Consumer-friendly scoring
Recommendations for action

SVRV Sachverständigenrat
für Verbraucherfragen

Subject matter and objective

Scores are numerical ratings used to predict or steer people's behaviour. These numerical ratings are usually calculated via algorithmic processes based on a broad range of data.

Scores are not the same as decisions. Rather, they are used to inform *decisions* made by the responsible actors, who take scores as the basis for their assessments (on a case-by-case basis, for example when approving a bank loan).

Scoring is by no means unique to the digital world. Traditionally, people have been "scored" not only in terms of their creditworthiness, but also in the grades assigned to them by educational institutions. The spread of digital technologies is opening up more and more possibilities for assessing and rating individuals. This has unlocked both risks and opportunities for consumers. Formalised scoring systems can be verified and audited; they have the potential to reduce the number of errors and discriminatory judgments in decision-making. In this respect, scoring has the potential to deliver results that are superior overall to informal decision-making processes steered by individuals.

Algorithms also have their origins in the analogue world, and can be utilised in all manner of contexts, from simple decision-making and mathematical processes to *artificial intelligence* and *deep neural networks*. Algorithms are to be found in more and more areas of our lives (e. g. personalised pricing in internet retail, search results ranking, in future: self-driving vehicles). This report focuses on the use of scoring algorithms in three fields: credit ratings, health, and vehicle insurance. Nevertheless, the present assessments and recommendations can be modified and applied to other areas, e. g. personnel selection procedures and employee scoring (people analytics).

The potential offered by scoring systems can be exploited to the full only when society's various legitimate expectations are met. Only then will consumers accept and benefit from these systems. *Consumer-friendly scoring* is the state that will be achieved when these conditions are satisfied. This will include applying data

protection in a way that prevents/minimises the risk of mistaken identity, and providing simple and effective ways for individuals to appeal their score. Protected characteristics (such as gender) may not be used as a basis for unwarranted discrimination – directly or indirectly. When scores are calculated for predictive purposes, the quality of the criteria applied and the reliability of the predictions must be demonstrated. Moreover, predictive powers of this calibre should remain stable across a variety of socio-economic groups. Furthermore, the communicated objective of a score should not be misleading to the affected individuals. The predictions made must correspond to the objectives of the scoring system concerned, and should not be applied frivolously to areas other than those for which the score was calculated. Above all, however, scoring systems must be comprehensible to those who are scored.

This report seeks to set out and explain the conditions required for a system of consumer friendly scoring, and advance proposals for how these could be implemented.

In order to do this, not all scoring services must be examined in equal depth. Rather, this report focuses on those scores that have a considerable impact on consumer well-being. Scores can affect people's lives in sensitive areas, such as their ability to obtain a loan and the conditions imposed by the lender, as well as their insurance premiums; furthermore, scoring may deliberately target health-related behaviour. These areas are particularly relevant. So far, there has been a lack of clear distinction between those and less sensitive areas (such as fraud detection in e-commerce). We believe that such distinctions should be made. This would allow a more reasonable determination of the amount and degree of regulation required for consumer-friendly scoring.

The following recommendations build on the four main chapters of the report. Chapter B focuses on definitions, as well as the structures and mechanics of scoring. It explores the state of international research and discusses relevant areas for action. This is followed by three chapters dedicated to an in-depth analysis. Two of these build on our own empirical studies, to which reference is made in the recommendations that follow. Chapter C presents the results of a survey of companies offering creditworthiness scores, scores for third-party vehicle insurance (telematics policies) and scores for health

insurance bonuses. Chapter D presents the results of a representative survey of the general public, which investigated knowledge of and attitudes towards scoring. Chapter E examines the scope of (and gaps in) the legal framework for scoring services, focusing on substantive law and supervisory competences. This chapter is also directly relevant to the recommendations.

1. Making scoring comprehensible for consumers

Problem:

There is a considerable need for education and information among the German public with regard to data, algorithms, corporate decision-making processes and how these interact. For example, loan decisions are subject to comprehensive scrutiny by supervisory authorities, which means that a number of recommendations contained in this report have already been broadly addressed. However, the population has little or no awareness of this. Uncertainty and confusion prevail. The need for information and education efforts involving all stakeholders cannot be overemphasised.

The findings of the representative survey of the general public (Chapter D) demonstrate in detail that knowledge of the criteria applied in creditworthiness scoring depends – in addition to the person's age – on whether they have applied for access to their own credit rating within the last five years. On the other hand, formal education per se does not appear to have a relevant impact on a person's knowledge of the scoring system. It is also interesting to note that the majority of respondents were against receiving notification of their own score; just under half of all respondents indicated that they would like to be actively informed (with differences in the intervals at which they would prefer to receive such information).

Recommendations for action:

1. The Advisory Council for Consumer Affairs recommends that data protection authorities operationalise the *comprehensibility requirements* set out in the GDPR (cf. Article 15 para. 1 letter h) for scoring and score-based business processes. Comprehensibility should be measured according to the standards relevant to the average consumer. Where scoring entails a level of complexity that is no longer comprehensible to the individual consumer, measures should be taken to ensure that scoring processes can be understood not only by supervisory authorities, but, at the very least, by consumer bodies and non-state actors as well.

2. Scoring services should release clear and comprehensible information for consumers about the main criteria used to score them and, preferably, how these variables are weighted. Trade secrets, of course, must remain inviolable. The definition of which variables are considered crucial for consumers cannot be left exclusively to lawmakers: this task should additionally fall within the remit of consumer organisations, or, alternatively, the "market watchdogs" of Germany's consumer advice centres. At any rate, full disclosure to supervisory authorities of scoring systems and their attributes is a must (see page 5 of the Advisory Council's *Digital Sovereignty* report).

3. Some members of the Advisory Council advocate further-reaching transparency. They believe that all scoring variables should be disclosed to the consumer and that the relative weighting of each component should be indicated in the calculation of the score. To this extent, any interests on the part of scoring services and users in maintaining secrecy would take second place to the consumer's interest in receiving information. At the same time, the trade secret of how a scoring system has been developed and programmed would be maintained.

4. However, disclosure alone will not necessarily give consumers a better understanding of how scoring works. This will require a variety of meas-

ures, which include: providing examples of consumer scores and how they are tiered according to different variables; the production of visual teaching aids (e. g. by consumer organisations); general efforts to raise scoring-related competence among consumers. Any assessments of how comprehensible scores are to consumers should be based not only on expert opinion but on *empirical evidence*.

5. Consumers already have a right to tailored and meaningful written information whenever they are scored (see Article 13 para. 2 letter f, 15 para. 1 letter h GDPR). However, this right has not yet been set out in *more concrete terms*. Companies, supervisory authorities and consumer organisations should work together to develop *standards for scoring services*, which would help guarantee relevance and comprehensibility. The Advisory Council further recommends informing consumers of how their personal score is to be interpreted against the distribution of score values among the population as a whole (e. g. does my score put me in the "upper third"?).

6. Prompt, free-of-charge notification should be provided – or at least offered as an option for consumers – in the event of *major changes to a person's score* (e. g. if the person slips into a lower category). Naturally, there are certain limitations to this: in order to register a change in score, scoring services would have to retain historical score values. There are many practical applications (such as fraud recognition or determining possible payment modalities) for which this option will not be available. At banks and insurance companies, scores are calculated on an ad-hoc basis. This means that no score history is maintained, and potential changes are not apparent at the time the next "event" is registered. This proposal can therefore be implemented only at institutions where data collection is ongoing, e. g. credit scoring services and the Federal Motor Transport Authority in Flensburg (with its "Register of Driver Fitness", which already sends out such notifications).

2. Fostering knowledge and competence

Problem:

Consumers do not know a lot about scoring. This cannot be blamed entirely on a lack of transparency on the part of scoring services and users. However, knowledge is key and will determine the ability of informed consumers to scrutinise and appeal their scores, as well as to detect errors in the system. Supervisory authorities, consumer organisations and their "market watchdogs" must ease the burden placed on consumers.

Recommendations for action:

As recommended in the Advisory Council's Digital Sovereignty report, NGOs, consumer protection organisations and consumer protection projects should provide education on basic issues related to scoring in all its manifestations, as well as on the use of scoring in specific fields of business.

1. For this purpose, the Federal Government should develop *information and discussion* materials as part of its digitalisation strategy for the current parliamentary term, with the aim of improving skills on the part of consumers, multipliers and decision-makers. The underlying principles and quality aspects of scoring, as well as forms and causes of unequal treatment are just as much part of this basic knowledge as the rights enjoyed by those scored.

2. Measures should be taken to foster the competence people require in order to take informed decisions concerning their participation in a scoring process. This includes having the skills to identify scoring services and seek alternatives, as well as to verify, assess (e. g. is the information relevant to the consumer disclosed?) and utilise such services.

3. Identifying and revealing discrimination

Problem:

Just like informal decision-making scenarios, scoring is never perfect. It always has the potential for isolated or systematic error. Unjustified discrimination is one such flaw, and is systematic, since a certain group of consumers is wrongly assigned, as standard, a higher or lower value than another. The market study (Chapter C) reveals, for example, that most credit agencies use geoscoring. Furthermore, a number of credit agencies score on the basis of attributes that fall under the General Equal Treatment Act, such as sex and age. In the case of telematics-based vehicle insurance, certain groups of consumers such as shift-workers or city dwellers could be at risk of disadvantage because of factors beyond their control, such as time of day or location.

We have established shortcomings in reconciling anti-discrimination law with the General Data Protection Regulation and, above all, in enforcement. First and foremost, these enforcement problems arise due to difficulties in sourcing empirical evidence as proof of algorithmic discrimination against certain groups or individuals. In this context we should not fail to acknowledge that it is sometimes difficult or impossible to compare score distributions among different protected groups because there are data protection provisions in place which may prevent the collection and retention of information concerning certain attributes (see Article 9 of the GDPR).

Recommendations for action:

1. The Advisory Council for Consumer Affairs recommends that consumer information rights, as set out in Article 15 para. 1 letter h of the GDPR, be strengthened. In particular, consumers should be able to ascertain how scores are distributed among different groups with different protected attributes (to the extent that this can be established by the services themselves). This will allow consumers to provide evidence of algorithmic discrimination.

2. The Advisory Council also recommends strengthening the position of supervisory authorities (see recommendation 7).

3. Furthermore, it recommends that associations be given the right to pursue representative actions in cases of discrimination through scoring.

4. Ensuring that non-telematics based options remain available

Problem:

The market study for this report demonstrated that the potential for extending telematics-based tariffs in health and motor vehicle liability insurance is there. This would signal the use of online information on health-related and driver behaviour. The technology available for capturing further data is advancing by the day (eCall technology, connected vehicles, wearables). Providers will potentially be interested in extending their datasets to include more and more attributes (e.g. electronic patient files in the healthcare sector).

However, placing certain policy holders at a disadvantage by hindering or denying them access to reduced, behaviour-based tariffs – e.g. on the grounds of physical impediment – may jeopardise the principle of solidarity. The gap between insurance tariffs may widen as a result of differences in individual status, and freedom of choice may be restricted (e.g. for those opting for higher data protection standards). While Germany's social security schemes effectively preclude core violations of the solidarity principle, one might still question the bonus programmes offered by statutory health insurance providers.

The survey of the general public revealed that there is little acceptance overall for use of the behavioural attributes that would form the basis of scoring in the presented areas of vehicle and healthcare insurance.

Recommendations for action:

1. The Advisory Council for Consumer Affairs recommends the introduction of legal guarantees to maintain telematics-free options for those seeking insurance (especially motor vehicle liability insurance and health insurance). In particular:

2. Policyholders who do not use telematics-based tariffs may not suffer substantial disadvantage compared to the holders of telematics-based policies.

3. Most members of the Advisory Council for Consumer Affairs believe that telematics policies should be self-financing and should not be offered at the expense (even indirectly) of policyholders who do use telematics. Since solidarity objectives are relevant particularly in health insurance, steps would need to be taken to prohibit cheaper telematics tariffs that exist only because they attract policyholders with above-average health and do not significantly reduce the expenses incurred by insurers.

5. Ensuring score quality

Problem:

German law requires the use of "recognised mathematical-statistical procedures" (section 31 of the Federal Data Protection Act), but does not provide any further specification of what these are. The GDPR is no more specific. The use of mathematical procedures such as "multiple regression" does not guarantee that the resulting score will automatically be of high quality. Furthermore, complicated and opaque procedures are not necessarily superior to simple, comprehensible rules – or indeed pure chance. For this reason, providers must guarantee that their *scoring procedures meet high standards of quality* – not just in terms of methodology but also with regard to their objective. "Health scores" are problematic, for example, if little importance is attached to highly damaging forms of behaviour such as

smoking, but activities with unproven health benefits are given a higher weighting. A score must measure the attributes it purports to measure, while at the same time meeting the criteria set forth in these recommendations.

Recommendations for action:

1. The Advisory Council for Consumer Affairs recommends that ambitious quality principles be developed on the basis of best practices. This should be based on existing quality assurance initiatives for algorithmic processes. These quality principles should be developed and updated (drafted, implemented, monitored) on a collaborative basis by industry, supervisory authorities, consumer organisations and the market watchdogs of Germany's consumer advice centres.

2. Scoring services operating in sensitive fields should be obliged to file information with supervisory authorities that is verifiable in detail and reveals the high quality of their procedures. Only then will it be possible to test scores for consumer fairness. This obligation would apply to scores which use statistical measures to predict behaviour (e. g. false positive rates, hit rate, gini coefficient, area under the ROC) for the population as a whole and for relevant population groups (by sex, age, education etc.). This would also make it possible to identify discrimination and cases of questionable score quality.

3. As the situation currently stands, scoring procedures that pursue objectives which have not been appropriately identified to the public are prohibited by law. In addition to the role of supervisory authorities (see recommendation 7), consumer organisations or the market watchdogs of Germany's consumer advice centres could also apply their expertise and contribute to uncovering "falsely labelled" scores as well.

4. The use of proxy variables, as for example in geoscoring, requires special justification (there must be a causal connection!) and must be subject to the scrutiny of the relevant supervisory author-

ity. The use of proxy variables should be minimised. Where proxy variables are used, plausible reasoning must be given as to their substantive connection with the target variable.

6. Ensuring data quality

Problem:

The market study shows that all credit agencies offer consumers the possibility to correct information, should they find that data held on them is inaccurate or no longer current. In addition, best practice is provided in the study by those companies which (at least according to the information they provide) continuously monitor the data they have, and implement deletion periods in order to maintain the quality of their data stock.

Data quality is part of data governance, i.e. a framework that more and more companies and organisations are putting in place. This emphasises the significance of data (and not just personal information) in today's world, and reflects the need for a responsible and commensurate approach to handling the commodity of our age. Data governance involves defining responsibilities for all decisions concerning data and data use, as well as the relevant security and infrastructure.

Data quality is not a single factor in itself. Rather, it is made up of a number of components, e. g. the information used must be accurate, complete and up-to-date. In particular, data quality is a relative term. Data quality is high when data are *fit for use*. Requirements will be differently defined, depending on the type of application. High-quality scoring requires a very high level of data quality.

On the one hand it is vital that scoring processes are developed on the basis of data that meets the highest possible standards of quality. Data must be highly accurate, sufficiently up-to-date and available in the type of quantities required for estimations to be made. Furthermore, development and training data must be representative.

The data must reflect the composition, structure and behaviour of the group of individuals to whom the scoring process is to apply.

Even scoring processes of the highest quality will deliver inaccurate or even false results, if the available input data are of low quality. For example, the market study tells of cases of telematics insurance where false journey data were generated because maps were out of date, or there were technical problems with GPS signals or data transfer. For creditworthiness scoring, the main focus in guaranteeing data quality is on ensuring correct entity recognition, as well as ensuring that data is accurate, complete and up-to-date.

Recommendations for action:

1. When developing scores, a sufficient level of data quality must be ensured and documented for supervisory authorities.

2. Scoring services and users should enter into voluntary commitments to improve their data governance, in particular their data quality management, in accordance with the standards set in the quality principles.

3. In applying the procedure, measures must be taken to ensure that data is accurate, complete and up-to-date.

4. In its report on *Digital Sovereignty*, the Advisory Council for Consumer Affairs already outlined the option of a *data dashboard*, which would allow consumers to scrutinise their own data. This would facilitate consumer-oriented data management. The Advisory Council reaffirms its recommendation that this option be explored. Such explorations should cover current developments in the area of secure identity management via blockchain-based systems, which allow consumers to manage their own identity data securely and definitively.

5. The Advisory Council recommends that research be conducted promptly to appraise and, where applicable, improve the quality of data used in

relevant scoring processes, with a particular focus on entity recognition. Where necessary, improvements should be made via statutory provisions. Measures must be taken to ensure that a score calculated for a certain person is correctly assigned to that person. The duty for providers to inform individuals that they are being scored (see recommendation for action 1) will serve to minimise the risk of identity mix-ups.

In this regard there is clearly a conflict between the interests of scoring services and users, on the one hand, and data protection interests on the other. For this reason the Advisory Council recommends that the Federal Government's Data Ethics Commission discuss ways of improving entity recognition and develop concrete recommendations.

7. Improving oversight

Problem:

With current staffing levels and the technical resources available to them at present, supervisory authorities are far from able to fulfil their oversight duties in the areas (and on the scale) specified in our report. Legal know-how will not suffice. IT and statistical competences are also required. This is particularly true where the recommendations of this report would entail additional tasks for the authorities concerned.

Recommendations for action:

1. The Advisory Council for Consumer Affairs recommends that the Federal Government explore whether a *digital agency* (see the Advisory Council's report on "Consumer Law 2.0") could act as a competence centre to assist supervisory authorities in exercising their mandates. This might consist, for example, in setting up a federal institute as a centre of method expertise for quality assurance, which

could also be used for "non-digital" purposes.

2. The responsible supervisory authorities should be put in the position (both structurally and in part through salary improvements for specialists, especially in statistics and IT) to perform the aforementioned tasks. Developments at the Federal Financial Supervisory Authority (BaFin) over the last few years could serve as good practice. The responsible supervisory authorities should be granted the considerable financial resources required for them to perform the aforementioned additional tasks and test concrete scoring services.

3. To ensure that the present recommendations are promptly implemented, the Advisory Council for Consumer Affairs proposes the creation of a task force at the level of the Federal Government (for example at the Federal Chancellery) in order to develop guidelines for the elaboration of quality principles on the basis of existing procedures (e. g. at BaFin). This task force should be set up immediately after the Data Ethics Commission has finished its work.

8. Preventing "super scores"

Problem:

By 2020 the Chinese government is planning to assign everybody a *social credit score*. This "super score" is currently being developed by commercial enterprises using cutting-edge computer technology, and will serve to measure trustworthiness in financial, legal, social and political affairs. Citizens with a high score are to receive certain benefits, similar to those granted in a frequent flyers' programme. Those with a lower score would not be allowed to use express trains, for example, or send their children to the best schools. This super score is designed to improve moral comportment on the part of the population, reduce corruption and create "harmony".

Although the state introduction of a social credit system is not foreseeable in Germany, data traders and dataset de-anonymisation capabilities create new potential for data from different areas of people's lives to be collated into one database and at a single company for commercial purposes. This could even occur on a voluntary basis, if people are promised individual advantages for agreeing to their data being centralised. In the case of insurance companies offering cover in a number of areas, this potential is very real. It is at these sorts of companies that datasets could be merged and analysed in context. Those who live healthier lives could, for example, be offered cheaper income protection insurance.

Our survey of the general public found that up to a third of the adult population in Germany can imagine a merger of driver and health scores; the respondents indicated that this would be "fair", or that they would benefit personally.

Recommendations for action:

The Advisory Council for Consumer Affairs recommends that developments in China and in other countries which are experimenting with "super scoring" are closely followed and analysed. In particular, public debate is required on the change in social values and structures that such systems entail.

The development of "super scores" by international commercial actors may also have an impact on Germany. Lawmakers and supervisory authorities should prepare for an examination of whether measures can and should be taken to ensure that "super scores" cannot be offered commercially in Germany.

The Advisory Council recommends that an examination be carried out into the extent to which existing instruments (especially purpose limitation and the "no tie-ins" rule) contained in the GDPR may also be used to prevent "super scores".

Advisory Council for Consumer Affairs

The Advisory Council for Consumer Affairs is an advisory body of the Federal Ministry of Justice and Consumer Protection (BMJV). It was set up in November 2014 by the Federal Minister of Justice and Consumer Protection.

The Advisory Council for Consumer Affairs is tasked with using research findings and drawing on the Federal Ministry of Justice and Consumer Protection's practical experience to help shape consumer policy.

The Advisory Council is independent and is based in Berlin.