swissmedic 4.0

Regulator in the Loop: Trust in ML-based Systems

RSNN Annual Workshop, June 11, 2024

Michael Renaudin

Schweizerisches Heilmittelinstitut
Institut suisse des produits thérapeutiques
Istituto svizzero per gli agenti terapeutici
Swiss Agency for Therapeutic Products

Hallerstrasse 7, 3012 Bern, Schweiz
www.swissmedic.ch

swissmedic 4.0

7

2020 - 2025

Swissmedic 4.0 is an innovation lab, an experimental field for change and innovation. The aim of the initiative is to promote interdisciplinary work and to design new digital business models. Unhindered by bureaucratic structures and processes, technological innovations and new forms of organisation and collaboration are to be evaluated.

“A modern regulatory agency office with a diverse team of professionals using AI technology to assess pharmaceutical products. The scene includes scientists and analysts working on computers displaying complex AI algorithms, molecular structures, and data visualizations. A large screen on the wall shows a real-time dashboard of pharmaceutical assessments. The office is well-lit with a high-tech, organized environment, featuring shelves with pharmaceutical samples and documents. The overall atmosphere is one of efficiency and cutting-edge technology.”

Automated Data Analysis:

AI can quickly and accurately analyze large volumes of data, leading to faster decision-making processes.

Risk Assessment:

AI models can assess risks more accurately and quickly by analyzing complex data and identifying potential problems early.

Potential of AI in the Regulatory Field

Monitoring and Compliance:

AI systems can continuously monitor and ensure compliance with regulations by detecting deviations in real-time.

Detection of Illegal Activities:

By analyzing online content, AI models can identify suspicious activities and products that human controllers might overlook.

Challenges in Implementing ML Systems

The introduction of ML-based tools into regulatory processes presents a significant challenge.

Contradicting principles:

These models are often perceived as "black boxes," which contradicts the principle of traceability in regulatory environments.

Competencies:

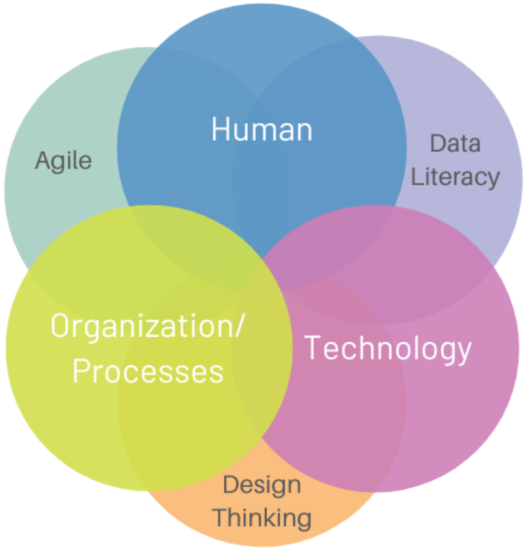
Regulators must understand not only the processes and data but also the functionality of ML models.

Building trust in these systems is therefore difficult.

Why Trust is Crucial

- Regulatory agencies operate in an environment where errors cannot be tolerated.
- Trust is crucial because regulators need to maintain complete control and traceability of decision-making processes.
- A system that regulators do not trust will not be used effectively and can lead to uncertainties and potentially serious misjudgments.
- Trust is built through transparency, traceability, and the continuous involvement of users in the development process.
- Comparison with human working relationships: Trust through openness and communication.

Our Framework



- We have developed a framework that supports the introduction and operation of ML-based systems in regulatory environments
- The central theme of our approach is trust. Trust in the technology and its results is essential, comparable to the trust that is indispensable in human working relationships.
- Our framework is based on two main components:
  - Technical Backbone: Performance metrics, transparency requirements, data management
  - Iterative “Show-and-Tell”: Agile methods, involvement of regulators, transparency, and trust

Use Case 1

/NightCrawler

- The goal of NightCrawler is to automate the search for illegal medical products on the internet.
- In the initial phase of the project, no suitable data was available. Therefore, we had to manually label cases, which was a time-consuming process.
- Over time, we collected over 50,000 manually labeled cases. These served as the basis for developing and training the ML model.
- The model was trained to identify and classify suspicious URLs. We placed great emphasis on the accuracy and relevance of the results.
- Through continuous testing and feedback from regulators, the model was continuously improved. A benchmark was established to ensure that the system found at least as many relevant URLs as human inspectors.

Use Case 2

/TRICIA

- The goal of TRICIA is the risk-based triage of medical device incidents.
- At the beginning of the project, we had approximately 10,000 entries available as the gold standard. These data were collected and assessed by various experts over four years.
- We analyzed the data and used unsupervised learning methods to identify patterns and text structures in the input data.
- The ML model was trained to assess incidents based on risk criteria.
- Two main KPIs were defined: the detection rate of high-risk cases and the F1-score multi-class evaluation.
- Through regular feedback from regulators and the adaptation of the model to new data and requirements, the system was continuously optimized.

NightCrawler Results

- The system significantly improved the efficiency of inspectors by searching large amounts of data and filtering out relevant cases.
- NightCrawler achieved an 80% hit rate in identifying relevant URLs
- NightCrawler demonstrates how an ML tool can be developed to meet the regulators' requirements through an agile and iterative development process.

TRICIA Results

- The system was able to standardize and accelerate the triage process, reducing manual workload.
- TRICIA achieved high accuracy in risk assessment and improved the consistency of assessments between different experts.

Trust as the Key to Success

Implementing ML systems in regulatory environments requires a careful and transparent approach.

Trust is the key to success in implementing ML systems in regulatory environments.

By building and strengthening trust through transparency, traceability, and continuous involvement of regulators, we can effectively utilize ML systems to fulfill regulatory tasks.

Trust in technology is as important as trust in human colleagues and is built through transparency and continuous communication.