REPORT CO-DEVELOPMENT PROCESS

SWISS DIGITAL INITIATIVE

CONTACT

Niniane Paeffgen
Managing Director
Swiss Digital Initiative
niniane@sdi-foundation.org

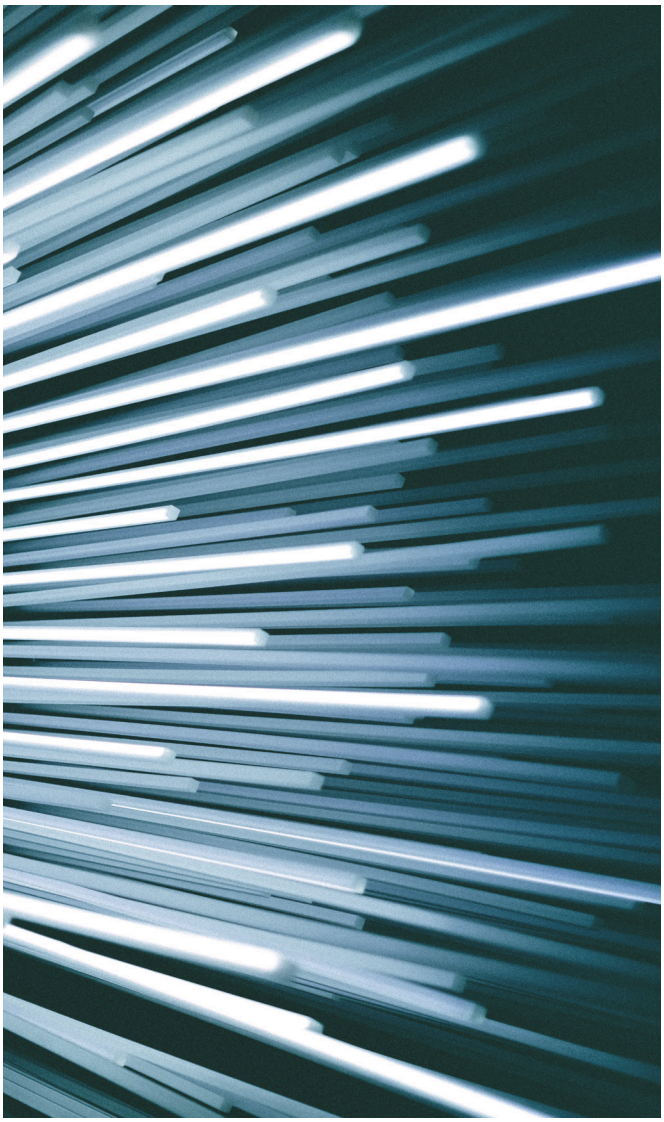
Dr Johan Rochel
Co-founder
ethix - lab for innovation ethics
rochel@ethix.ch

EXECUTIVE SUMMARY

From July to October 2020, ethix - lab for innovation ethics interviewed representatives from national and international civil society organizations and other interested organizations and individuals on behalf of the Swiss Digital Initiative. Over a hundred organizations were invited to take part in the process. Participants were invited to give feedback on the first draft of the Digital Trust Label prepared by the Swiss Digital Initiative. This co-development process is part of the ambition to produce a cutting-edge label that includes multiple perspectives on digital trust.

All involved participants underlined the relevance of digital trust, and the importance to address it from a consumers' perspective. Besides its use for consumers, the creation of a label presents an interesting opportunity to crystallize a catalogue of technical requirements that make up trustworthy services.

Participants emphasized that the legitimacy of the development process, but also of the label-giving body, is key. Transparency and diversity are crucial in this respect. As a general strategic question, the label needs to find its position with respect to other standards and legal requirements. Furthermore, the label must find its way between a high level of generality and sector specificities. Detailed feedback on the specific criteria of the label can be found in this report. The final section of this report entails recommandation by the ethix team.



ABOUT THE SWISS DIGITAL INITIATIVE

The Swiss Digital Initiative is a long-term, sustainable process to safeguard ethical standards in the digital world through concrete projects. It brings together academia, government, civil society and business to find solutions to strengthen trust in digital technologies and in the actors involved in ongoing digital transformation.

The initiative has a global focus and is headquartered in Geneva, Switzerland. It was initiated by the cross-sectoral association digitalswitzerland under the patronage of Federal Councillor Ueli Maurer.

digitalswitzerland.com/sdi

1
INTRODUCTION AND GENERAL REMARKS

The Swiss Digital Initiative (SDI) is developing a Digital Trust Label addressing the trustworthiness of digital services. A draft version of the label was created by an Academic Expert Group from EPFL, ETH and the Universities of Zurich and Geneva. On this basis, the co-development process was led from July - October. The aim of the co-development phase is to collaborate with a range of civil society stakeholders to challenge the first draft version of the label. This feedback reinforces the inclusion of consumers and civil society voices in the label's development. It increases the label´s quality and improves its legitimacy.

Three methods were used to gather feedback from national and international participants:

- **Face-to face interviews:** semi-structured interviews with national and international civil society participants, lasting approximately 45 minutes to one hour, were held by the co-development process team.

- **Online survey:** In addition to the interviews, feedback on the label's content was gathered by means of an online survey structured around the content of the label.

- **Workshop:** participants from civil society and academia met in Geneva on September 8 for a three-hour workshop and discussion of the label's content.

The report will be made public on the SDI website, inviting further commentary from participants. In the meantime, the report will be sent to the Experts' Committee in charge of revising the label and to the board of the SDI. The current draft version of the label will then undergo a revision, a second testing phase and user study.

2
FEEDBACK ON THE SCOPE OF THE LABEL

This section presents feedback on the overall scope of the label.

Scope: object of the label?
In its current form, the label is conceived to apply to 'digital services'. Participants agreed that the label should be applied to a specific service and not certify an institution or company. However, there was disagreement over how useful a general label, applicable to any type of digital service, would be. It could fail to account for the challenges of a digital service in a specific sector and remain too superficial. A more specific approach would require the development of distinct sector-specific labels.
Although it should not certify an institution or company, the label could only apply to one or a few services offered by a company. It implies that some of the services offered by company X could be labelled, while others would not. Furthermore, participants were uncertain as to who are the final users of the label (exclusively consumers or B2B users also thinkable?).

Scope: national, regional, global?
Several participants mentioned the issue of differing values worldwide. Concepts used in the label such as "privacy" and "reliability" are understood differently depending on the political, legal or cultural contexts one is found in. This means that decisions need to be taken for the development of the label, mainly whether the label should be applied in a broadly uniform context and whether it should be designed to be acceptable across different political, legal and cultural contexts. One of the key questions is here the opportunity to adopt the European standard of data protection as standard for the label.

Scope: beyond legal obligations?
A majority of international participants pointed out the similarities between the SDI Trust Label and other existing standards (e.g. ISO, ECI, GDPR, convention 108+). As companies are legally obliged to comply with national and international regulations such as GDPR, participants suggested that a label goes beyond these legal requirements. Otherwise, the label would provide little added value to consumers and service providers alike.

3
FEEDBACK ON THE DEVELOPMENT PROCESS

This section presents feedback on the label's development process.

Transparency and participation
A number of participants noted that the inclusion of consumer representatives and civil society in the development process has been insufficient. While the co-development process is considered a good step, participants recommend that this inclusion is institutionalised for the further development of the project.

Re-assessing the label itself
To ensure legitimacy and the highest quality, the label itself needs to undergo periodic audits. Clear mechanisms to continuously adapt the label and its criteria to the current needs of users and the development of technology should be developed.

Third party control
Third party control mechanisms were underlined by many participants as crucial for the quality of the label. Self-assessment approaches are not enough. This was mentioned in a number of comments towards the specific label's criteria, particularly for fair data management and security indicators. For example, a few respondents suggested that security parameters should always be reviewed independently, by an external auditor.

4
FEEDBACK ON THE CONTENT OF THE LABEL

General Remarks
The label is a list of criteria which a specific service ought to fulfil. These criteria are organized into four main categories: Security, Data Management, Service Reliability and User Interactions. These categories are taken as working categories to organize and give structure to the label's content. On the basis of a user-study conducted in 2019, these categories were identified as corresponding to what digital trust means to users. Each criterion is then categorized into indicators, which are operationalized through several variables each. During the interviews and through the online survey, participants judged how adequate these variables are at operationalizing the specific concept and were invited to comment.

Open source
Many participants mentioned Open Source and Open Data as important concepts that need to be taken into account by the label. While transparency is mentioned in fair data management, with regard to informing users on certain aspects of the service, it is considered by a few as not sufficiently incorporated in the security or reliability categories.

Ambiguity of measurements
Almost all survey respondents perceived some words within certain variables as requiring further precision. Wordings such as "best practice" or "deemed sufficient" are not comprehensively defined and cause ambiguity. Some variables are believed to still need additional information to be measurable. The use of "state of the art" instead of "best practice" is considered preferable by some participants, as it usually implies a higher standard than simply what is done within an industry.

Product re-assessment
Digital services evolve very quickly. A few participants emphasize that services and tools that have been granted the label, have to be periodically reassessed and undergo surprise audits. Relying on self-reporting and self-assessment of companies is not deemed sufficient by participants. The criteria to be evaluated, but also the periodicity of doing so, is crucial for trust.

Missing elements
The four categories and their criteria are seen by most as a good representation for digital trust. However, many respondents have pointed out important aspects that are still missing:

- Four participants both through the survey and during the workshop have discussed that trust should also include considerations of climate and health – how environmentally friendly a digital service is. Sustainability matters and the label's content should, according to these participants, integrate it as part of what trust entails.

- New developments in Big Data weaken solidarity, for example in the insurance sector. Therefore, one participant suggested including the notion of value sensitive design from the beginning.

- Responsibility of service providers to assure no misuse of their products for unethical purposes. (e.g. crypto banking services have the responsibility to assure that their services are not abused for money laundering).

SPECIFIC REMARKS

4.1 Security

a. Secure Communication, data transmission and storage
b. Secure user authentication
c. Secure service set up, maintenance and update
d. Vulnerability/ breach monitoring/ reporting

Specific feedbacks:

- A few participants mentioned that security by design and security by default should be included from the beginning.
- One participant sees it as unnecessary to specifically mention passwords. Passwords are considered an outdated and inherently insecure method. "Authentication" could be used instead.
- One participant mentioned that certain aspects of encryption, maintenance or authentication are often outsourced to third-party suppliers. Yet, that in most cases, this is not communicated to users. It is not clear if the security variables also apply to third-party providers.
- Auditing in security issues, as in other issues, can be a very costly process only big companies may afford.

4.2 Fair data management

a. Privacy policy and user content
b. Data collection, storage, exploitation
c. Data retention, access, rectification

Traceability, usage and ownership of personal data
A few participants underlined the importance for consumers to be able to see what happens to their data, who is using it, and whether third parties can be traced. Furthermore, it is not enough to "inform about rights to request their personal data". A service needs to actively make it possible for users to download and obtain their data in a format they can understand.

Portability of data
According to several participants, personal data should also be transferable to a different service provider as is the case under GDPR. A participant mentioned the principles of mydata.org as a good benchmark when it comes to data portability.

Consent v information
One participant from Switzerland sees the concept of consent as used in the label as problematic and not working in practice (e.g. consent-tiredness through cookie policies). He suggests the focus be put on information. Is consent necessary for any kind of data collected? Several participants on the other hand, pointed out that the level of consent represented in the variables does not go beyond the GDPR.

Higher standards for vulnerable people
A few participants have pointed out that children and other vulnerable people need higher standards and stronger protection than adults. Some believe that this should not be the responsibility of companies but remain solely with parents, while others think that it has to be included in the design as well. For example, no third-party cookies should be placed on websites targeted at children.

Generic vs differentiation
A majority of participants perceive the label as very generic, which offers one standard for all services. They point out the need to differentiate between different types of data. Different services require different standards. For example, data related to location or personal health requires a much higher level of security and/or anonymization than data collected by a gaming app.

Privacy indicator
The PEP Foundation is working on developing generalizable privacy indicators and a graphic representation (icons, colour schemes etc) of these indicators. One participant suggested cooperation with this partner.

Data sold to third-parties
If anonymized data is being sold to third parties, these sellers should put in place contracts that prohibit the de-anonymization of data. The German data ethics board has recently proposed that [1].

Specific feedbacks:

- While companies may make their practices transparent, this does not change the way they use data. Therefore, transparency may not be sufficient.
- The concept "data protection" is understood differently in the United States than in Europe. While in the US, exploiting data for commercial purposes is acceptable and government exploitation has many restrictions, the opposite is the case in the EU.

1 Opinion of the Data Ethics Commission – Executive Summary. Standards for the use of personal data no.20: https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.pdf?__blob=publicationFile&v=2

4.3 Reliability

a. Reliable service updates
b. Resilience to service outage
c. Functional reliability
d. Accountability

Distinguishable software versions
A user gets access to the software but does not get the software itself. It is becoming harder and harder for users to identify the precise software they are using, which makes user complaints extremely difficult.

Terms and Conditions
A few participants mentioned that when updating their terms and conditions, companies should make clearly visible which specific parts have changed.

- Could the label incorporate that a certain standard ought to be provided beyond simply having terms and conditions?
- How can "extensive descriptions" be made compatible with user friendly information?

Specific feedbacks:
- Notion of automated patch management could be included (Indicator 3.3).
- Duty to inform if there has been a data breach
- For certain services (e.g. whistle blowing platforms) it is necessary that the service provider cannot be identified (Indicator 3.3, variable a).

4.4 Interactions with users

a. Non-discriminating access
b. Fair user interfaces
c. Fair use of AI-based algorithms

Fair use of Ai-based algorithms
One participant mentioned that the label should address chatbots specifically: a company should always have to inform users when they are interacting with a chatbot (specifying indicator 4.3, a). Several participants believe that explainable AI is crucial for fair user interaction.

Data bias
When dealing with algorithms, non-discrimination plays a particularly important role. Evaluating the quality of the data used for an algorithm and the impact of this algorithm should be a permanent requirement (specifying indicator 4.3, d).

Accessibility and inclusion
A number of participants pointed out the importance of accessible service and information – e.g. is the service and information about the service accessible in the language of the country it is operating in? Furthermore, it is unclear as to what "non-discriminating access" means (specifying indicator 4.1)

Specific comments:
- Speaking about "user management" frames user interaction in a paternalistic way. Wording should be chosen carefully with central focus on empowering users through information.
- Within the setup of the label, technical questions come before social questions – interactions with users.
- What are the boundaries between influencing behavior and manipulation, and how is user manipulation defined in the label?
- Artificial Intelligence is currently an ill-defined concept. Instead the wording "Automatic Decision Making" could be used. The effect on the user is decisive, as a complex AI may have the same effect as a simple algorithm.
- When things go wrong: user empowerment was an important discussion during the workshop and interviews. Particularly, the need for companies to have comprehensible and easily accessible complaint processes. E.g. is there a place, and a person a user can talk to?

5
CONCLUDING REMARKS AND RECOMMENDATIONS

Overall, participants considered the current content of the label solid and a good representation of digital trust. However, open questions and doubts remain on its operationalization and the labeling process more generally. The need to communicate more transparently about the development process and the need to include civil society stakeholders in all stages of the development and deployment process of the label were highlighted. Interestingly, international participants mentioned that the label remained very close to existing regulation (e.g. GDPR in European Union) and would therefore provide only a small added-value for consumers in those areas. Both national and international participants mentioned the importance to better embed the label in existing standards and regulation, as well as to join forces with similar initiatives around the globe.

This following section presents the recommendations based on the feedback and inputs received during the public consultation process. These recommendations are formulated by the ethix team and are directed at the Experts Committee, in charge of further developing the label's content, and at the SDI Board, in charge of strategic decisions on the positioning of the label.

Firstly, the gathered feedback shows the crucial importance of **building trustworthy processes**. The processes underlying the label project must embody the ambition by the SDI to foster trust in digital services. With regards to the processes, trust means the capacity to make room for diverse voices and to take them into account in the development of the label (principle of diversity). In order to avoid a one-sided view on the topic and to avoid an exclusively national focus on the issue, academics (represented in the Experts Committee), business representatives, public servants but also civil society representatives and consumers representatives should have the opportunity to participate in the design of the label. Furthermore, the SDI should aim to be as global as possible and integrate voices from Switzerland, from Europe, from other industrialized states but also from developing countries. The legitimacy and the quality of the label are at stake.

The principle of diversity should apply to the development process of the label, as shown by this co-development process. It should equally apply to the (post-) deployment process, i.e. when it comes to keeping the label up-to-date. The actualization of the label's content should be open to diverse voices. Ideally, the process should be open to comments with the aim of constantly improving the label and adapting it to changing, multicultural realities.
In the same way, this diversity of voices should be represented in the decision-making bodies of the SDI. Diversity is nothing without the capacity to have a formal say in decision-making procedures and to influence the power relations. This concerns mainly the Board, in charge of strategic decisions, but also the Experts Committee, in charge of recommending changes to the label's content.

Secondly, the feedback underlines the requirement to take a strategic decision on the **definition of trust** upon which the label is built. The objective of the SDI is not to develop a label evaluating whether a digital service is ethical in general, but whether it is trustworthy. The focus is more specific. The label needs to address all the dimensions linked to trust in the context of digital services. It is not clear that sustainability - as proposed by some participants - is part of trust, though it is highly important for other reasons. The danger of integrating further elements is to make the label too comprehensive. This would mean to evaluate every possible aspect of a digital service being described as "ethical" as part of trustworthiness. If possible, further investigation, e.g. user-study, should try to find out if users associate trust with sustainability.

Thirdly, the feedback shows that the criteria were **complex to assess and evaluate** by a third-party. Beyond some binary criteria that can be answered with a yes/no, most criteria require a careful evaluation and depend on continuously evolving 'state-of-the-art' standards. This evaluation should be done according to standards that are made transparent to all actors involved. Even with the transparency requirement fulfilled, the criteria call for a complex evaluation to be done by audit specialists. The required evaluation cannot be compared to information which only makes explicit what the content of a product is (e.g. nutrition facts). The trust label is, and must be, a true label in the sense of defining a benchmark to be reached and measuring if a specific service reaches the defined benchmark or not.

Fourthly, the feedback shows that the SDI Board should take a strategic decision on the **broad standard applicable** to the label. The issue is especially important for data-related issues, as to whether the GDPR (or the Convention 108+) should represent the basis for the label. To take this decision, the Board needs to address the underlying question of the overall goal of the label. Should the label be a quality label which only deals with digital services which are already GDPR-compliant, de facto excluding digital services which do not take European users as customers? Or should the label be a way for qualitatively less developed digital services to improve their standard? Both goals are interesting and contribute to the reinforcement of trust as a key feature of digital services. However, the main added value for the digital trust label as developed by the SDI is, at least for the time being, in Switzerland and in Europe. It should focus on providing added value for consumers, businesses and public institutions in these markets. In that sense, it should take European legal standards as basic standards for the label.

Fifth, the feedback makes clear that a label dealing with digital trust has no choice but to be international in its ambition. For the SDI, it means the requirement to cooperate with like-minded projects and organizations in trying to level the playing field. Taking advantage of the ecosystem found in Geneva, the SDI could be one of the key actors to make a new consensus emerging on the criteria to operationalize the concept of digital trust.

6
LIST OF PARTICIPANTS

The aim of the co-development process was to include an array of diverse perspectives. Over a hundred organizations were invited to take part in the process. Of these, 35 organizations were based in Switzerland. The 71 international organizations invited were predominantly based in Europe or the United States. Invitations were sent to organizations in Germany, the United Kingdom, Belgium, Denmark, the Netherlands, Finland, Spain, France, Canada, South Africa, Colombia, India, and Pakistan.
Out of over a hundred invitations sent, 26 actively participated, giving feedback on the Swiss Digital Trust Label. Many of those who declined participation stated lack of resources and time as reason.

The following people have participated in the co-development process, either through interview, survey, or the workshop in Geneva.

- Jean-Yves Art, Microsoft
- Yaniv Benhamou, UniGe
- Peter Bihr, The Waving Cat, ThingCon, Germany
- Sage Cheng, Access Now
- Christophe Ebell, VertiAI, Switzerland
- Elisabeth Ehrensperger, TA Swiss
- André Golliez, Association Opendata, Switzerland
- Elea Himmelsbach, Open Data Institute, UK
- Mathias Holenstein, Stiftung Risiko Dialog, Switzerland
- Michael Kende, Graduate Institute
- Valérie Khan, Digital Equity Association
- Helena Leurent, Consumers International
- Arié Malz, ISSS association, Switzerland
- Hernani Marques, Pep Foundation, Switzerland
- Giacomo Mazzone, RAI
- Rohinton Medhora, Centre for International Governance, Canada
- Jean-Henri Morin, UniGE
- Nicholas Niggli, Canton Genève
- Stephanie Nguyen & Ginny Fahs, Consumer Reports
- Cailean Osborne, Centre for Data Ethics and Innovation, UK
- Jacub Samochowiec, Gottlieb Duttweiler Institute
- Thomas Schneider, BAKOM
- Martin Steiger, Digitale Gesellschaft Schweiz, Switzerland
- Christoph Stueckelberger, Globethics
- Pernille Tranberg & Birgitte Kofod Olsen, DataEthics.eu, Denmark
- Nicolas Zahn, Operation Libero, Switzerland

ethix

Lab for Innovation Ethics
Laboratoire d'éthique de l'innovation
Laboratorio per l'etica dell'innovazione
Lab für Innovationsethik

Authors:

www.ethix.ch

ethix – Lab für Innovationsethik
Zweierstrasse 100
8003 Zürich