

Statistical awareness promoting a data culture

Walter J. Radermacher

Ludwig-Maximilians-Universität München, Germany

E-mail: wjr@outlook.de

Abstract. Good statistics can do a lot of good: They help to base decisions on factual arguments, they can simplify conflict resolution. This requires an understanding of the opportunities and risks, the strengths and limitations of statistical facts. Overestimation leads to exaggerated expectations and disappointments, underestimation to missed opportunities, risks. Even worse is the trouble if facts are influenced or manipulated with political intentions or if even the impression of arbitrariness is created with so-called ‘alternative facts’. The very bad excesses of political misuse of statistics are carried out with intent and not negligently. Nevertheless, it can be argued that the virus of false and manipulated information flourishes when the statistical literacy of the population is at a low level. On the less serious scale of missed opportunities or too high expectations regarding statistics, there are, of course, also observations that suggest that an improvement in statistical literacy would be very good for politics, both on the part of the population and on the part of politics itself. Overall, the aim must be to promote and nurture a culture in which a conscious and experienced approach regarding data and statistics has become the standard.

Keywords: Data culture, statistical literacy, evidence-informed policy-making, data worlds

1. Introduction

In today’s digitised (and globalised) realities and societies, an understanding of data is critical for everyone. Knowledge about data and knowledge from data are equally important for processes and progress in technology, business, administration and science. Last but not least, this knowledge can and should also positively contribute to a public discourse based on existing facts. In view of such fundamental importance, it should be assumed that statistical knowledge, as an essential prerequisite for equal participation in public life, is given a correspondingly high priority, be it on the part of (education) policy, as subject of research or also on the part of the economy. It is even more astonishing to observe how, during the pandemic crisis, data, facts and indicators have not contributed to objectification but rather to confusion. The question therefore arises as to how the objective of improving ‘digital skills’¹ can be filled with life and achieved by means of concrete activities.

This paper does not seek to add value to the vast literature on statistical education and teaching of statistics.

Nor will it be about addressing the question of whether there is a difference between (classical) statistics and (newer) data sciences in terms of education [1] (to the extent that such categorisations are helpful at all [2]). Rather, in a narrower sense, it is about the interrelationships between the generation and use of evidence in the form of quantitative information for the decision-making and communication processes of the society as a whole. It is therefore about governing through information in mutual relation with institutional change due to information flows. This interrelation of governance by the numbers is influenced by new processes of information construction, new technologies (e.g. new social media) and other driving forces thus making it even more complex [3]. In order to further narrow down the subject matter of the paper, it should be emphasised that for the aforementioned questions, the focus will primarily be on statistics that are operated by public institutions (i.e., official statistics).

Before we turn to the question of statistical literature and culture in relation to the use of evidence for policy-making, the terrain needs to be prepared. It makes little sense to juxtapose two monolithic blocs, namely evidence on the one hand and politics on the other. Rather, it is necessary to make a sufficiently fine decomposition so that the web of relationships that is decisive

¹For example in *Europe’s Digital Decade*. https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en.

for the various forms and levels of statistical education becomes visible. We will approach the topic by first characterising the different data cultures that exist and are relevant to our topic. Secondly, we will describe the (circular) value creation process from data to knowledge, in the phases of which different competences are used and required. Finally, we will introduce the term data culture, which we use to describe an overall social and political climate.

2. Trust in numbers in the time of datafication

The conditions under which public statistics operate are marked by the difficulties of their time. In the dictatorships of the 20th century, statistics were put on a short leash by state interests; in the Thatcher era, they were shrunk as superfluous ballast. Both had a devastating effect on the quality of evidence and the public trust based on it. Tim Holt, then President of the Royal Statistical Society and Director of the Office for National Statistics, commented on the establishment of modern statistical governance through the Statistics and Registration Service Act 2007: “Statistics, their production and interpretation, really do matter. It is uncommon to find some area of public debate that does not draw on a body of statistical information. As a result, public confidence in the statistical system and its outputs is vital. In particular this is essential if people are to have confidence in the decisions that are made on their behalf by elected leaders. For example, people will not accept the case for closing a school or hospital if they do not trust the statistics on which the case is made. More generally, no one will trust government if it is seen to rely on statistical information which is perceived as untrustworthy.” [4] Today, a new threat is that the *statistical logic* of the past two centuries is being replaced by a *data logic*: “The new apparatus of number-crunching is well suited to detecting trends, sensing the mood and spotting things as they bubble up. It serves campaign managers and marketers very well. It is less well suited to making the kinds of unambiguous, objective, potentially consensus-forming claims about society that statisticians and economists are paid for.” [5] Statistics and the significance of statistics with the function of authority are disappearing from the public consciousness and the political agenda. If nothing takes over this role of statistics with the stamp of *officiality*, the door is opened for any form of arbitrariness, not least also and especially on the basis of data sciences and a *new generation of data* [5,6]. Imagine that the GDP in its

current form would be up for disposal. Instead of an extremely elaborate accounting based on international standards, it could be replaced by an estimate based, for example, on night-time light intensity from remote sensing measurements. Instead of a single *official* indicator with possibilities for comparison over time and space, we would find ourselves in a maze of parallel solutions, all claiming to be based on statistical and data science methods.

In 2019, the Royal Statistical Society has set out the following comprehensive and well-balanced recommendations in a *Data Manifesto*:

1. “Ensure official statistics are at the heart of policy debate.
2. Commit to greater data sharing between government departments for statistics and research purposes.
3. Champion basic training in data handling and statistics for politicians, policymakers and other professionals working in public services.
4. Maintain the commitment to keep pace with other leading scientific nations on investment in research and development.
5. Give the Office for National Statistics and the wider Government Statistical Service adequate resources.
6. Prepare for the data economy by skilling up the nation.
7. Involve the public in shaping the conversation about how data is used.
8. Misinformation needs countering but without undermining free speech.
9. Move beyond averages and break down data to a much more granular level.
10. Keep data regulation updated to protect the public [7].

What is (positively) striking about this Manifesto and the list of recommendations is the close interlocking between statistics/data sciences and their application. A holistic view of the interaction between the three historically important drivers of development in statistics, namely science, statistics and society [8], ensures an adequate response to the challenges of our time.

2.1. Revolutions, evolutions, mega-trends

As a prerequisite for promoting improved data practices, there needs to be an analysis and understanding of what role data plays in contemporary society. In this context, it is by no means sufficient to devote attention to the technical processes, the infrastructure, the eco-

nomics of data or similar. Likewise, it is not enough to focus the ethical and legal questions on the protection of individual privacy. With a proliferation of data, new dynamics have been triggered that have led to the replacement of the previous well-ordered (infra)structures of statistical production based in the public sector by other forms of (decentralised and individual) use and production of data. This may be accompanied by a great promise, as a result of which ubiquitous amounts of data would also include more equality and less dependence on authoritative and technocratic institutions. In a direct sense, the fulfilment of such a promise would depend on everyone actually having a sufficient level of competence to be able to draw good statistical analyses from the data sets. That is however less the point here. Rather, the indirect and accumulated risks associated with such an approach that relies significantly on decentralised competences and processes are of interest, i.e. when there is no longer a common statistical language based on standardised methodologies, classifications, etc. [9,10].

In a traditional division of roles, there are those who produce data, facts, indicators, etc. and those who use such results, producers vs. consumers, to put it simply. Today, the situation is a little more complicated. The traditionally divided roles have become mixed, consumers can also become producers using widely available data and evaluation tools, everyone is a potential 'prosumer'. In principle, this development is to be welcomed; it opens up diverse opportunities for participation and emancipation. At the same time, however, this development is also accompanied by substantial dangers:

- There is a trend decline in trust in public institutions and governments; people feel that they have little or no say in public decisions and that public policies do not address their concerns: this undermines trust in democratic institutions.
- New media and information channels have altered how people access, transmit and share information, blurring existing institutional arrangements.
- Big Data discussions are often characterised by a scientific and technical bias, at the expense of a consideration of the socio-political implications.
- When statistics produced by state institutions are replaced by individual evaluations, comparability gets lost; standards (and authority) are substituted by diversity.
- There is a proliferation of on-line data and content that can lead to mis-use, mis-information, disinformation and distrust in information.

- There is an erosion in the social consensus regarding the role ethical behaviour (respect the general validity of the Rule of Law, trust in institutions) plays in the good functioning of democracies and societies.
- New actors interfere and can be at odds with long-established rules and principles in the making of numbers.
- In a do-it-yourself movement, professional quality standards are quickly lost; especially when the necessary education is lacking.

These trends and phenomena can combine to form a cocktail that poses a threat to democratic societies. All too easily, so-called 'alternative facts' are manufactured in the alchemy kitchens of populism to be held up (as equivalent) to scientifically based facts produced with high quality.

2.2. Data worlds

Apparently, there is no single answer to the question of how statistics relates to democracy, and certainly not a simple one according to which more statistics would generally lead to more democracy. What is obvious, however, and has been described many times scientifically [11–13], is the interaction between the state in its very different forms on the one hand and statistics on the other. Statistics and the making of the state have been closely married for two centuries and since the time of the Enlightenment.

Today, the situation is more complex and difficult to assess insofar as there are different *data worlds* [14] in which diverging ideas and conventions have emerged regarding the quality of information and the role of statisticians in today's society. "The concept of the *world* here emphasises objectification: the practice of action and the evaluative networking of things and people creates its own objectivity to which the practice of justification can refer, whereby *objects*, *qualities* and *categories* as well as their interlinking appear as *given by nature* and as a network like an *ontology* in social reality and thus appear situationally evident to the actors as a *world*."² Measurement, argues the *Économie des Conventions* [16–18], is necessarily grounded in conventions and norms, even if (or precisely because) they are not in fact visible or made to be transpar-

²Translation from source in German language 15. Diaz-Bone R. Die conomie des conventions – Ein neuer institutionalistischer Ansatz in der Wirtschaftssoziologie, 2009.

Table 1
Benefits and limitations of decentralisation and standardisation

Decentral statistical analysis (big data world)	Standardised statistical analysis (official statistics world)
<i>Pro</i>	<i>Pro</i>
<ul style="list-style-type: none"> – Flexibility in the design and choice of methods – Quality responsibility and sovereignty – Relevance for the individual information need – Competence through acquired factual knowledge 	<ul style="list-style-type: none"> – Homogeneity through standards – Comparability (time, region) and consistency with scientific frameworks – Efficiency, sustainability – Quality standards – Low level of statistical literacy required (equality)
<i>Contra</i>	<i>Contra</i>
<ul style="list-style-type: none"> – Heterogeneity of solutions – Costs – Risk of inequality due to digital and social gaps – Comparability and consistency with other analysis not guaranteed 	<ul style="list-style-type: none"> – Industrialised production, small room for customisation to individual needs – Fixed statistical programme and methods, slow adaptation and limited flexibility – Technocratic power limiting transparency and participation

ent. “Boltanski and Thévenot [19] emphasise that there are a multitude of different conventions and associated forms of justification/orders of value/worlds. While these claim to be a complete basis for action in their own right, they can never completely eliminate the competing claim of the other value orders to be equally relevant as a basis for action. Social situations are therefore mostly characterised by a plurality of value orders. This plurality can take the form of a conflict between conventions. Conventions then enable each other to question the other ontologies, i.e. to criticise (denaturalise) their qualities/values. On the other hand, forms of compromise between the conventions can occur, as they become possible especially in organisations.”³ The fact that communities differ in terms of their values, convictions, quality concepts and norms rarely plays a role when it comes to defining or contrasting their respective professional profiles. Rather, the focus is usually directed towards a distinction regarding the methods, technologies and tools used. In this article, however, we will be concentrating on the other characteristics that distinguish *data worlds* from each other.

Production of evidence

For our considerations, several such data worlds are relevant, namely first of all that of *big data* versus that of *official statistics*. Superficially, the two worlds differ in the genesis and ownership of the data; collection of mass data for commercial purposes here, targeted collection of data for government policy there. In the course of time, professional profiles have emerged in line with this, namely that of data science (strongly oriented towards predictions using algorithms and AI) and that of statistics (oriented towards controllable data

inputs and evaluation processes). It is interesting to note that it is not so much any methodological-conceptual differences that need to be bridged. Rather, they are two concepts and approaches to tasks of quantifying aspects of reality that are as complementary to each other as individual transport and rail transport are responses to questions of mobility. So, what essentially distinguishes the two is a decentralised organisation of statistical processing for the *Big Data world* on the one side and a centralised, standardised logic embedded in quasi-industrialised processes for the Official Statistics world on the other. It should be noted that the point here is not to evaluate one approach against the other, but to emphasise their complementarity (Table 1).⁴

Both are in fact based on a third, namely the *academic data world*. This academic data world (in universities and research institutes) develops new methods and technologies under the guiding principle of scientific independence.

In our context, it is interesting to reflect on the mutual influence on content, the networking of experts and the possibilities of successful (or unsuccessful) cooperation between the three spheres and communities presented. What is a promising combination of methodological challenge and prospects of success for a young scholar within the framework of the current scientific incentives? Does the (increasing) proportion of staff trained in mathematics, science or computer science in research and teaching influence the narrower or broader interpretation of the boundaries of the subject ‘statistics’? Is the financial support for methodological developments too focused on fundamental research in statistics and data sciences, while applied developments have no chance of success in the competition for limited funds? For all

³Translation from source in German language 15. Ibid.

⁴Source 20. Radermacher WJ. Standardisation and Statistics. The Survey Statistician. 2021; 84: 24-31.

Table 2
Attitudes and values in data worlds

Technical trends		
Data revolution	Big data, internet of things, AI	New opportunities, challenges and risks
Societal trends		
Management world	Evidence-based decision-making, neoliberal governance, technical experts and elite	Addiction to evidence, no major concerns on the quality of information
Civic data world	Social movements, NGOs critical views on the adequacy of indicators	Pro evidence, with different focus and content
Big data world	Do-It-Yourself, data scientists and 'amateurs'	Pro my own evidence
Populist data world	Post-truth politicians, supporters of conspiracy theories	There is no such thing as objective facts

these reasons, is the interest in empirical, descriptive statistics with its methodological questions (especially survey statistics) about to lose weight and interest in the academic data world? What does all this mean for the education of the coming generations of highly paid and qualified experts in business, research and administration?

Use of evidence

Likewise, divergent worlds can be observed in the way quantitative information is assessed and exploited on the so-called user side. The guiding principle of management (in business and administration) today is the doctrine of evidence-based decision-making: If you don't measure it, you can't manage it! In this *management world*, decisions are augmented with the input of facts, increasingly automated with the use of AI and in any case partially outsourced to the services of evidence provision. Actually, the guiding principle of evidence-based decision-making is trivial; who would want to oppose it? Its risks and side effects only become apparent on closer inspection, namely that the relationship between the process of measuring and that of decision-making can be influenced by feed-backs, that the boundary line between them is blurred and - not least - that the impression of a dominance of the data producers (statistics, data science, etc.), i.e. that of a technocracy, can arise.

The impression that scientifically based and statistically validated evidence can create is that there is no alternative for the decisions based on it. This gives rise to a resistance, first against (unpopular) policies, then consequently against the factual basis on which they base their argumentation. An essential element of populist politics is therefore the discrediting of the sources and suppliers of evidence.⁵ In a *populist data world*, the existence of neutral, objective and high-quality in-

formation is denied and replaced by so-called alternative facts that are subjective and, above all, in line with political goals and beliefs.

Another important variant in our context is that of the *civic data world*, in the broadest sense, forms of civil society engagement, i.e. social movements, non-governmental organisations (NGOs) that report on environmental, social or similar issues, based on factual evidence. These organisations aim to counter private and administrative forms of governance and to build agency and critical capacity around data and data transparency [14].

The social groups summarised in Table 2 each pose special challenges to official statistics and to data literacy.

3. Ingredients for a statistics education and culture

3.1. *The data value circle: From data to knowledge*

“The data value chain describes the evolution of data from collection to analysis, dissemination, and the final impact of data on decision making.” [21] The process of statistics, from their preparation, through the collection of data and their processing, to their dissemination and evaluation, is regularly presented as a linear sequence of steps, e.g. in the Generic Statistical Business Process Model [22]. Although these seem to be rather details, it is essential that statistics firstly do not start with the collection of data and secondly should not be understood as a single and unique process. Rather, the beginning of a statistic is characterised by a question, a need for information or a model from a theory that calls for quantification (unemployment, inflation, need for schools, etc.). In the design phase of a statistic, in which the theoretical terms are translated into quantifiable constructs, one often does not start from scratch. Instead, one orients on the existing and tries to learn

⁵People have had enough of experts <https://www.youtube.com/watch?v=GGgiGtJk7MA>.

from it and add improvements or extensions.⁶ Statistics are not typically individual gems, but are part of a larger informational collection or jewellery of diverse, related facts.

What we call ‘data value circle’ corresponds to the statistical investigation or statistical enquiry or data investigation cycle and the long-established advocacy by both professional statisticians and statistical educators that teaching must reflect, and be built around, this. It is an essential criterion for assessing an educational structuring model in statistics whether it takes this broader view or whether it reduces itself to the narrower, purely technical process. Katharina Schüller’s Data Literacy Framework [23] approaches the question in this comprehensive sense. The data processes in the narrower sense are integrated into learning cycles that focus on evaluation and (re)design.

In a production process – and that is what the production of facts and evidence is – expertise and skills of several kinds are needed at the different stages and steps of the process. In later stages of processing, it is important to have a reasonable understanding of the upstream supply chain, even if one has to take its outcome more or less for granted. This roughly outlines what users of public statistics results should contribute at their stage: a reasonable understanding of what information product they are being provided with, what its quality profile is, what they can (or cannot) use it for, etc.

3.2. *Competences, competence levels*

“Data literacy is the cluster of all efficient behaviours and attitudes for the effective execution of all process steps for the creation of value and/or decision making from data.” [23] With this broad description of data literacy, it becomes very clear that it is about much more than just being familiar with techniques and methods.⁷ Rather, competence in a comprehensive sense involves a combination of knowledge, skills and attitudes. Furthermore, it is important at what level such competences must be available, which in turn – comparable to language competence – depends on the context in which and the purpose for which such competences are to be applied.

⁶See GSBPM Trainings Material <https://statswiki.unece.org/display/GSBPM/GSBPM+Training+Materials>.

⁷One such narrow and not to be generalised approach for empowerment through literacy is, for example, the one presented by the DATA-POP Alliance 24. DATA-POP Alliance. *Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data*. New York: DATA-POP ALLIANCE; 2015.

3.3. *Critical thinking skills*

In the interest of empowering civil society with the aim of broad participation in evidence-based political discourse, it is not least a matter of being able to deal critically with statistics in addition to technical and methodological skills. This task can be approached in relation to statistics in general by analysing the characteristics of statistical literacy from the perspective of critical thinking [25]. With the eight critical thinking skills distinguished here, it should be possible to successfully develop a soundly based culture of data and to counter ‘crank’ science using sophisticated pseudo-statistics [26]. In the specific domain of civic, public and official data, the task of trustworthiness, transparency and institutionally based authority from the perspective of the users of statistics arises with particular urgency, far-reaching consequences for the democratic order and often without the existence of alternative data sources. What this means in concrete terms for statistics in the public sphere is explained in detail in the article “Literacy in statistics for the public discourse.” [27]

3.4. *Statistical and data literacy skills required in policy-making*

As briefly described in the introduction, policy has several roles and tasks in the processes of evidence-based decision-making. Just as there are data and facts for policy, it is conversely important to emphasise the different policies regarding the framework for data on the one hand and for statistics on the other [28]. Accordingly, different competences are also required in the political sphere, depending on whether someone is preparing decisions on the analytical side of policy advice with the help of evidence, whether it is a matter of communication or whether the legal-political determinants for actors and action in informational value creation are to be designed.

For the one side, namely the use of evidence for policy-making, the competences and literacy required for doing so can largely be referred to what has already been elaborated. More interesting, however, is the opposite direction of activities, namely the politics for data and the politics for facts. In this respect, qualities and competencies that generally apply to good management are (additionally) important. Deming has established his criteria for profound knowledge, according to which successful management is based on the following four competences: “Knowledge about variation, Psychology (psychology of individuals, groups, society

and change), Theory of knowledge, Appreciation for a system.” [29,30] In the field considered here, the system for which ‘Appreciation’ is required is about a comprehensive and deep understanding of the processes and interrelationships related to statistics as

- A science of variation, data and uncertainty, with the extraction and interpretation of knowledge involving constant querying of all aspects, including issues, data, models, assumptions and analyses.
- A (modern) technology, reliable data sources, efficient processing methods.
- An infrastructure that needs to be regularly maintained and modernised.
- A common language between producers and users.
- All framed by and based on values, ethics, governance.

4. Governance, institutions and trust: Data culture

Trust is built on knowledge and experience as opposed to *blind faith*.

Trust in democratic institutions, such as public statistics, is characterised by a mutually reinforcing relationship between *Wertschöpfung* (value generation) and *Wertschätzung* (value appreciation). Creating *Wert* (value) in the form of informational products and related services is the task of public statistics. Whether and how well this can be achieved depends not least on structural preconditions such as governance, budget, competencies, etc. *Wertschätzung* (appreciation) on the part of the users of these products and services is influenced by concrete experiences, but also by opinions, attitudes, values and not least by (statistical) literacy.

For public statistics, danger arises when either the conditions for *Wertschöpfung* (value generation) or the factors influencing *Wertschätzung* (value appreciation) are unfavourable. Recent experience from the Covid pandemic has shown weaknesses and risks on both sides: On the one hand, weaknesses in the coordinated and qualified provision of relevant indicators, on the other hand, considerable gaps in statistical literacy. As a result, trust has been lost. This conceals an essential implication: Of course, an appreciation of public statistics in the policy-making circles is a prerequisite for them to initiate appropriate programmes and measures. In the absence of statistical literacy and an understanding of the role of public statistics in the democratic process, greater trouble is inevitable. Winning trust back is an objective whose importance can hardly be overstated.

This can only be achieved through political initiatives and investments on both sides.

Let us summarise: There is a social and political dimension to literacy in data and statistics. Let’s call it *data culture*. If this important prerequisite is missing, then political programmes will be one-sidedly focused on the elements to which the public’s attention is devoted in the short term. Comprehensive knowledge of structures and forward-looking investments in public infrastructure will be lacking, value creation and appreciation will suffer, trust will be (further) lost.

“Official statistics have shown an excellent record in the role of trustworthy authority in the intersection of three fundamental rights: privacy (the right of a person to privacy), freedom of information (the right of a person to open and transparent information) and statistics (the right of a person to live in an informed society⁸.” [8] In order to be able to play the important role from the past also in the future,⁹ some (partly new) principles must be observed (Table 3).

Towards a data culture

Several targets are required to achieve the abovementioned objectives. Some of the most important ones are listed here:

- To understand the process chain of generating knowledge from data under today’s conditions with its sub-processes and their characteristic requirements and tasks.
- To develop an adequate structure of educational programmes that is geared to the different process stages in this value chain as well as actors, supposed to be competent at a respective stage and level.
- To impart not only the knowledge and skills but also the values and attitudes that are required at the respective process level in this data literacy training.
- To review the existing principles for ethics [31], good governance [32] and quality in statistics [33] so that they can also be relevant and guiding for data scientists.
- To offer easy-to-understand, yet solid forms for the communication of facts and figures.

⁸The right of a person to have access to and be a stakeholder in the provision of good quality statistical information intended to steer and govern a society.

⁹See also <https://hdsr.mitpress.mit.edu/pub/1g514ljw/release/3?readingCollection=016f5798>.

Table 3
Objectives for (official) statistics in a data culture

Statistics is key for <i>people empowerment</i>	Statisticians should be aware of data's power to provide information and, hence, knowledge
<i>Open data</i> fundamental for open societies	Statisticians should ensure open and transparent access to data and metadata, and monitor their actual use for information and knowledge
' <i>Datacy</i> ' key enabler for citizens	Statisticians should promote data literacy in society at large, and regularly monitor the levels of understanding
<i>Trusted smart statistics</i> is the future	Statisticians should continue to invest in methods, algorithms and a business architecture that enhance the quality of data for statistical services tailored to users' needs
<i>Participation</i> in the design, production and communication of statistics	Statisticians should foster a greater involvement of civil society in all stages and processes of statistical production
More influence means more <i>responsibilities</i>	It is the duty of statisticians to explore the link between statistics, science and society and to lead reflections on statistical and AI ethics

- To actively involve citizens in the design and production of statistics; co-creation and co-production can help to constructively bridge the gap between technically skilled experts in statistics production and the lay public.
- To scientifically process the interrelationship between statistics and the state [34], between data and society [18] in a sociology of the quantification [17] and of the politics of numbers [35].

Overall, it is about actively promoting a data culture. This goal goes far beyond conveying 'digital skills'. In fact, in many cases, corresponding activities have already been initiated and have produced concrete results. Such activities should be built upon and further networked and strengthened. The following actions (with reference to exemplary examples) are recommended to proceed and achieve the goal of a data culture:

- Data manifesto.¹⁰
- Standardisation¹¹ of a data literacy framework.
- Data literacy charter.¹²
- Sense about science¹³ type of activity focussed on data and statistics.
- Digital ethics¹⁴ as orientation for the review of professional ethics in statistics.¹⁵
- Citizen Science [36]:¹⁶ opportunities to involve

- civil society in the production of statistics.^{17,18}
- Facilitating access to data of public sources without compromising privacy protection; provision of scientific and public use files for research and training in statistics.^{19,20}

This is an extensive catalogue of demands, which presupposes the political impulses to be set in the corresponding direction. At the same time, however, it is also crucial to set the appropriate course in science, for example by promoting the necessary interdisciplinary research.²¹ Finally, it will also be crucial to motivate the institutions of official statistics and enable them to play an active role in this programme.

References

- [1] Gould R. Data literacy is statistical literacy. *Statistics Education Research Journal*. 2017; 16(1): 22-25.
- [2] Donoho D. 50 years of data science. *Journal of Computational and Graphical Statistics*. 2017; 26(4): 745-766.
- [3] Soma K, MacDonald BH, Termeer CJ, Opdam P. Introduction article: Informational governance and environmental sustainability. *Current Opinion in Environmental Sustainability*. 2016; 2016(18): 131-139.
- [4] Holt DT. Official statistics, public policy and public trust. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2008; 171(2): 323-346.
- [5] Davies W. How statistics lost their power – and why we should fear what comes next. *The Guardian*. 2017.
- [6] Wise AF. Educating data scientists and data literate citizens for a new generation of data. *Journal of the Learning Sciences*. 2020; 29(1): 165-181.

¹⁰<https://rss.org.uk/policy-campaigns/policy/our-data-manifesto/>.

¹¹<https://standards.ieee.org/ieee/7015/10688/>.

¹²<https://www.stifterverband.org/data-literacy-charter>.

¹³<https://senseaboutscience.org/>.

¹⁴<https://centri.unibo.it/digital-ethics/en/centre>.

¹⁵<https://unstats.un.org/unsd/statcom/53rd-session/side-events/documents/08022022-M-UNSC-Side-Event-IAOS-proposal.pdf>.

¹⁶<https://unstats.un.org/unsd/statcom/53rd-session/side-events/03022022-M-Engaging-citizens-in-monitoring-and-reporting-on-the-SDGs/>.

¹⁷<https://unstats.un.org/unsd/statcom/53rd-session/side-events/03022022-M-Engaging-citizens-in-monitoring-and-reporting-on-the-SDGs/>.

¹⁸<https://unstats.un.org/unsd/statcom/53rd-session/side-events/documents/10022022-concept-note-on-CGD.pdf>.

¹⁹<https://opendatawatch.com/about/>.

²⁰https://ec.europa.eu/eurostat/cros/content/public-use-files-eurostat-microdata-0_en.

²¹<http://arithmus.eu/>.

- [7] Royal Statistical Society. The Data Manifesto London: Royal Statistical Society RSS; 2019. Available from: <https://rss.org.uk/policy-campaigns/policy/our-data-manifesto/>.
- [8] Radermacher WJ. Official Statistics 4.0 – Verified Facts for People in the 21st Century. Heidelberg: Springer Nature Switzerland AG; imprint Springer; 2020.
- [9] Radermacher WJ. Medium [Internet]. Policy D, editor: Medium. 2021. [cited 2021]. Available from: <https://medium.com/data-policy/official-statistics-language-for-public-dis-course-part-1-294d31e56b6e>.
- [10] Radermacher WJ. Medium [Internet]. Policy D, editor: Medium. 2021. [cited 2021]. Available from: <https://medium.com/data-policy/official-statistics-language-for-public-dis-course-part-2-b9e912621523>.
- [11] Desrosires A. The politics of large numbers – A history of statistical reasoning. Cambridge Massachusetts: Harvard University Press; 1998.
- [12] Porter TM. Trust in numbers: The pursuit of objectivity in science and public life. Princeton, N.J.; Chichester: Princeton University Press; 1995.
- [13] Ruppert E, Scheel S. Data practices – making up a European people. London: Goldsmiths Press; 2021; 368.
- [14] Diaz-Bone R, Schrr S. Algorithms, conventions and new regulation processes. in: Democratic Frontiers. Filimowicz M, ed. London: Routledge; 2022; 23.
- [15] Diaz-Bone R. Die conomie des conventions - Ein neuer institutionalistischer Ansatz in der Wirtschaftssoziologie. 2009.
- [16] Diaz-Bone R, Horvath K. Official statistics, big data and civil society. Introducing the Approach of economics of Convention for Understanding the Rise of New Data Worlds and Their Implications. Statistical Journal of the IAOS. 2021; 37(1).
- [17] Henneguella A, Jatteau A. Sociologie de la quantification. Paris: La Dcouverte; 2021. 128.
- [18] Beaulieu A, Leonelli S. Data and society: A critical introduction: SAGE publications; 2021.
- [19] Boltanski L, Thvenot L. De la justification: les conomies de la grandeur: Gallimard; 1991.
- [20] Radermacher WJ. Standardisation and statistics. The Survey Statistician. 2021; 84: 24-31.
- [21] Data2X and Open Data Watch. The Data Value Chain: Moving from Production to Impact. Available from: https://opendatawatch.com/wp-content/uploads/2018/03/Data_Value_Chain_WR-1803126.pdf.
- [22] UNECE. Generic Statistical Business Process Model Geneva: UNECE; 2022. [Available from: <https://statswiki.unece.org/display/GSBPM/Generic+Statistical+Business+Process+Model>].
- [23] Schller K. Future skills: A framework for data literacy. Working Paper No. 53. Berlin: Hochschulforum Digitalisierung; July 2020.
- [24] DATA-POP Alliance. Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data. New York: DATA-POP ALLIANCE; 2015.
- [25] Koga S. Characteristics of statistical literacy skills from the perspective of critical thinking. Teaching Statistics. 2022; n/a(n/a).
- [26] Lee C, Yang T, Inchoco GD, Jones GM, Satyanarayan A. Viral visualizations: How coronavirus skeptics use orthodox data practices to promote unorthodox science online. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; Yokohama, Japan: Association for Computing Machinery; 2021. Article 607.
- [27] Radermacher WJ. Literacy in statistics for the public discourse. Statistical Journal of the IAOS. 2021; 37: 747-752.
- [28] Radermacher WJ. Governing-by-the-numbers – Reflections on the future of official statistics in a digital and globalised society. Statistical Journal of the IAOS. 2019; 35: 519-537.
- [29] Moen RD, Norman CL. Always applicable – Deming’s system of profound knowledge remains relevant for management and quality professionals today. Quality Progress. 2016; 2016(June).
- [30] Deming WE. Out of the crisis. Cambridge MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study; 1994.
- [31] ISI. ISI declaration on professional ethics. The Hague The Netherlands: InternationalStatisticalInstitute; 2010.
- [32] United Nations. Fundamental principles of official statistics. New York, 2014.
- [33] Eurostat. European statistics code of practice luxembourg: Eurostat; 2018. Available from: <http://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice>.
- [34] Saetnan AR, Lomell HM, Hammer S. The mutual construction of statistics and society. New York, NY: Routledge; 2011.
- [35] Mennicken A, Salais R, eds. The new politics of numbers. Cham: Palgrave Macmillan; 2022.
- [36] Haklay M, Fraisl D, Greshake Tzovaras B, Hecker S, Gold M, Hager G, et al. Contours of citizen science: A vignette study. Royal Society Open Science. 2021; 8(8): 202108.