

# Final Report: Researching a new form of compliance test for upcoming technology legislation through a pilot event: The First University of St. Gallen Grand Challenge – The EU AI Act

*Thomas Burri, University of St. Gallen, 15 November 2023*

*This is the final report under the project referenced in the title (contract no. 8003535283), i.e. the last deliverable under this project.*

**Context:** The European Union is currently negotiating legislation on artificial intelligence.<sup>1</sup> It will likely adopt an A.I. Act this fall and declare it applicable from 2025, making it the first comprehensive legislation on A.I. worldwide. Meanwhile, A.I. and, most notably, large language models develop rapidly.

The E.U.'s A.I. Act will impose stringent regulation depending on the risk A.I. applications pose.<sup>2</sup> In the legislative process, the question of what the provisions of the Act mean in practice – whether it can be adequately implemented – hardly received any attention.<sup>3</sup>

**The project:** To explore the A.I. Act's feasibility and implications in practice, we organized a challenge which Armasuisse S+T supported by means of a research contract.<sup>4</sup> Challenges are typically used in the natural sciences to foster technological solutions to real-world problems.<sup>5</sup> Our challenge, in contrast, was to apply upcoming law to real-world technology, situating it in the social sciences. In the “First University of St. Gallen Grand Challenge: The EU A.I. Act 2023”, we presented existing A.I. applications from technology companies to teams and asked them to assess the applications under the new A.I. Act.<sup>6</sup>

We recruited twelve teams in March 2023, each comprising six professionals in computer science, law, and other social sciences. During the Grand Challenge, technology companies confronted each team at the event with seven A.I. applications

---

<sup>1</sup> See the latest draft of the EU A.I. Act from the European Parliament, P9\_TA(2023)0236, Artificial Intelligence Act, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))1 -- A9-0188/2023).

<sup>2</sup> Michael Veale and Frederik Zuiderveen Borgesius, ‘Demystifying the Draft EU Artificial Intelligence Act’, (2021) 22 Computer Law Review International (4) 97-112.

<sup>3</sup> Ben Green, ‘The Flaws of Policies Requiring Human Oversight of Government Algorithms’, (45) Computer Law & Security Review (July 2022) 105681.

<sup>4</sup> See also our lab experiments: Serhiy Kandul, Vincent Micheli, Juliane Beck, François Fleuret, Thomas Burri, Markus Christen, ‘Human control redressed: Comparing AI and human predictability in a real effort task’, (10) Computers in Human Behavior Reports (May 2023) 100290.

<sup>5</sup> E.g., Darpa's Robotics Challenge (2015) or Subterranean Challenge (2021). Hackathons, even in the law, also aim at technological solutions.

<sup>6</sup> The rulebook of our law challenge is available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4503312](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4503312); the version of the EU A.I. Act we used in the challenge and further information at [www.thegrandchallenge.eu](http://www.thegrandchallenge.eu), our videos at <https://www.youtube.com/@grand-challenge>.

from their portfolio, ranging from health care, manufacturing, transport, strategic advice, and telecommunication to robotics.<sup>7</sup> Teams interacted individually with each technology company to learn the details of the A.I. applications.

The teams' task was to assess each A.I. application in the light of the latest available draft of the A.I. Act. Teams had to look for answers to questions they defined themselves, such as: Does a specific piece of technology qualify as high-risk A.I. under the Act? Does the technology company comply with the Act's duties, e.g., to guarantee human oversight, manage the risks and data, establish transparency, and guard against cyber risks? How could compliance be improved? Each team delivered a comprehensive A.I. assessment report for use by the technology companies.

Armasuisse S+T did not only support the Grand Challenge with a research contract, but also organized an intense practice session, the "Boot Camp", which was integrated into the ARCHE 2023 week. This Boot Camp took place a week before the Grand Challenge Final in St. Gallen. It allowed the Grand Challenge Teams to gain some practical experience, before the competition actually started. We also carried through one of the A.I. applications from the Boot Camp to the Final. This aspect was of particular importance for the Teams, as it allowed them to work for almost a week on one A.I. assessment.

In aggregate, the assessments done by a dozen teams on numerous A.I. applications subjected the A.I. Act to an in-depth stress test. In the end, a jury of internationally renowned law practitioners awarded the prize of 100,000 CHF (approximately 100,000 USD) to the team whose A.I. assessment report it considered most accurate under the Act and most useful to the technology companies.

**Output and deliverables:** The main deliverable under the research contract with Armasuisse S+T was the Grand Challenge itself, i.e., that the event took place. Three research groups/companies with whom Armasuisse S+T collaborates closely, namely, Ascento, Gravis Robotics, and Swiss-Mile, benefitted directly from the A.I. assessment the Grand Challenge teams delivered to them. But beyond this most tangible output, the Grand Challenge also further delivered, both in terms of insights and results, as follows.

While, on substance, the teams' assessment reports identified legal shortcomings of specific A.I. applications, the reports, in aggregate, revealed that the A.I. Act was difficult to apply. The Categorization of an A.I. application as "high-risk" according to

---

<sup>7</sup> The A.I. applications were provided by [Ascento](#) (security robotics), [Deltia AI](#) (manufacturing), [Deutsche Telekom](#) (telecommunication), [Gopf](#) (strategic advice), [Gravis Robotics](#) (robotics in construction), [Ovom Care](#) (health care), [Mercedes-Benz](#) (transport), and [Swiss-Mile](#) (delivery robotics).

the specific use cases in Annex III of the Act proved controversial in more than one case. This is problematic as the Act attaches most duties to this category. Moreover, some of the duties, e.g., guaranteeing human oversight and managing risks, turned out to be too varied and vague to be applied effectively to A.I. applications, which are inevitably specialized. Consequently, technology companies enjoy a broader margin of discretion than the Act on a superficial reading indicates. Conversely, companies face legal uncertainty – whilst the authorities can issue steep fines for infractions.

The teams' A.I. assessment reports were useful to start-up and mature technology companies. They appreciated the legal assessments early in law and product development. Yet, the unease of technology companies with the A.I. Act was palpable, even though the Challenge clarified it to some extent. This heralds a more profound impact on innovation than E.U. legislators are ready to admit. The Act intends to increase legal certainty, but it currently sows *uncertainty* and chills A.I. development and innovation. The three to five years it will take to dispel the uncertainty coincide with the most crucial phase of A.I. development. In all likelihood, once the European economy is ready to rush for A.I. based on a firm legal framework, the global race will be over. The winners will be elsewhere.

As an alternative to the A.I. Act, a set of high-level principles aimed at best efforts could be proclaimed, such as those in article 4a of the draft A.I. Act. Such principles would be legally binding but not subject to hefty fines in case of breach. Less grand than a full-blown A.I. Act, principles would give technology companies, especially start-ups, two to three years to innovate and develop. The legislature, in turn, should use the period to adapt existing sectoral legislation to A.I. and adopt measures to prevent large language/foundational models from escaping sectoral law. This painstaking, meticulous work is indispensable. The A.I. Act does not do it – perhaps it even distracts from it.

On an organizational level, the Grand Challenge also provided a number of lessons, which we list here in the form of recommendations.

1. *Announce a few months before the law challenge.* We published our call for participation in January 2023 and scheduled the main event for July 2023. Six months was enough time to bring on board teams and technology companies. In the natural sciences, teams typically need more time to develop new technical solutions.

2. *Emphasize diversity.* A.I. reflects it, but the technology sector struggles with it. This is why we turned the spotlight on diversity. We announced it as a criterion (in terms of gender, age, and general background) on an equal footing with expertise and credibility to select teams. Consequently, the 30 teams which applied reflected a fully diverse

society. Moreover, the 12 teams we admitted not only hailed from South Africa, Singapore, the U.S., and all parts of Europe but also included highly qualified law and computer science professionals with nationalities from across the globe. We tried to include more teams from the global south, but some practical difficulties proved insurmountable. Yet, our challenge shows that lack of diversity can be overcome with appropriate effort.

3. *Synchronize with the legislative process.* As our challenge stress-tested the E.U.'s upcoming A.I. Act, it had to be in sync with the ongoing legislative process. We managed to have the challenge on the day the E.U.'s inter-institutional negotiations of the Act began. Owing to this fortunate timing, we had a well-developed, up-to-date draft of the Act and could feed substantive learnings from the stress test to the legislative process.

4. *Offer a monetary prize.* We put up 100'000 CHF (approx. the same in USD/EUR) for the winner. Since prizes are uncommon in the law, it helped spread the word, even to professional service providers and made a short timeline feasible. A challenge in the natural sciences may be viable without a prize,<sup>8</sup> but in the law, it is probably not (or not yet).

5. *Prepare for differences.* To make the challenge demanding, attractive, and insightful, we cooperated with companies from different sectors to provide A.I. applications, including health care, manufacturing, transport, strategic advice, telecommunication, and robotics.<sup>9</sup> The companies ranged from early start-ups interested in tailored and affordable advice on the new A.I. legislation to mature listed companies keen on marking their presence in A.I. and law. However, each had to be cajoled in various ways into making available technology,<sup>10</sup> e.g., by promises to keep the A.I. or the assessment confidential or have all participants sign a non-disclosure agreement (which took an enormous effort to implement).

6. *Offer teams enough time.* In our challenge, each team had 24 hours in the first round to assess four A.I. applications and submit a written report. The two best teams advanced to the final and had two hours for an oral assessment of another A.I. application. We released some information on A.I. applications to avoid overload and offered the one-day Boot Camp, together with Armasuisse S+T and the SDRC, to practice assessment a week before the main event. Twenty-four hours for the first round turned out to be sufficient. Two hours for the final proved too little, though. Hence, it

---

<sup>8</sup> See, for instance, the ETH Cybathlon, [www.cybathlon.ethz.ch/en](http://www.cybathlon.ethz.ch/en).

<sup>9</sup> The A.I. applications were provided by [Ascento](#) (security robotics), [Deltia AI](#) (manufacturing), [Deutsche Telekom](#) (telecommunication), [Gopf](#) (strategic advice), [Gravis Robotics](#) (robotics in construction), [Ovom Care](#) (health care), [Mercedes-Benz](#) (transport), and [Swiss-Mile](#) (delivery robotics).

<sup>10</sup> Some big tech companies, e.g., Microsoft and Nvidia, proved impossible to convince.

was no surprise that the Jury split the prize between the two teams. Furthermore, after the competitive part of the challenge was over, more time should have been allocated to building a community *between* teams.

7. *Specify the criteria.* The two criteria to determine the winner were: accuracy of the legal assessment of the A.I. and its usefulness for the technology companies. We specified them relatively late to account for the priorities of technology companies. Yet, teams would have appreciated earlier clarification, as the performance metrics in law are not evident.

8. *Secure funds early.* Our challenge did not fit easily into the funding landscape. Confidential A.I. was hard to sell, sponsorship conflicted with being a team member or A.I. provider, and the categories funding institutions applied did not fit. We count ourselves fortunate to have received funding from longstanding partners, first and foremost Armasuisse S+T and SDRC.

The press reported favourably on our challenge, including SRF Tagesschau, at 19.30h, and Neue Zürcher Zeitung.<sup>11</sup> All media reporting is linked on our webpage [www.thegrandchallenge.eu](http://www.thegrandchallenge.eu), where a video reviewing the Grand Challenge, many pictures and quotes are also available. Finally, perhaps the most significant achievement of the Grand Challenge is the publication of an article on it in *Nature Machine Intelligence*: “A Challenge for Law and Artificial Intelligence”<sup>12</sup>.

We are currently considering another iteration of the Grand Challenge. For now, we are reduced to waiting, though, for the final result of the negotiations of the EU AI Act from the Council, Commission, and the Parliament. Yet, other legislation could also benefit from competitive stress tests in the manner of our challenge. Regulation of supply chains, financial technology, climate change, etc., could be made more practicable. In any case, another Grand Challenge would not be doable earlier than January 2025. If we decide to give the Grand Challenge and the EU AI Act another try, Armasuisse S+T will be the first partner we will approach. Our partnership has been tremendously beneficial, and we are deeply thankful for the support.

---

<sup>11</sup> SRF Tagesschau, Swiss national prime time news, 19 July 2023, <https://www.srf.ch/play/tv/tagesschau/video/hsg-was-darf-kuenstliche-intelligenz-was-nicht?urn=urn:srf:video:73f91b72-fb62-4b3a-b928-da1a7e398a71>; Neue Zürcher Zeitung, 21 July 2023, <https://www.nzz.ch/technologie/kuenstliche-intelligenz/dieser-roboter-koennte-in-zukunft-ihre-pakete-liefern-welche-standards-muss-er-dafuer-erfuellen-ld.1748033>.

<sup>12</sup> [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4620764](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4620764).