





Contribution to the United Nations Global Digital Compact

by NGO Identity Valley Research gUG, Unkel, Germany, April 29th, 2023

From Re-Action to Re-sponsible-Action. From BIG Tech to RESPONSIBLE Tech.

DIGITAL RESPONSIBILITY GOALS – A SHARED VISION FOR A BETTER DIGITAL TOMORROW

On the cusp of the age of artificial intelligence, our society is faced with daunting questions: how digital will our future be, and how will this affect our humanity?

In the last two decades, technological progress has tested the social fabric of our communities. Privacy concerns, disinformation, biased algorithmic decision-making, a lack of digital self-determination or the manipulation of our behaviour have led to distrust, polarisation, and disorientation. Only with a shared global vision for a better digital tomorrow are we able to get back on track towards a digitalisation that works for people and planet.

To address this need for a shared vision, IDENTITY VALLEY has developed a framework of guiding principles for a digital future that prioritizes human identity, social cohesion and trust. This framework, developed through a multistakeholder process involving experts from different areas, is comprised of seven Digital Responsibility Goals (DRGs), each designed to reduce complexity in our increasingly digitised lives and make responsible behaviour in the digital world visible and comprehensible.



Figure 1: Digital Responsibility Goals

Those DRGs are ...

- **easy to understand**, making complex characteristics of digital applications, such as cyber security or data fairness, more approachable to non-professionals,
- **measurable**, allowing for comparability and a concrete path towards trustworthy and human-centric digital solutions, and

• **principle-based** - and therefore future-proof (e.g., also applicable to emerging technologies, like Web 3.0, digital identities, artificial general intelligence...).

Using the DRGs as guiding principles in the conception and operation of digital services or applications facilitates trustworthiness "by design" and can provide, for the first time, a recognizable badge of digital responsibility to users of technology. This will not only incentivize companies, organizations, and administrations to strive for digital trustworthiness as an added value but will also direct users towards digital solutions prioritizing digital responsibility and human identity.

FROM DECLARATION TO ACTION

Decision-makers are aware that we are moving in the wrong direction. Governments and companies around the world have issued many well-intentioned declarations on protecting privacy, empowering users vis-a-vis digital services, or striving for explainable AI. However, these initiatives are often insular and lack commitments. To effect change, we need concrete verifiable commitments.

The Digital Responsibility Goals provide a holistic framework for such commitments – easily trackable and transparent. Their principle-based approach ensures that the objectives at their core won't need to be adapted as technology evolves further. The values on which they are based are universal and thus transcend linguistic or cultural barriers. This means that everyone can contribute to transform our increasingly complex virtual world into a trusted digital ecosystem in which social and economic interactions can flourish.

Governments can use the DRGs as a compass for their digital strategies, businesses can tailor their digital services to meet clear objectives of trust, civil society can boost awareness, and individuals can demand trustworthiness in their increasingly digital lives. Additionally, the DRGs can serve as an addition to the set of decision criteria of investors, who will for the first time be able to identify businesses valuing digital responsibility and contributing to a sustainable digital transformation.

IMPLEMENTING A DIGITAL RESPONSIBILITY INDEX

To add applicability to the DRGs we propose a Digital Responsibility Index (DRI) that enables the scoring of digital solutions based on guiding criteria derived from the overarching DRGs. Those guiding criteria (find them in the Annex) empower producers and providers of digital technologies to take concrete measurable actions to align with the DRGs.

Because the guiding criteria are tied to pre-defined metrics, a dynamic evaluation of digital responsibility for a specific digital solution can be achieved (for details see Annex). By ranking the results of these evaluations, digital solutions can be transparently listed and easily compared within an emerging Digital Responsibility Index.

Contrary to this approach, existing certificates, and regulations with their selective assessments of particular areas are not sufficient to spark a holistic transition towards responsible technology. Only a framework that merges the most significant certificates, policies and regulations and visualizes them in a simple way can kickstart comprehensive global change.

Driven by the users' desire and the societal necessity for a responsible approach to digital technology, companies, organizations, and administrations will be incentivized to improve their digital responsibility. In particular the proposed index will enable a non-legislative "race to the top". A labelling system based on the DRI could empower users and stakeholders further to make informed choices about the digital services they use, promoting a trustworthy and human-centered digital ecosystem.

THE IMPORTANCE OF TRUST IN THE DIGITAL WORLD

The building of trust is at the core of our initiative. Trust is essential for human societies to function. It enables cooperation and underpins social and economic interactions. However, trust is not innate; it must be built over time. Unfortunately, trust has been eroding between groups within societies, between people and public institutions, and in digital technologies.

To achieve sustainable development through digital transformation, we must reverse this trend and rethink the digital transformation. First and foremost, technology needs to serve planet and people, not the other way round.



Figure 2: Human- and planet-centred impact model by Identity Valley

Especially, the rapid proliferation of powerful AI systems has shown both the necessity and complexity of agreeing on a global governance mechanism to rein in this powerful technology. In the face of disagreement on how to achieve a multilateral accord with clear commitments for the future digital world, our approach proposes a voluntary framework that could be implemented fast and from the bottom up.

ANNEX:

DIGITAL RESPONSIBILITY GOALS & THEIR GUIDING CRITERIA IN DETAIL

The 7 Digital Responsibility Goals (DRGs)



DRG#1 - DIGITAL LITERACY:

Digital Literacy and unrestricted as well as competent access to digital services and infrastructure are prerequisites for the sovereign and self-determined use of digital technologies. They are the basis for all of the Digital Responsibility Goals.



DRG#2 - CYBERSECURITY:

Cybersecurity protects systems against compromise and manipulation by unauthorized actors and ensures the protection of users and their data - from data collection to data utilization. It is a basic prerequisite for the responsible operation of digital solutions.



DRG#3 - PRIVACY:

Privacy is part of our human dignity and a prerequisite for digital self-determination. Protection of privacy - with a consistent purpose limitation and data minimization - allows users to act confidently in the digital world. Privacy by design and default enable responsible data usage. Users are given control and providers must account for how they protect privacy.



DRG#4 - DATA FAIRNESS:

Non-personal data must also be protected and handled according to its value. At the same time, suitable mechanisms must be defined to make data transferable and applicable between parties. This is the only way to ensure balanced cooperation between different stakeholders in data ecosystems.



DRG#5 - TRUSTWORTHY ALGORITHMS:

Once data has been collected, it must be processed with the aim of trustworthiness. This is true for simple algorithms as well as for more complex systems up to autonomously acting systems.



DRG#6 - TRANSPARENCY:

Proactive transparency for users and all other stakeholders is created. This includes transparency of the principles that underlie digital products, services, and processes, and transparency of the digital solution and its components.



DRG#7 - HUMAN AGENCY AND IDENTITY:

Especially in the digital space, we must protect our identity and preserve human responsibility. Now. Preserving the multifaceted human identity must be a prerequisite for any digital development. The resulting digital products, services, and processes are human-centered, inclusive, ethically sensitive, and sustainable, maintaining human agency at all times. Only in this way can digital technology promote the wellbeing of humanity and have a sustainable impact.

The guiding criteria of each of the DRGs

DRG#1	DRG#2	DRG#3	DRG#4	DRG#5	DRG#6	DRG#7
Digital Literacy	Cybersecurity	3 Privacy	4 to Sairness	Trustworthy Algorithms	6 Transparency	Human Agency & Identity
designed individually and in a way that is suitable for the target group.	operators of digital products, services and processes assume responsibility for Cybersecurity. Users also	Operators and providers of digital products, services, and processes must take responsibility for protecting the privacy of their users.	When collecting data, proactive care is taken to ensure fair representation of the context in which it is collected.	provide a maximum of	organizations establish transparency about their digital business ventures and solutions - for the	The preservation of the multifaceted human identity is a basic requirement and must be the basis for any digital development. The resulting digital approaches are always user centric. They respect personal autonomy and dignity, and limit commoditization.
services, and processes must be reliable and barrier-free.	solutions are responsible for appropriate security measures and are constantly developing them further. Products,	personal data basic principles of data protection are respected, in particular strict purpose limitations and data	In digital ecosystem structures, the mutual exchange of data between all parties involved must be clearly described and regulated (data governance). The goal must be fair participation in the benefits achieved through the exchange of data.	The individual and overall societal impact of algorithms is regularly reviewed and the review documented. Depending on the results, proportional measures must be taken.	implemented in interactive	Sustainability and climate protection must be part of digital business models and implemented in practice (especially in accordance with the UN Sustainable Development Goals).
processes must be proactively considered in design and operation. This	implementation of cybersecurity are considered along the	Privacy protection is considered throughout the entire lifecycle and should be considered as default setting.	Developers, providers, and operators of digital solutions must clearly define and communicate the purpose with which they use and process data (including nonpersonal data).	The results of algorithmic processing are comprehensible and explainable. Where possible results should be reproducible.		responsible, non- manipulative communication. Where possible, communication
the digital transformation is essential - everyone is entitled to education on	products, services, and processes must account for how they provide security for users and their data - while maintaining necessary trade secrets and	over their personal data and its use - this includes the rights to access, rectify, erase, restrict processing, avoid automated	The "FAIR" data principles are satisfied, especially for use cases relevant to society as a whole - "FAIR" stands for Findable, Accessible, Interoperable, Reusable.	AI systems must be robust and designed to withstand subtle attempts to manipulate data or algorithms.	transparency for users, transparency should also be provided for other	Digital technology always remains under human authorship and control - it can be shaped throughout its deployment.
information offered should be designed to create awareness of related topics such as sustainability, climate protection, and diversity/inclusion (e.g.,	authorities, civil society and science must collaboratively shape the framework for cybersecurity with	Providers must account for how they protect users' privacy and personal data - while maintaining necessary trade secrets and information security.		control of their mode	Organizations must outline how they will	Technology may only be applied if it is of use to individuals and mankind and promotes the wellbeing of humanity.

Figure 3: Table of the DRG guiding criteria

A POTENTIAL MECHANISM FOR MEASURING AND LABELLING

Towards measuring Digital Responsibility

For the purposes of outlining a potential mechanism to measure digital responsibility according to the DRGs, in the following example let's assume we want to evaluate a fictional conversational "AI companion" app.

From Guiding Criterion to Metric

If we want to assess the app's compliance with DRG#5 Trustworthy Algorithms, a simplified matrix for translating the guiding criteria into a concrete metric might look like the following:

DRG#5 Trustworthy Algorithms - Guiding Criteria	Governance		Interpretation / Maturity Level	Num. Score
5.1 Algorithms, their application, and the datasets on which they are used or trained are designed to provide a maximum of fairness and inclusion.	Gender diversity of data scientists and software developers	workforce in software development	0-5 High 6-10 Advanced 11-20 Basic 21-100 None	10 5 2 0
5.2 The individual and overall societal impact of algorithms is regularly reviewed and the review documented. Depending on the results, proportional measures must be taken.	Evaluation and audit policy, and documentation		3+ High 2 Advanced 1 Basic 0 None	10 5 2 0
5.3 The results of algorithmic processing are comprehensible and explainable. Where possible results should be reproducible.	Explainable AI techniques are employed		Yes - High No - None	0
•••	•••	•••	•••	

Figure 4: DRG#5 Trustworthy Algorithms - From guiding criterion to metric (extract) - Prototype

From metric to visualization of digital responsibility

The "maturity level" of an evaluated guiding criterion is visually reflected in a different intensity in coloring. The more color in the labelling the more digital responsibility can be expected.

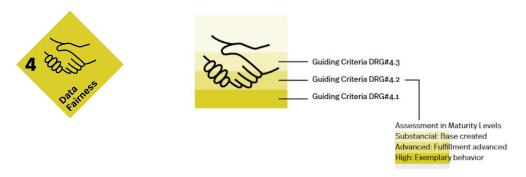


Figure 5: Visualization of DRG evaluation – example of DRG#4 Data Fairness

A full-scale measurement model will employ several metrics per guiding criterion enabling more granular comparison and scoring. Additionally, a numerical score allows for precise comparability by categories or in total, enabling the creation of a ranked index.

About IDENTITY VALLEY

As a non-profit organization, Identity Valley Research engages thought leaders in academia, policy and industry for a value-based future of the Digital World through networking, advocacy and communication. Identity Valley advocates for a data economy based on trust, privacy, and personal identity, derived from the humanistic tradition of Europe. In this, the organization is partly a response, partly an evolution of Silicon Valley. It is about both the possibilities of technology and the accompanying assumption of responsibility – by companies, institutions, and states. In the process, the uniqueness of multi-faceted human identities replaces "silicon," – until now probably tech's most important raw material. Identity Valley evolves the question "What can technology do?" to the question "What should technology do?".



Contact

Identity Valley Research gUG Pützgasse 6 53572 Unkel, Germany https://identityvalley.org/ #itsallabouttrust



Ferdinand Ferroli
Director Policy & Research
f.ferroli@identityvalley.org
https://www.linkedin.com/in/ferdinand-ferroli/



Jutta Juliane Meier
Founder & CEO
jj.meier@identityvalley.org
https://www.linkedin.com/in/juttajulianemeier/