Based on latest research, the Council on Robotics and Artificial Intelligence has prioritised several fields of action in a detailed discussion. These priorities are meant to serve as a basis for the development of an Austrian Strategy for Robotics and AI. We see an urgent need for action in the following four areas:

These four fields of action are closely inter-connected, therefore it is not always possible to draw a clear line between them.

5.1. Technology, R&D and Economy

An internationally competitive and "fit for the future" economy is the basis of Austria's past and future prosperity. This position can only be maintained if a framework is created that allow companies to succeed in global competition.

Tackling digitisation requires Austrian economic stakeholders to be willing to innovate and invest.

However, they will only do so if they can expect a good return on their investment. In a high-wage country like Austria, companies must occupy creative niches and build on their existing strengths to succeed on global markets. Areas in which both the academic field (knowledge generation) and the commercial field (knowledge utilisation) are already strong are to be consolidated and consistently developed with the help of robotics and AI. To this end, both the potential of small and medium-sized enterprises (SMEs) and large corporations must be equally accounted for by political decision makers.

The foundations of prosperous development must be provided, ranging from efficient infrastructure to the creation of business-friendly and innovation-promoting conditions. This also includes access to data as a business resource. Political stakeholders have to ensure a high degree of legal certainty for business activities, while protecting the human rights for individuals. All in all, these basic principles are key factors of competitiveness, which the Council believes should by all means be incorporated into all future measures.
Austria has various research institutions that focus on robotics and AI, ranging from basic research to the development of marketable prototypes. They cover a wide range of topics, from motion control of manipulators and mobile systems to the perception and understanding of environments; from human-machine interaction to machine learning and AI. Latest studies provide more detailed information

Although there are individuals offering outstanding expertise in their respective field, Austria lacks an overarching strategy with a clear vision.

In view of the limited size of the Austrian research landscape and the scarcity of available human and financial resources, it is essential to define clear priorities and focus on application areas which are relevant for Austria.

In the future, it is to be expected that robotics and AI technologies will be deployed in almost every area of life.

Digital assistance systems, which facilitate human decision-making by providing information that is generated from data according to a given set of rules, are the first step. In other areas, however, AI systems should be enabled to operate autonomously, i.e. to solve tasks independently, to make independent decisions within a given framework, and to react adequately to unforeseeable events.

The timeframe within which such developments are expected to be implemented depends on various factors and varies from application to application. In general, however, it is safe to assume that tasks requiring a high level of physical dexterity, flexibility and sensitivity in addition to cognitive abilities will be more difficult to solve than tasks that can be performed with purely software-based systems.

Significant growth rates are predicted in all robot application areas, with "social robots" in particular expected to achieve very high growth rates. However, these developments also entail some hazards. In addition to classical safety problems with robots, such as collision-free navigation or the danger of manipulation, social aspects must be taken into consideration too. Machine learning algorithms could potentially be dangerous, as it is
currently not clear which categories these algorithms can extract from data sets. For example, it is known that deep neural networks trained on image recognition of traffic signs for applications in autonomous vehicles are completely misled by even a slight modification of the image, for example by a sticker on a traffic sign, and can no longer recognize a stop sign as such.

The question of the extent to which the behavioural characteristics of autonomous robots must be formally specified and verified before they can be used in practice is being intensively discussed by a wide variety of interest groups and requires further research, but also a clear legal framework. At the end of the day, Austria and the EU have to formulate their legal positions in order to guarantee the safety of autonomous systems for their citizens.

For a high-tech country like Austria, that has highly specialised professionals, disruptive changes such as digitisation offer the chance to be at the forefront of key technologies. As a very small research and development community by global standards, however, Austria should not measure itself against the ambitions of big nations (China, USA, Japan, Germany...), but—as part of an international network—should occupy niches in research on these key technologies.

Therefore, the Council recommends targeted public investment of significant amounts, measures to foster innovation in robotics and AI technologies, and provision of the infrastructure needed for the use of these technologies. The question of using data collected by public authorities (e.g. health data made available for research purposes) must be discussed and an appropriate strategy has to be developed, which takes into account the interests of citizens' worthy of protection.

The Council recommends to define key application areas ("use cases"), relevant fields of technology and future priorities in R&D as soon as possible. In doing so, it is to be ensured that investments permit rapid implementation and realisation of the potentials of robotics and AI, and enable a rapid learning process.

When selecting use cases, the controllability of the framework conditions and risk must be taken into consideration. A distinction between unproblematic (no risk) and sensitive areas (high risks) of application should be made. “Sandboxes” and “testbeds” should be set up for sensitive, high-risk application areas and/or if the impact of these new technologies cannot yet be safely predicted.

The aim of these measures is to create fast learning processes and knowledge transfer to stakeholders from all backgrounds (research and development, business and politics).

The experience gained from such use cases has to be continuously reviewed, so that after multiple feedback loops it can be used to create effective guidelines for government funding, development and innovation policy.
5.2. Workplace and Qualification

Robots and AI systems are increasingly able to perform complex cognitive activities that previously only humans were capable of.

The ever-stronger intertwining of virtual and physical processes in all areas of life—humans, machines, organisations, businesses, society or the state—change life and work situations and require new ways of thinking, working practices, new ways to cooperate and additional skillsets.

First and foremost, this means basic digital literacy of the general public and extends to specific expertise in the development, use and optimisation of intelligent machines for a given purpose. In addition, innovation and creativity, logic and integrated thinking will become more relevant in contrast to pure factual knowledge. Empathy, social, intercultural and communication skills play an increasingly important role with regards to human-machine interfaces and the evaluation, selection and sharing of results provided by digital assistance systems.

In light of these factors, the Austrian Council on Robotics and Artificial Intelligence regards the working world of the future as a key field of action. The key question is how opportunities for the labour market can be leveraged and how the working conditions of employees, work organisation and income development can be addressed.

It is still not fully clear what effects robotics and AI will have on the future working world. Forecasts on the potential impact of of new technologies on the labour market vary widely, ranging from predicting enormous job losses to positive effects on employment. Very pessimistic expectations concerning labour markets can be found, for example, in Ford (2015), Brynjolfsson and McAfee (2011) or Frey and Osborne (2017). These authors predict extensive automation of routine tasks in almost all industries and expect almost every second to be job affected. Although they are frequently quoted in the media, these forecasts are highly controversial among experts because the forecasting models used do not distinguish between individual tasks and entire professions or job profiles.

Therefore, it is safe to assume that automation potentials tend to be overestimated. Rigidly structured tasks such as routine workflows are indeed easy to automate. In contrast, less structured tasks cannot be easily taken over by machine systems.

The latest IHS and OECD forecasts see 9% to 14% of jobs—in particular those involving simple activities that can be easily taken over by robotics and AI—at risk. In any case, it can be assumed that simple physical activities in particular, which can be easily completed by robotics and AI, will be taken over by technical systems, depending on the economic and legal framework and cultural acceptance.
Recent OECD forecasts see about 14% of all jobs at risk. The Institute for Advanced Studies (IHS, 2017) predicts that 9% of jobs in Austria will be automated. Other studies predict the creation of new jobs, particularly in connection with the development, optimisation and maintenance of new systems.

According to the latest report of the World Economic Forum (2018), corporations expect a steep increase in new tasks and fields of work as a result of digitisation. At the same time, it is clear that every second member of the workforce will need significant additional training and retraining within the next four years. The Council also sees an urgent need for education, training and retraining in Austria.

**Projected average need for retraining in enterprises commensurate to the number of employees, 2018–2022**

![Rationale for retraining needs](image)

A redistribution of paid work can already be observed today. However, a detailed examination reveals clear differences depending on the level of income and education. While in the past mainly simple manual tasks requiring low skill-levels were automated, now cognitive robotics and AI systems can also automate tasks that previously required a higher level of education and were therefore performed by so-called knowledge workers.

In certain sectors, such as banking or retail, automation is already very advanced on the level of administrative staff. Similar developments can be expected in other industries.
The economist David Autor (2015) assumes that paid work will not decrease overall, but a polarization effect might occur. Robotics and AI create jobs for well-educated people with a specific skillset: analytical abilities, abstraction, ability to plan and good judgment, orientation and problem-solving skills, creative intelligence and social skills such as empathy and self-reflection are required. On the other hand, there will also be a large number of tasks which include simple manual activities that cannot be automated, or for which automation is not economically viable.

Deploying low-skilled workers, who are enabled to perform more complex tasks by intelligent assistance systems, could also contribute to this polarisation effect.

This could lead to more precarious work (job insecurity) and increase social inequalities. Appropriate measures have to be taken to prevent increased job insecurity. On the other hand, digital assistance systems could help preserve the jobs of low-skilled workers, which are often considered to be at risk.

Overall, it is a challenging and important task to implement measures for the qualification of workers and social policies, which take into account the qualitative effects on working conditions, work organisation and forms of work as well as on job structures. In fact, the OECD (Nedelkoska & Quintini, 2018) sees a high probability of change in around 32% of all jobs in terms of work organisation.

It is important to identify risks - for example with regard to the physical or mental health of workers or safety at work. However, it is also necessary to understand how robotics and AI applications could improve working conditions. Therefore, best practice examples and research on people in the work process are of particular interest. Equally important is the integration of employees’ experience and knowledge into the design of new working environments.

The Council regards close monitoring of these developments as a key task in order to ensure that society a whole can benefit from these changes. Short, medium and long-term measures are to be developed in order to cope with the expected changes in the working world. Training and re-training measures are the first priority in the short term. Industry representatives stress that the prevailing lack of skilled workers is inhibiting growth.
From a technical point of view, professional education and training should focus in particular on mechanical engineering, mechatronics, automation technology, electrical engineering, mathematics, information technology and statistics. In addition, the high complexity and strong integration of machines requires interdisciplinary skills in systems engineering and project management, process analysis and design skills, and the ability to integrate new technologies into a given application scenario.

In view of the rapid pace of technological development, skills required for future jobs will be constantly changing. Therefore, the willingness to learn, flexibility, and the ability to assess unclear cause-effect scenarios and elasticities (heuristically or probabilistically) are required at all qualification levels.

It is clear that basic skills for using digital technologies, robotics and AI are required for all member of society. In addition, specialist knowledge on the use and shaping of these technologies is required. This starts with the basic MINT competences (competences in mathematics, computer science, natural sciences and technology) required and ranges from the methodological tools of technical scientific domains to practical knowledge and skills in technology areas such as data science, robotics, machine learning, etc.

In contrast to basic digital skills, these advanced technological skills will be required to varying levels. In addition, there will be an increasing need to strengthen human-centric skills such as interdisciplinary discourse, cooperation, willingness to learn and flexibility. With most areas of life and work being digitised, individuals and society as a whole have to learn to use new technologies.

Particular attention must be paid to gender-relevant aspects when developing a suitable education and training strategy. Initiatives that inspire girls and women in particular to study technical, mathematical and scientific subjects and to work in technical professions should be fostered and intensified.

Existing barriers - both implicit and explicit - to career development in the education system and in professional life are to be eliminated. Encouraging and integrating women into technical professions is regarded by the Council as an essential factor not only in averting the looming lack of skilled workforce, but also for ensuring the sustainable and responsible development of robotics and AI.

In view of the rapid technical change, an appropriate qualification strategy must also take into account the responsiveness of the education and training system, as well as opportu-
nities for support provided by active labour market policies. This also means that education and training should not be confined to formal institution. Other stakeholders in the education and training system—e.g. businesses and interest groups—must also be involved. Improved training for teachers with modern teaching and learning methods is also required, as they are the ones to guarantee the quality of future education and training.

Right now, it is also essential to conduct a detailed analysis of the medium- and long-term effects of technology use:

In addition to the fundamental questions mentioned above on shaping and distributing work in future, socio-political considerations must also be taken into consideration.

5.3. Law and Society

In addition to their use in production, manufacturing and other highly specialized work environments (international space stations, operating theatres, etc.), autonomous robots and AI systems are increasingly being used in social contexts such as nursing and health care, psychological counselling (online or over the phone), customer care and telemarketing.

Robots and AI technologies are also expected to play an increasingly important role in the leisure sector. Digital, artificially intelligent assistance systems (Amazon’s Alexa, Google Assistant etc.) as well as cleaning or mowing robots are already showing considerable user numbers in the private customer segment. So-called “personal robots” or “companion robots”, which have a mobile physical shape and can provide simple assistance and entertainment services (e.g. taking photos, playing games), are not yet widespread. However, they are expected to become more widely available and more affordable in the coming years.

Unlike industrial robots, which perform clearly defined tasks in highly specialized work areas, the functionality of these social robots is for the most part not restricted to a specific environment or task.
In addition, their appearance is often similar to that of humans or animals and they are equipped with abilities for natural language communication and for simulating human emotions or human character traits. Anthropomorphic design features such as faces, eyes and arms are used to transmit non-verbal signals known from interpersonal communication to the robot, thus promoting a "social" connection with the machine.

In fact, relevant social science literature suggests that people—depending on personal and situational factors—tend to perceive intentionality or other human characteristics in robots and AI systems, identify with them, sympathize with them, show affection and trust, or even lie for them. These effects occur particularly with humanoid robots, but also sometimes with non-humanoid robots (e.g. vacuum cleaner robots). At the same time, numerous empirical findings indicate that highly humanoid robots often trigger strong adverse reactions and that society is widely sceptical about robots that assume social communicative functions perceived as "typically human".

The Council suggests to monitor research on the acceptance of different "social", "emotional" or "humanoid" robots and suggests to take state-of-the-art findings for policies into consideration. The Council also points out the need for further research into the psychosocial, social and societal implications of such "humanoid" robots.

Apart from the exemplary developments in social robotics, questions of human-machine interaction and technology acceptance are also increasingly relevant in other contexts.
The current trend towards collaborative robotics, the aim of which is close cooperation between humans and machines in the workplace, or the expected deployment of autonomous robots in public spaces (e.g. autonomous cars, transport robots, drones) leads, among others things, to the phenomenon of very different people, sometimes with little affinity for technology, increasingly coming into contact and interacting with robots. A broader definition of the concept of the "social robot" seems useful for the future. It will be relevant wherever people without special qualifications are confronted with autonomous and intelligent machines.

From a social perspective, the increasing use of robots and AI systems is associated with both opportunities and risks.

Opportunities could lie, among other things, in the use of such systems as supporting tools, e.g. to take over mechanical repetitive activities (e.g. cleaning), to help with strenuous physical work (e.g. by exoskeletons), to enable people with disabilities or the elderly to lead an independent life (e.g. with driverless cars or service robots), or to counteract the detrimental effects of social isolation (e.g. by telepresence robots).

Risks lie, among other things, in a perceived loss of control, in uncertainties in dealing with autonomous machines, in new forms of psychological stress in highly automated work environments, in one-sided emotional ties to machines or in the threat to self-esteem when robots and AI systems take over competencies from humans. As intelligent agents also simulate social-communicative behaviour and users build some kind of emotional relationship with the machine, questions of susceptibility to emotional manipulation or a possible loss of interest in actual interpersonal relationships arise.

It must be ensured that decisions made with the help of algorithmic processes and AI systems do not discriminate against anyone.

In addition, risks are reflected in reservations and fears that exist in the public towards autonomous technologies. In addition to fears of "being replaced" and, in particular, the substitution of socio-emotional ("human") competences, scepticism also prevails in society due to the fear of dependence on machines, decision-making by machines and the associated dominance of algorithmic classification schemes ("fear of being categorized"). Today, AI is increasingly used in so-called algorithmic prognosis and decision-making processes (e.g. predictive analytics). This means that large amounts of (historical) data are subjected to machine-aided statistical analysis and learning processes in order to identify possible correlations that help to categorise current events or predict future events.
These processes have enormous economic potential, as they can be used to forecast the future purchasing behaviour of customers and to adjust inventories accordingly, or to precisely predict and plan the maintenance and replacement of machine parts. This method provides important opportunities for medical applications, for example for diagnosing diseases such as certain types of cancer. The expectation in this field is to have earlier and more accurate prognoses.

Whenever predictive analytics methods are used to predict individual behaviour or characteristics of people and these forecasts are used as a basis for decision-making, ethical questions might arise. In the United States, these methods are used to predict the probability of recidivism among criminal offenders. A considerable number of companies already use AI systems in HR management to analyse video-recorded job interviews and biographical information of job candidates to draw conclusions on the personality and establish whether they are fit for a certain position. In addition, such predictive analysis methods are used to decide whether to grant insurance or consumer credits. There are several reasons why this is problematic: first, the accuracy of this prognosis strongly depends on the quality of the training data. Second, the weighting of individual factors used in the prediction is often not (or no longer) comprehensible.

One of the risks is that existing stereotypes and prejudices against certain groups will be reproduced in the data used by such technologies, leading to unfair and discriminatory decisions. Therefore, people's fears of being confronted with technologies which they do not understand or their fears of becoming "transparent" due to tracking of personal data by robots and AI systems must be taken seriously.

The Council on Robotics and Artificial Intelligence urges stakeholders to prevent or mitigate the risks listed above by taking appropriate measures. It is also necessary to consider existing and new developments in robotics and AI with regard to psychological, social and socio-cultural effects for Austrians and to assess the associated opportunities and risks. To be able to do so, the Council on Robotics and Artificial Intelligence sees a need for further research.

In addition, there is also a clear need for legal regulation and the need to proactively incorporate appropriate proposals into European legislation. The Council therefore recommends that the existing ethical and legal framework be reviewed in the light of upcoming changes and, where appropriate, add new rules and standards to ensure the safe use of robotics and AI for individuals and society.
The key problems are summarised below (cf. Borges, 2018):

It is generally assumed that the use of robotics and AI reduces the probability of errors and potential damage, e.g. in road traffic\(^{24}\). However, this must not lead to the neglect of ethical questions which rarely occur in practice, but are nevertheless of great significance (such as the "dilemma situations" already mentioned above). This is a matter of maintaining the high ethical standards currently in place in Austria. Liability regulations must therefore be adapted to these new challenges. This does not mean that the principles of tort law are to be changed, but rather a modernisation of the rules and standards for the use of robotics and AI. Even if a robot acts autonomously, a person must assume liability for illegal behaviour. As of now, this person could be the holder, the owner, the user or the manufacturer. Under current law, the breach of law can only stem from the behaviour of these persons, but not from an error made by an autonomous system (liability cannot be transferred to autonomous systems).

It may therefore be necessary to adapt existing regulations (legislation such as the "Machinery Directive" or technical standards) with regard to responsibilities and duties. Stricter maintenance obligations could be introduced or lawmakers may rely more strongly on product hazard and product liability regulations. Decisions by autonomous systems need to be checked for compliance with legal and ethical standards. The Council recommends the development of appropriate certification and auditing/compliance tools for robotics and AI technologies.

According to the General Data Protection Regulation (GDPR), no person may be subject to an automated decision which has legal or other significant consequences for them, subject to certain exceptions (conclusion or performance of a contract, consent or legal regulation). However, the right to explanation and transparency does not include the obligation to disclose the algorithm that was used.

Questions on appropriate tools and measures for ensuring the transparency of AI systems should be discussed at the European level.

Autonomous systems produce a vast amount of data that are of interest to manufacturers, owners, drivers and third parties. Regarding data usage rights, additional research and regulation is needed, for example, regarding personal data, where the principle of informational self-determination pursuant to the GDPR applies. Data may only be processed with consent, in fulfilment of a contract, on the basis of legal provisions or in case of a legitimate interests. Transparency must be ensured in any case (additional data processing is possible in compliance with "Privacy by Design and Default").

It is essential to maintain the high ethical standards in Austria by adapting the legal framework with regard to liability, responsibility and duties.
Finally, the use of autonomous systems changes the standard of due diligence under criminal law. User’s duty of care will be alleviated, while manufacturers and holders have to assume a higher level of liability. From the Council’s point of view, more precise regulation is needed. For example, the risk that in certain situations no one would be criminally responsible could be tackled with the concept of shared responsibility.

Although not a priority at present, the Council believes that the question of the legal personality of robots should be further pursued. However, for now it is more important to formulate rules and standards for the use of robots in order to determine liability issues.

5.4. Awareness Raising, Communication and Public Relations

Robotics and AI are topics that people have been concerning themselves with for a long time. Science fiction has also contributed help to achieve a high level of visibility. Due to the increasing performance of cutting-edge computer technology, which in connection with digital high-speed networks can process huge amounts of data in a very short time, a previously rather visionary picture of the future quickly becomes reality, either now or in the foreseeable future. However, very often only specialists understand the technical functionalities of AI. Many people do not have detailed knowledge of these technologies and are therefore not fully able to distinguish between reality and fiction. This lack of understanding, combined with science fiction information from the past and little objective information in the media, leads to doubts and fear among the general public. Thus it is not surprising that potential negative effects dominate discussion over chances and opportunities.

In order to be able to meaningfully shape the future with these technologies, it is urgently recommended to provide people with reliable and comprehensive information. This is the only way for citizens to assess the opportunities and risks of robotics and AI. Information should focus on meaningful future scenarios and on raising awareness for the social relevance of individuals’ deeds. We should strive for a way of life in which we as human beings can use their individual abilities and enjoy autonomy.

Citizens should be able to form their opinion on the risks and opportunities of robotics and AI. Therefore, Austrian citizens should be encouraged to participate in the creation of the country’s Robotics and AI Strategy to provide information and education, and thereby gain acceptance.

Robotics and AI create opportunities that allow people to spend more time applying human strengths and abilities. This includes creative activities, solving unstructured tasks, interaction with others in need of empathy and social intelligence, and much more.

From the Council’s point of view, it is indispensable to create such images from that contrast science fiction narratives in order generate public awareness and to enable meaningful dialogue.
The Council therefore recommends the broad participation of Austrian citizens in the development of the Austrian robotics and AI strategy to inform, educate and increase public acceptance. We recommend the "Montreal Declaration on Responsible AI" as reference.

Within the framework of citizen participation, measures must be developed as soon as possible which are suitable for bringing people into contact with robotics and AI technologies and informing them of their potential and risks. One of the first measures could be, for example, to invite Austrian kindergartens and schools to visit research facilities. Universities and research institutions must provide the resources for such visits. Creating a well-designed, up-to-date, interactive website at the Council on Robotics and Artificial Intelligence could be another measure for short-term implementation. Additional platforms are to be created to ensure that all stakeholders are heard and have the opportunity to participate from the beginning.
6. Strategy Process and Governance

The members of the Council on Robotics and Artificial Intelligence consider this white paper as a first impulse for the development of an appropriate robotics and AI strategy for Austria. It is therefore not a fully finished concept and requires ongoing adaptation to incorporate new findings and developments as well as to the results and outcomes of a broad public discourse with all relevant stakeholders. Due to the complexity of the problems and the dynamics in the fields of robotics and AI, long-term planning on the basis of complete information on the current and future situation is impossible from the Council’s point of view.

Against the background of these influences, it is not expedient to develop an all-encompassing strategy in the sense of a "big bang". Instead, the Council recommends incremental strategy development and implementation with institutionalised feedback loops, and continuous research and technology monitoring ("strategising").

With the objectives mentioned above (1) to promote technological and industrial performance, (2) to prepare for socio-economic changes associated with robotics and AI, and (3) to ensure an appropriate ethical and legal framework, the route is clear. In the interest of rapid success and feedback, short-term measures to leverage potentials ("quick wins") and urgently needed measures to prevent and mitigate negative consequences should be implemented as soon as possible.

The Council recommends a broad involvement of stakeholders from the all fields of action (e.g. research, technology development and production, business, education and qualification, health, safety, housing, mobility) and forming strong networks with the strategic activities currently taking place at the European level (European AI Alliance and the AI High Level Expert Group).

The fundamental structural changes addressed in this document, for example in RTI, in connection with the digital transformation of the Austrian economy, the school and education system, questions of social security, etc., are to be initiated promptly by initiating appropriate institutional processes.
End Notes

1 e.g. in the Charter of Fundamental Rights of the European Union (2012)

2 Principle of shared responsibility, see for example WEF (2015)

3 In values of the European Union in particular; the Communication of the European Commission (2018) contains helpful references and sources and announces a draft version of ethical guidelines for artificial intelligence by the end of 2018.


5 Cf. 5 Cf. by Schomberg (2011)

6 Cf. Owen, Bessant & Heintz (2013)

7 The literature provides numerous references to ethics in research, see for example the University of Vienna. (o.D.) or by Unger & Narimani (2014)

8 The term "robot" with its Czech origin in the word "robota" first appeared in almost 100 years ago in Karel Čapek's play R.U.R., in which artificially created, human-like machines revolt against their enslavement.

9 “[Autonomous] Cognitive Systems are systems that perceive, understand, learn and develop through individual or social interaction with their environment. [...] As a scientific discipline, Cognitive Systems seeks to provide an enabling technology for robotics and automation, natural language understanding, man-machine interaction and complex real-world systems (Europäische Union, 2002) und “[The objective is] to construct physically instantiated or embodied systems that can perceive, understand (the semantics of information conveyed through their perceptual input) and interact with their environment, and evolve in order to achieve human-like performance in activities requiring context-(situation and task) specific knowledge.” (US Department of Defense, 2015)

10 AI systems should store all information on algorithms and decision-making rules (on ethical questions) in an "ethics black box" in order to ensure that decisions can be "explained". See e.g. UNI Global Union (2017)

11 Cf. Wachter, Mittelstadt & Russell, 2017


14 Cf. 14 See Pichler et al. (2017); Čas, Rose & Schütter (2017)

15 Cf. 15 See Pichler et al. (2017)

16 Think, for example, of taxi drivers: with the help of a navigation system they no longer need to know where their destination is located to perform their job.

17 Cf. Urry (2017); Ho & MacDorman (2010); Mara & Appel (2015); Mori, MacDorman & Kageki (2012)


19 Cf. Scheutz (2011)

20 Cf. Perset et al. (2018)

21 Cf. 21 See The Economist (2018); Accenture, for example, uses HireVue, an AI program that analyzes video interviews with to assess whether they match the position; Johnson & Johnson uses the HiredScore AI system to evaluate candidates.

22 Cf. Bryson et al. (2017) and Perset et al. (2018)

23 Cf. Bryson et al. (2017)

24 Cf. 24 Eisenberger, Lachmayer & Eisenberger (2017)

25 Montreal Declaration on Responsible AI (https://www.declarationmontreal-iaresponsable.com)

26 For example, the Institute for Automation and Control Technology at TU Wien and the Department of Industrial Engineering at the Vienna University of Applied Sciences already offer robotics workshops for schools where children can experience robots and also program them. Similar efforts are being made by the PRIA Institute (https://pria.at/) and other institutions in Austria.

Note:
REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)
Glossary

**Actuators**
Actuators convert electrical signals into mechanical motion or other physical quantities.

**Algorithm**
Specification on how to solve a task.

**Autonomous driving**
Using a driverless transport system with autonomous behaviour.

**Drone**
Unmanned aircraft operated autonomously or controlled from the ground.

**Exoskeleton**
External support structure for an organism that can also be used to increase performance with the support of technology.

**Humanoid robot**
Robots with human-like shape or characteristics.

**Collaborative Robot**
Industrial robot that works together with people, i.e. in the same workspace.

**Sandbox**
An isolated area where actions have no effect on the external environment.

**Science Fiction**
A narrative that depicts scientific and technical speculations on a distant future.

**Software agents**
A computer program that is capable of certain (specialised) independent and self-dynamic (autonomous) behaviour.

**Strategising**
Includes all practical measures taken by individuals to develop long-term goals, plans and actions.

**Testbed**
Scientific platform for experiments and research.
List of References


Online Resources

Last accessed on 01.10.2018


Montreal Declaration on Responsible AI (o.D.). Calling from https://www.declarationmontrealiaresponsable.com


The Economist. (2018). Non-tech businesses are beginning to use artificial intelligence at scale (GrAlt expectations). Downloaded from https://www.economistuni-global-union-artificial-intelligence.pdf


Imprint (German)

Editor and Media Owner
Austrian Council on Robotics and Artificial Intelligence
1010 Vienna, Pestalozzigasse 4
www.acrai.at

Design
Haemmerle & Luger
www.haemmerle-luger.com

Illustration
Alexandar Savić