



Executive Summary

Challenges of artificial intelligence

Among the most promising developments of digitalisation is the possibility of using artificial intelligence (AI) in new ways, beyond the many successful applications that already exist, e.g. in the field of image recognition, medical diagnostics, language translation or mobility. As a basic technology, it has the potential to pervasively alter all areas of the economy, offering considerable innovation and growth potential.

While the mathematical foundations of artificial intelligence were laid decades ago, it was only the availability of enormous amounts of data and the rapid development of computing power that enabled the meaningful and marketable use of data with AI methods. There is currently no end in sight for technological developments. Many other areas of application for AI are already emerging. These include applications in drug development, real-time monitoring of machines and production processes, cyber security or medical research.

For Switzerland, it is important to exploit this as yet untapped potential by creating the best possible conditions to establish and consolidate Switzerland's position as one of the world's most innovative locations for AI research, development and applications. At the same time, steps need to be taken to fully understand the risks of AI and take corresponding measures in a timely fashion.

Good starting position

Research and development are an important means of maintaining competitiveness in the face of rapid technological change. This report shows that Switzerland is in a good starting position when it comes to research and development in the area of artificial intelligence. While it is true that the volume of AI research in Switzerland is slightly below average (measured in terms of the share of publications per inhabitant and the relative share of total publications worldwide), Swiss research institutions are among the top players in terms of leading publications. Switzerland is also among the most dynamic countries worldwide in the area of development and application, as measured by patent and start-up activities.

Despite this good starting position, it still faces considerable challenges caused by the speed of developments and has not yet tapped into the considerable application potential of artificial intelligence technologies. Research, innovation and education have a key role to play in addressing these challenges. It is therefore important to maintain and improve AI expertise to keep up with these developments.

Aim of this report

In addition to the rapid pace of new AI technologies, the main **challenges** relate mostly to the **new applications** envisaged thanks to artificial intelligence. The first challenge is to ensure that these new applications can actually be brought to fruition; after all, R&D efforts cannot be fully deployed if major obstacles hinder implementation of the resulting technologies. The second challenge is to craft the legislative framework in such a way that enables avoidance of the undesirable consequences of applying new technologies while still keeping up with technological progress. It is therefore also important to determine whether the consequences arising from a given application of AI require regulatory adjustments.

This report takes stock of the relevant regulatory conditions affecting the rollout of AI technology, highlights the specific challenges associated with various AI applications in policy areas affecting the whole of the Federal Administration. It also discusses certain adaptations that may be required at federal level.

The present report is the result of the work done by the Interdepartmental Working Group on Artificial Intelligence (IDAG AI), which the Federal Department of Economic Affairs, Education and Research (EAER) created in autumn 2018 at the request of the Federal Council. IDAG AI has itself formed various thematic working groups, which have also consulted numerous external experts as part of their work.

Structural features of artificial intelligence systems

There is no generally valid and accepted definition of artificial intelligence. In order to discuss the possible need for action on the part of the Confederation, the report therefore does not focus on the technology itself, but rather on the various ways in which this technology may be used today or in the foreseeable future and what implications these entail.

With this in mind, AI may be characterised (rather than defined) by various structural features that are typically found in the current applications of AI systems. This report identifies four key structural features, which appear in varying degrees depending on the area of application. AI systems are therefore capable of:

- (1) Evaluating the complexity and quantity of data in a way that would not be possible with other technologies currently available, i.e. algorithms that learn to automatically identify relevant statistical features in data;
- (2) Making predictions as an essential basis for (automated) decision-making;
- (3) Replicating abilities associated with human cognition and intelligence;
- (4) Acting largely autonomously on this basis.

Of course, some of these features can also be found in various forms in non-AI applications. However, it is only when combined with AI that entirely new applications become possible (e.g. facial recognition, fully automated vehicles).

Machine learning (ML) algorithms are the key universal technology behind the successful development of current AI systems. ML algorithms are extremely powerful **AI methods**, but their function is limited to recognising patterns in data and making simple predictions. In contrast, an **AI system** is capable of solving complex problems that were previously reserved for humans. This is done by dividing complex problems into a series of simple prediction tasks, each of which can be handled by a 'simple' ML algorithm.

Nowadays, such AI systems are each specifically developed for a particular application context. Unlike ML algorithms, which have increasingly become a universal technology, specific expertise on the exact areas of application needed to combine ML components of complex applications into a complete solution will therefore not be automated in the foreseeable future and still requires significant human input.

While AI has certainly enabled applications that mimic cognitive and perceptual human abilities and simulate certain aspects of intelligence, AI in its current form is far from comparable to human intelligence.

Specific artificial intelligence challenges

Artificial intelligence approaches are confronted with well-known statistical problems that become amplified with currently used methods. In particular, with some AI methods it is no longer **possible to comprehend** how a certain prediction or a certain result comes about or why an AI system has arrived at a certain answer to a given question. Moreover, AI-based applications are dependent on the quality of

the data and algorithms. For example, **systematic errors** in data or algorithms (e.g. hidden biases, such as when a population group is over- or under-represented in data) often cannot be detected given the size and complexity of the data used.

The business and science experts consulted in preparation of this report noted that these challenges were significant. While mostly technical in nature, in certain areas of application, these challenges can also create serious problems from a social or legal perspective. This would be the case, for instance, if groups of people are improperly and systematically discriminated against on the basis of AI decisions or if the result of given analysis cannot be explained in delicate contexts (e.g. the use of AI within the justice system).

Although technical solutions have already been found to mitigate these problems, there are certain drawbacks to using them. Current AI systems are optimised to recognise correlations on their own. An increase in traceability therefore comes at the expense of AI system performance, which may not be very desirable in some cases (e.g. in medical diagnostics).

The **challenges are very different depending on the area of application**. For example, while the traceability of an AI decision to recommend a given music title is not a serious matter, limited traceability of an AI system's assessment of the likelihood of a suspect or convicted person's committing a given offense becomes serious because the outcome of that automated decision may restrict the fundamental rights of the person concerned.

The draft revision of the Data Protection Act¹ takes these challenges into account and stipulates various obligations that must be met by the persons or institutions responsible for making AI-based, automated decisions. For example, a person affected by an automated decision must be informed of such a decision if it entails a legal consequence or significantly affects him or her. The data subject may also request that the decision be reviewed by a real person or that he or she be informed of the logic on which the decision is based.

The ability of AI systems to increasingly **act autonomously** is also putting the existing legal framework to the test. The Federal Council has already reviewed existing regulations and the current state of AI technology in this regard on several occasions. However, it has concluded that the existing legislative framework is adequate, particularly in the area of civil and criminal liability and private international law.

On the whole, the general legal framework is adequate

This analysis shows that the general legal framework in Switzerland is adequate when it comes to dealing with new AI applications and business models. The use of new and innovative technologies does not take place in a legal vacuum and must fully comply with applicable legislation. The relevant legal principles are usually formulated in a technology-neutral way so that they can also be applied to AI systems. The existing legal framework thus permits and limits the use of AI in principle. This also applies in particular to discrimination that may arise as a result of AI decisions. **Responsible use** is therefore determined by the value system underlying the legal norms and guaranteed by their observance within the legal system.

Thus, there is no need for fundamental adjustments to the legal framework. However, given technological dynamics, it cannot be ruled out that this assessment could change quickly.

Challenges in policy fields

Even if the general legal framework is fundamentally suitable from today's perspective to cover AI applications, new application possibilities may nevertheless call into question the existing **regulatory framework in specific policy areas**.

¹ Draft of the total revision of the Data Protection Act and data protection amendments, 15 September 2017, BBL 2017, 7206 ff.

Under its mandate, the IDAG AI has taken stock of the AI-related challenges relevant to the Confederation. In this regard, 17 topics were identified that are relevant from today's perspective and should be given priority.

The topics were handled under the responsibility of the relevant office. Since the challenges of using AI vary greatly depending on the topic, clarifications and consultations were needed to varying degrees. While some topics had already been sufficiently addressed by the offices, broadly representative project groups were created to handle other topics.

A total of seven larger interdepartmental working groups were set up and numerous external stakeholders and experts from science and industry were consulted. In accordance with the Federal Council's mandate, the considerations for clear and responsible use of artificial intelligence had to be taken into account.

Clarifications covered the following 17 topics:

1. International bodies and artificial intelligence
2. Clarification of Swiss interests in European AI activities (Digital Europe Programme)
3. Changes in working life
4. Artificial intelligence in industry and services
5. Artificial intelligence in education
6. Use of artificial intelligence in science and research
7. Artificial intelligence in cyber security and security policy
8. Artificial Intelligence, media & general public
9. Automated mobility and artificial intelligence
10. Artificial intelligence in healthcare
11. Artificial intelligence in finance
12. Artificial intelligence in agriculture
13. Energy, climate, environment and artificial intelligence
14. Artificial intelligence in government
15. Changes to the general legal framework with regard to artificial intelligence
16. Artificial intelligence in the justice system
17. Artificial intelligence, data and intellectual property law

Challenges largely addressed

In some cases, the developments in artificial intelligence pose major challenges in the various policy areas. The analyses show that there is still a great need for clarification and adaptation in many areas, but that the actors in Switzerland have already largely accepted the topic and responded accordingly. This applies in particular to education and research as well as to the economy. In these areas, a large number of measures have been initiated to fundamentally address the identified challenges. This also applies to the Federal Administration, which has already largely addressed the issue of AI.

In terms of *additional* need for action, this report recommends that clarification work be intensified in certain areas and stresses the urgency of clarifications in some cases. In addition, another new important area of action is improving the level of coordination, networking and monitoring in response to rapidly advancing technological developments. Given the fast pace of change, strategic guidelines should first be developed at federal level on the basis of this report. Artificial intelligence should not be viewed as an isolated technology, but rather as an essential component of the ongoing digitalisation of the economy

and society. The relevant sector-specific measures of federal departments and offices should therefore be taken into account in the Confederation's general digitalisation strategy ('Digital Switzerland' strategy).

Because technological advances come quickly, the topics considered in this report should only be perceived as a **snapshot**. This applies not only to the selection of topics, but also to the foreseeable need for adaptation of sectoral regulation.

Proven policymaking principles

This report confirms that the central premise that a technology-neutral legislative and regulatory approach has proven its worth, especially in the rapidly changing technological environment of AI, whose development can only be predicted by lawmakers to a limited extent. Where there is a need for regulation of specific uses of AI, the regulations should be formulated in such a way as to ensure that all technologies are addressed equally.

The Confederation therefore intends to continue to pursue a fundamentally technology-neutral policy, which largely refrains from promoting specific technologies and tries to avoid technology-specific regulations as far as possible. Such openness on the part of the state towards new technologies allows the potential of new ideas and innovations to be optimally exploited.

In the Federal Council's view, this approach is a key factor in Switzerland's success. Based on these proven policymaking principles, Switzerland therefore seeks to maintain its position as an appealing location for research, development and application of new technologies by maintaining legal certainty, efficient regulation and its solid reputation.