Smart Data

Data as an economic asset

European data economy or rights to data?

Legal information

Published by:
Smart Data Research Department
FZI Research Center for Information Technology
Branch office Berlin
Friedrichstr. 60, 10117 Berlin
www.smart-data-programm.de

Editing and concept design
Legal framework specialist group in the
Smart Data Research Department

Final editing and layout
LoeschHundLiepold Kommunikation GmbH

Status
September 2017

Printed by
WIRmachenDRUCK, Backnang

Illustrations
sdecoret – Fotolia.com (title), Nataliya Hora – Fotolia.com (S. 9), xiaoliangge – Fotolia.com (S. 10), 75tiks – Fotolia.com (S. 11), Gorodenkoff – Fotolia.com, (S. 12), nd3000 – Fotolia.com (S. 17), alphaspirit – Fotolia.com (S. 22), Björn Wylezich – Fotolia.com (S. 40), Olivier Le Moal – Fotolia.com (S. 41), by-studio – Fotolia.com (S. 42)

Supported by:

Federal Ministry
for Economic Affairs
and Energy

on the basis of a decision
by the German Bundestag

Content

Foreword

An economic consideration of data as one of the central assets in the context of big data, Industry 4.0 or the Internet of Things creates major challenges for the legal system, and the introduction of a consultation process on the creation of a "European data economy" by the EU Commission now also indicates its growing importance in the European dimension. Should information be a freely available public resource on principle — at least as long as free access is not hindered by data protection law, copyright law, sui generis database rights, the protection of trade secrets or criminal law? Could the creation of a general "right to data items" over and above the existing rights provide extra stimulus for the future structure of the data economy? Or will the typical legal consequences of exclusivity rights prove to be largely counterproductive? How could the design for competitive claims for access to data and platforms and any supplementary contractual provisions be organised to create a coherent overall legal framework — which is worth promoting from our point of view — and yet take data protection requirements into account?

This topic has been intensively discussed in the legal framework specialist group. As an interim result we present five proposals which were agreed in the specialist group.

We would like to express our thanks to our authors Dr. Alexander Duisberg and Mr. Patrick Bunk who have supported this publication with solid scientific content and valuable experience. In the first article, current opinions in the legal discussion are analysed and the existing protective framework for data collections in copyright and unfair competition law is outlined. The second article focuses on the technology of text and data mining, which is very relevant in the context of smart data, and its conflict with copyright law.

The final section offers brief explanations of the main concepts and aspects of the legal framework related to "Data as an economic asset". Non-personal data certainly do not exist in a legal vacuum. Even in addition to the data protection laws which are much discussed in the development process for the upcoming General Data Protection Regulation (GDPR), there are important legal questions which could influence the use of data. For example, the use of data which form part of a protected database, a text or image work or a media production may only be permissible with the approval of the relevant author or creator. And data may also be legally protected as company or trade secrets. It is also worth noting that data interception and espionage are liable to criminal prosecution.

PD Dr. Oliver Raabe und Manuela Wagner
Accompanying research for the technology programme "Smart Data — innovation from data"

Five postulations for the European data economy

The legal framework specialist group of the accompanying research department for the technology programme “Smart Data — innovations from data” consists of representatives of the 16 flagship projects of the technology programme and experts from the political, economic and science sectors. In regular workshop meetings of the specialist group, these experts discussed the creation of a European data economy and formulated five central postulations.

Impulse 1: Access to anonymous machine-generated data should be improved

This year, the EU Commission started a consultation process on the question of how a European data economy can be created.[1] The central theme is the improvement of cross-border data exchange. On the basis of the insights gained in the pilot projects, the majority of participants in the legal framework specialist group in the smart data research department agree with the goal postulated by the EU Commission that an improvement in access to anonymous data is sensible:

If machine-generated data are shared, reused and aggregated they can create added value, become sources of innovation and pave the way for a variety of business models.

The EU Commission defines machine-generated data as follows:

Data are generated by machines without any direct human intervention in the course of computer processes, applications and services or by sensors which receive information from virtual or real devices or machines or from a software program.

The commission proposes the following possible ways to create incentives to share data and thus improve access to anonymous machine-generated data:

- Preparation of EU guidelines on the legal situation in the member states
- Promotion of the development of technical solutions for reliable identification and the exchange of data
- Standard contractual clauses to create legal certainty for exchange of data
- Creation of a right of access to non-personal data in the public interest or for scientific purposes
- Creation of a right of the "data producer": the owner or possessor of the machine would then be entitled to use non-personal data and to either permit others to use such data or to exclude them from such use.
- Creation of access rights in return for payment

The majority of the participants in the legal framework specialist group see potential in the promotion of technical solutions and the creation of access rights in the public interest and for research purposes. The members of the specialist group tend to oppose the concept of ownership rights for data. The question of whether data access can actually be promoted by an owner-like legal status of the "data producer" must be critically scrutinised, and it must be determined whether such a status would pose the risk of rising transaction costs and a strengthening of the lock-in effect. An (exclusive) ownership of data would also be problematical from the perspective of freedom of opinion and information, and it could lead to an unintended monopolisation of information. In addition, there are likely to be questions about the delineation of authorship. Data are normally created in an interaction between machines and people, or between machines and other machines or their environment, so there are likely to be challenges if the data need to be allocated to a single "data producer". The data may constitute personal data if the data also permit references to identifiable natural persons, so there is likely to be an overlap with the provisions of data protection law.

The greatest challenge will therefore probably be to ensure a legal basis for access to data under non-discriminatory terms and conditions. So the mechanisms of unfair competition law may need to be used to create incentives for data exchange beyond the boundaries of business enterprises and national borders. Standardisation and ensuring interoperability will probably be the fundamental technical requirements for a functional data exchange system.

Some smart data projects have the goal of creating open platforms for the exchange of data across company boundaries. The challenges here include the need of companies to protect their company and trade secrets and to ensure that they do not pass on any personal data without legitimation. The development of intelligent filter mechanisms, anonymisation tools and concepts for data access control and usage could provide solutions for these issues.

The smart data research department is involved in promoting a cultural transition towards an open data economy for a smart society.[2] The concept of "open data" stands for technical and legal openness, i.e. the data must be available in a machine-readable and standardised format and must be suitable for use without any legal restrictions (i.e. the data must be general accessible and not subject to any unreasonably restrictive licensing provisions).

Impulse 2: Development of technical solutions for reliable identification and data exchange should be promoted

Data can be reproduced to an unlimited extent, so any genuine control of the use of accessible data can only be facilitated with technical solutions which permit the transparency, traceability and identification of the data sources. Licence models and open data concepts both initially need standardised logging systems, interfaces and data formats. Often, data are not accessible without licensing requirements or technical limitations. If data are available with varying granularity and in different formats or have already been provisionally interpreted, this fragmentation makes it more difficult to analyse data for the whole of Germany. The development of data trustee concepts and the standardisation of machine-readable and free formats, especially for data in the public sector, would be possible solutions here.

To create confidence in the system and permit statements about the quality of the data, it may also become necessary to define reliable protocols, which are standardised as far as possible, for the seamless identification of data sources.

Open, standardised and well documented application programming interfaces (APIs) can promote the establishment and development of an ecosystem for application and algorithm development and thus enable access to data owned by business companies or public authorities. To ensure that this access complies with data protection law, the development of anonymisation tools and testing procedures needs to be promoted at the same time and supported by technical guidelines.

Impulse 3: Future solutions should minimise any lock-in effects

A lock-in effect occurs if the expected costs to change supplier effectively prevents a change of supplier. Such barriers to change can on the one hand be used deliberately to encourage customer loyalty, but on the other hand they could pose a market entrance obstacle for small new competitors.

In the establishment of a European data economy, the different negotiating perspectives of companies with great market power, less powerful companies and private consumers need to be taken into account. Lock-in effects especially need to be prevented for small and medium-sized companies, start-ups and private consumers.

The provisions of European and German unfair competition law currently only come into play if the abuse of a dominant or powerful market position is ascertainable. The relevant questions of market delineation, market concentration and the thresholds for abuse must be adapted to take the change towards more versatile data-based markets into account.

The 9th revision of the German Restrictive Practices Act, which came into force on 9 June 2017, aims to create a modern unfair competition law in the age of digitisation. The reform especially aims to address the scale factors and network effects based on data, which can lead to market concentration, and to improve consideration to the question of access to data which are relevant to competition. This should enable the competition authorities to evaluate the market position of a company and any possible abuse not only based on revenue, but also take the changing Internet-based data services into account.

Another discussion point is the question of whether there should be a general right to data portability, comparable with Article 20 of the General Data Protection Regulation.

Impulse 4: Right of access to data in the public interest or for scientific purposes

Research in the area of "Smart Data" is often based on non-personal data from the company context, and researchers require access to such data in order to generate new innovation-enhancing findings. Similarly, the functionality of the public sector can be improved by an analysis of statistical data. For example if statistics offices were granted access to business data, this could reduce the amount of work involved for business participants to comply with any necessary reporting obligations. The resulting optimisation of the infrastructure could then have a positive overall effect on Germany as a location for business.

In the creation of a right of access, care must of course be taken to address the problem that the dividing line between personal data and anonymous data is in a constant state of flux and may change over time due to the addition of extra knowledge or the improvement of analytical methods. A right of access which gives due consideration to competition law should therefore achieve a balance between the conflicting interests and mainly focus on non-discriminatory access and on interoperability. This should be established in coherence together with the planned General Data Protection Regulation.

At the same time, the interest of business companies in preserving their company and trade secrets must also be taken into account. In addition to the protection of personal data, anonymisation methods could for example be used to remove any mention of the company from a data record.

Impulse 5: Creation of a non-discriminatory legal framework for text and data mining and for webcrawling

In the proposal for a directive on copyright in the Digital Single Market[3], the EU Commission defines text and data mining as:
"any automated analytical technique aiming to analyse text and data in digital form in order to generate information such as patterns, trends and correlations".

The concept of "webcrawling" stands for the automated evaluation of online sources in the world wide web, usually by means of text and data mining technology.

Due to copyright, the legal protection of databases or protection of press publications concerning digital uses, the rightholders are entitled to exclude reproductions and/or publication of their work under the respective legal conditions. Thus, text and data mining can also be excluded in case this requires a (premanent) copy or the publication of original data (e.g. snippets). Allowing the use in exchange for a licence fee is in principle at the free discretion of autonomous private market participants, so there is no compulsion to make such contracts unless a refusal to do so would be an infringement of competition law. Large market participants (such as the search engine "Google") thus often have better conditions (sometimes even free licences) because publishers, for example, are dependent on Google's indexing, and this means that providers with less market power are discriminated against. One danger of this situation is that publishers may demand exclusive contracts (or a guarantee that no comparable licence terms will be agreed with certain competitors), so competitive relationships might be moved to secondary markets, to the disadvantage of the webcrawlers.

It is controversial whether pure information extraction by means of text and data mining should even constitute a form of use that can be licensed. Unlocking new knowledge on the basis of data available on the world wide web could also be understood as a new business model for data refinement. The fundamental question here is whether authors or information aggregators such as press publishers should receive a share of the value created by data refiners. On the one hand there is the expectation of data value chains based on license-based business models. On the other hand, there is a danger that exclusive protection could hinder progress and the exchange of knowledge, and that this could lead to an information monopolisation. Therefore, the question of whether information itself should remain common property is also relevant in this context, as well as the debate about the relevance of "data ownership".

The key central question is therefore how legislation in the area of "Smart Data" can create effective freedom from discrimination. This problem would not arise if there was a general permission for text and data mining which would permit temporary acts of reproductions exclusive for the extraction of information. Alternatively, mechanisms would need to be established under competition law to permit access to information under non-discriminatory terms.

Analysis of the current discussion on the “right to data”

A decisive factor in creating a European data economy is access to and the usability of non-personal data. The current legal debate revolves around the creation of “data ownership” with the goal of increasing the marketability of data. But it must also be taken into account that data ownership implies an unambiguous allocation of a data item to a specific owner and a clear distinction between data personal and non-personal data. The creation of exclusivity rights could therefore lead to difficulties in demarcation and to unintended legal consequences.

The following article deals with the legal derivation of such a right to single data items and the associated consequences. Besides there is already a legal protection of databases. The result of this consideration shows that the creation of a right to single data items is unlikely to provide an adequate solution.

“Data sovereignty and the right of the database maker” – the right in individual data items vs. rights in data collections

*Dr. Alexander Duisberg, Bird & Bird LLP**

I. Objectives and context

This working paper aims to describe the legal requirements and parameters for the transactional treatment of data as an economic asset under existing law (de lege lata) and under future law (de lege ferenda). The topic can be considered at two levels. The first level is the question of the legal attribution and right of disposal for the individual data item, the second relates to rights of disposal, access rights and treatment rights for data collections. The first step in this consideration will be to outline the state of the discussion about the concept of “data sovereignty” (deliberately avoiding the commonly used term of “data ownership” because it is legally misleading) (see section II.). This will be followed by a consideration of the rights of the database maker, which is currently the major legal instrument at the centre of the discussion (cf. section III.).

II. Data sovereignty — the state of the discussion

1. Preliminary considerations: “open” and “shared” versus proprietary data domains

The discussion about “data ownership” or other proprietary “rights to data” captures the imagination of lawyers. But it must be noted that our legal system – like practically all other current legal systems – does not recognise any absolute and exclusive “ownership” of data or datasets as such, neither under property law nor in any other form (Section 903 sentence 1 and Section 90 of the German Civil Code/BGB). This means that all statements in the current discussion are focused on possible future law (de lege ferenda) and the question of whether there should be a concept of “data ownership” under civil law.

1.1 Socio-economic factors
This question goes far beyond the question of whether and how such an absolute data right could be justified – in fact it is deeply socio-economic in character and cannot be fully clarified by purely legal considerations. The main goal in this respect is to move as far as possible towards an open, innovation-oriented legal culture and to examine whether possible exclusive rights in data could actually hinder innovation.[4] Just like the open source approach in the world of software development and software applications plays a decisive role for the innovation, scaling and growth of entire ecosystems (for example in the expansion of the app economy), we can equally well imagine – at least in theory – that an “open” or “shared” approach could become a central key to the success of certain models in budding the data economy. Amongst other things, the Recitals of the Public Sector Directive (“PSI Directive”) underline exactly this point.’[5]

1.2 Factors under competition and antitrust law
The associated questions about technical standards, open platforms, regulated non-discriminatory access and interoperability[6] are partly related to questions of competition and antitrust legislation both under existing law (de lege lata) and under future law (de lege ferenda) (cf. the separate considerations of the legal working group). They express and reflect the factual reality that there are a number of proprietary data domains which are owned and controlled by individual companies or groups of companies – and in some cases data monopolies and oligopolies. The legal answers to these questions are currently incomplete, but in the framework of this discussion they need not be discussed in detail. But they will become significantly more important in the future.

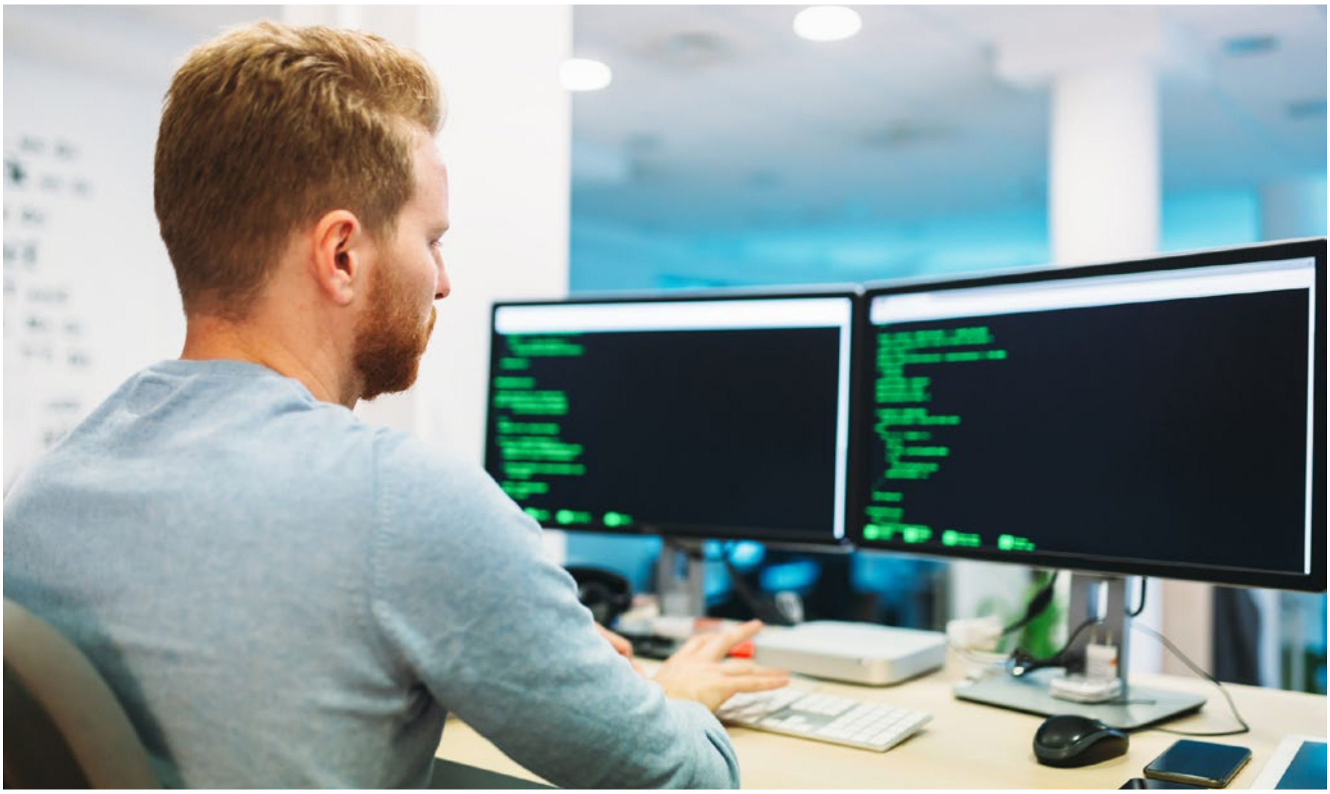
1.3 Linear value creation chain versus digital ecosystem
At the same time, there are statements to the effect that “dominion over data” and the “fight for data sovereignty” are one of the factors, or even the decisive

factor, for the development and establishment of value creation chains in the digital economy. This value creation is no longer mainly linear in structure. Rather, in the process of digitisation we are moving towards value creation in data ecosystems, where the economic gain for the individual participant is defined by the level of networking, and its evaluation by the other participants in the network.

The party that "has the data" thus determines the rules by which the other users and data exploiters must act. In this context, aiming for "data sovereignty" is best understood as competing for the meta data.[7] On the one hand, collecting, accessing and evaluating meta data provides a time advantage and first access to information, which can consolidate the participant's own value creation chains and business models. On the other hand, controlling the meta data allows the participant to secure its own company and trade secrets which are contained in the meta data. Against, third party access. In light of these aspects, we can see a need for guidance and possibly regulation (in case self-regulating market forces do not accomplish) in two areas – which have in fact been articulated in the industry – (i) introducing appropriate protective mechanisms for "proprietary" data content, and (ii) legal clarity as to how to suitably structure value creation chains and data ecosystems.[8]

1.4 "Information" as an asset to be protected

First of all, the conceptual and material difference between "data" and "information" is relevant here. A simple data point does not have any significance as information (e.g. "sensor data point 19"). The information content of a data point can only be determined by means of its assigned reference features (meta data) and the context of this data point in relation to other datasets.



Against this background, we may ask whether this need for protection and legal clarity (and regulation) should apply more to the data as such, or to the potential of the information which is embodied in the data or can be obtained by evaluating the data. The latter option could indicate that legal protection should be applied as the "right level" – i.e. not so much at the level of the individual data point or data set as such, but rather at the level of the contextuality of data. In other words, the question is why the focus of legal protection should be placed on the protection of individual data points (the "bare raw data") if the quality of data that might be worth protecting actually only arises from the contextuality (especially meta data and/or the correlation with other datasets) and, building on that, the resulting content deriving from the data.[7]

This could lead to the conclusion that an approach which does not define proprietary data, or even deliberately omits the concept of proprietary data in favour of a generally free approach – although this would not necessarily also lead to an obligation to permit open participation or "sharing" – could be consistent with protective mechanisms which apply (only) at the level of the context or the potential information – thus including the meta data which describe the quality of the individual data points.

As a consequence, legislation approaches which do not by default tackle the data point level are likely to provide greater flexibility and thus facilitate innovation, scaling and multilateral value creation through platforms and ecosystems, as well as by (possibly) adopting "shared" or "open" models. The opinions outlined in chapter II do not essentially deal with these considerations. Therefore the following ideas – without this reference – are only immanently summarised and briefly evaluated.

1.5 Validity of special rights

Of course, all the approaches discussed here are subject to the proviso that special rights to data which arise from the nature of their content (e.g. copyright and intellectual property rights to music files) are not affected. If we were to assume the existence of proprietary rights to "data as such", this naturally does not invalidate the existing special rights to the content of the data. This will be clearest in the following if we think of "pure raw data", such as data collected by sensors in an industrial environment (e.g. machine measurement data).[9]

On the contrary, such special rights and the question of data rights should be considered separately. The example of copyright shows that the underlying goals of the special rights cannot simply be transferred to the question of "data sovereignty". According to the classical legal theory of "copyright", copyright law is designed as protection from unauthorised reproduction. But the originator / author does not typically have the capital and the business means to publish and duplicate the work himself, so copyright law aims to enable him to have an economic share in his work. This leads to linear value creation chains (author-publisher-bookseller; composer-music publisher-record company-broadcaster/concert organiser etc.). Compared with this system, the direct costs of the reproduction and sale of data and digital content in an age of digitisation have in practice decreased to threshold costs of zero.[10] The underlying principle of copyright law is only partly suitable – if at all – as a means to justify, reflect or protect (non-linear) value creation chains in the age of digitisation.

1.6 Considerations under data protection law

From the perspective of data protection law, there is another problem if individual data items also have the quality of identifiable personal data, or could achieve this quality in the course of big data processing. Under Section 35 (2) sentence 2 of the Federal Data Protection Act (BDSG), the affected person is entitled at any time to demand the deletion of such data, and this right is constitutionally anchored in the personality rights of the individual and the right of informational self-determination. This brings up the question of what ownership rights to individual data items are worth – even though they are a protected basic right under Article 14 of the German Constitution – if they can be withdrawn and destroyed at any time by an arbitrary unilateral declaration of an unknown third party.

2. Proprietary approach

Several commentators in the literature advocate the creation of ownership rights or property-like exclusivity rights to data, but they do it with a variety of theoretical approaches and reasons. The reasons given for such rights especially involve creating incentives for companies to collect, store and share data and thus to develop their own data market.[11] It is suggested that such a market is not attractive without a clear legal allocation of data because data lose their value as soon as they are known to a third party.

2.1 Right of ownership by analogy with the law on tangible property

2.1.1 "Data creator"

Some authors, especially Zech, propose that a right of data ownership should be inferred directly from the provisions on tangible property (Section 903 sentence 1 and Section 90 of the German Civil Code (BGB).[12] It is suggested that such a right of ownership is based on the creation of data by import or coding and should therefore be assigned to the data creator as the originally entitled party. The decision on who is the data creator is then based on economic factors, so in a service contract the client or employer would be classed as the data creator.

The creation of a (transferable) exclusivity right is mainly justified by suggesting that it would permit a clear allocation in the use of the data and a clear assignment of claims for compensation (e.g. damages or unjust enrichment). It is proposed that although different contractual provisions can be agreed, a right of ownership would at least provide a starting point for contractual provisions and a fundamental decision-making criterion in the absence of any provisions.[13]

This approach may be plausible for practical reasons, but it brings up the question of whether the assumption behind the analogy, i.e. the unintended gap in the legal provisions and the relevant interest in an analogy-based solution, actually exists. Undoubtedly, the authors of the German Civil Code (BGB) could not anticipate that the existence of digitised data would be a subject of legal discussion 100 years later. But now there are indeed numerous other special provisions for data, so the existence of an unintended gap in the provisions must be denied.[14] And the justified interests of the parties are also different, because tangible property rights require an exclusive assignment of property, but the Federal Constitutional Court underlines that the individual does not have any rights in the form of an absolute and unlimited dominion over "his" data, but rather has a personal identity which unfolds within the social community and depends on communication.[15]

2.1.2 Section 950 of the German Civil Code (BGB)

Ensthaler also advocates that provisions from the law of tangible property should be applied to data law. But he suggests that there is not a right of ownership as an all-embracing entitlement under Section 903 sentence 1 of the German Civil Code (BGB).[16] Instead, he proposes that the ownership question should be based on Section 950 of the Civil Code, i.e. that the person who processes information as "raw material" should acquire ownership unless the value of the processing is less than the value of the raw material. Similar to intellectual property rights, the processing party is thus not merely granted a contractual claim for compensation from the owner, he actually becomes the owner.

He suggests that this is reasonable because the raw data normally only become valuable when processed. This means that the company which provides the technical ability to collect and transmit the data on the relevant device would then be regarded as the processing party.[17]

2.2 Right of ownership by analogy with protection under criminal law

In a much discussed essay, Hoeren has attempted to justify a right of ownership in accordance with the value judgements in Sections 303a and 303b of the German Criminal Code (StGB), at least under future law (de lege ferenda).[18] He argues that the protection provided in Sections 303a and 303b of the Criminal Code is based on the possession of recording devices and processes. Thus he suggests that a right of ownership should be inferred from the "act of scripting" and should be the sole prerogative of the "scribe". In a nutshell: the person who records the data can keep them and exclude others from using them, or make such use dependent on his approval.

Here, the question is whether the protection provided under criminal law can really be claimed to be similar to the protection of a concept of ownership under civil law which focuses on proprietary use and establishing the creation of value. And Heymann — not without justification — sees the risk of a circular hermeneutic argument. In practice, the scripting-based approach also seems to create more questions than it solves. If we consider the complete virtualisation of data control and data recording that is achieved in many applications — including the associated reproduction by service providers and sub-contractors — every autochthonous recording would be an "act of scripting" which would lead to proprietary rights, and the resulting rights may then need to be re-transferred or reassigned to the "first scribe" or to the user of a cloud solution by means of contractual agreements, although the process involves the same datasets or datasets with fully identical content.

Boesche/Rataj[19] propose a solution which would allow the data ownership to be allocated without recourse to contractual provisions. This would happen in two stages. In a first step, the data must be distinguished according to their type and purpose. In case of data concerning the status of the terminal device, the data are likely to be allocated to the producer. But if the data deal with patterns of use, this would mainly be assigned to the third party (e.g. the service provider). This process is suggested as a way of determining where the main focus of the act of scripting lies and who is responsible for the main activities involved in the act of scripting.

2.3 Emoluments (Section 100 of the German Civil Code/BGB)

By contrast with the above positions, Heun/Assion consider on the one hand that individual datasets are not suited for an ownership concept. Rather, they propose that the pecuniary advantage which arises from the factual availability of the individual data should be allocated as a "proprietary" right by recognising it as usage of the data medium. In contrast to the act of scripting they do not focus on the act of creating or collecting data, instead they focus on the actual authority or title to the data medium that contains the individual data.[20] The assumption — following the basic idea of Section 100 of the German Civil Code — is that this then leads to an exclusive right to the "use" of the data contained on the data medium. In this connection, however, Heun/Assion emphasise that there can never be uniform answer to the question of who the data "belong to", but that the answers must always be found on a case by case basis.[21]

This approach is interesting, but at the same time it reminds us of the early period in software law when the rights to the software were also closely linked with the ownership rights to the data medium containing the software — and this even extended to the justification of the idea of software as a legally recognised object in general terms of business and the provisions of the law of obligations.

If we follow the logic of virtualisation and see how legal practice of the software law has departed from its insistence on a data medium as a justification for independent rights to software (naturally also as a result of the implementation of EU directive 2009/24/EC on the legal protection of computer programs), we should spare ourselves this "detour via the hardware" as a way to to justify possible rights to data. The progress of virtualisation and distributed computing processes (cloud & Co.) means that this "recourse to the hardware" would take us back to the 20th century, not forward into the 21st century.

Nevertheless, the idea that data should have separate legal protection as emoluments — which would not have the same legal force as an absolute right — is still an interesting approach which is worth further consideration.

2.4 The fruit of an object (Section 99 of the Civil Code/BGB)

In a similar view to Heun/Assion, Grosskopf[22] proposes that data are the fruit of the object which produces them and therefore belong to the owner of the object (Section 953 of the Civil Code). This would mean that the right to the fruit arises from ownership of the object which produces the fruit.

One argument against this approach suggests that fruits can only be physical objects and that data can therefore not be classified as fruits.[23] Other authors suggest that the data should not be seen as a product of the object which generates the data but as a product of the object or person which the data refers to.[24] This would mean that data do not necessarily belong to the owner of the object. Even if the data were deemed to be such fruits of the object which generates the data, this would not automatically make the data into an object or lead to a right in the data.

2.5 Right of ownership of the data subject

Other authors advocate ownership rights or at least quasi-ownership rights for the data subject, but detached from Sections 903 ff. of the German Civil Code. The existence of such a right is justified in various ways. Some authors base it on personality rights as a form of the right of informational self-determination, others derive the right of ownership implicitly from the comprehensive data protection rights and entitlements of the data subject.[25] In both cases, the data subject receives an absolute legal position in relation to all third parties, which is a typical feature of the right of ownership.

2.6 Legislative initiative for a "Data Act"

The Federal Ministry of Transport and Digital Infrastructure (BMVI) is possibly following a proprietary approach to the specific area of vehicle data. A strategy document presented by the Ministry in March 2017[26] points out that on the one hand data are not objects in the legal sense and therefore are not suitable for ownership.[27] However, in conclusion it suggests that data should be deemed equivalent to objects and thus be clearly allocable to natural persons or legal entities as "property".[28] In future, it proposes that the relevant rights of disposal should be allocated to the party "to which the creation of the data is attributable."[29]

The practical importance of this strategy document is especially seen in the area of mobility, in particular in vehicle data. A modern mass-produced car already produces up to 25 gigabytes of data per hour, for example regarding the weather, routes, traffic congestion and hazardous situations.[30] As a matter of principle this data should "belong to" the vehicle holder who has purchased the vehicle. Without the (revocable) consent of the affected party to the use of his/her personal data, the Ministry considers that any processing and networking of the data may only be done in an anonymous and pseudonymized form.

3. Open data approach

Within the community of authors fostering the open data approach, the notion of owning individual data is rejected and an approach that favours the free use of data is mainly taken.

Advocates of this opinion currently see no need to create such legal regulations, they consider that the instruments which are already available are both legally[31] and economically[32] sufficient. They especially believe that contractual agreements can provide enough protection. Moreover, they think that a data owernship right would create economic uncertainties. They consider that a generalised allocation of rights in data without striking a balance through rights of access and participation would hinder innovation, in particular given that big data applications depend on large volumes of data.[33] In its outcome, this view can draw on the reasoning given by the Federal Constitutional Court in its national census ruling of 1983, in which it has regarded information (including personal data) as an "image of social reality [...] which cannot be allocated exclusively to the affected person"[34]

In furtherance to that consideration, the main argument for an exclusive right in data, which appear to state that data processing and the development of a data market would lack incentives without such rights,[35] is not convincing provided that businesses maintain valid options to preserve their secrets.[36] In fact, exclusivity rights are superfluous in that context, where usage rights (such as in data) are non-rival.[37]

In order to ensure an effective data protection of data (whether under personality law or for economic reasons), data rights are recognised at different levels of intensity.

3.1 Protection of the information represented in data

Hoppen rejects an ownership right for data simply because the protection of such rights is not feasible on principle due to the lack of a physical embodiment of the data.[38] He suggests that the prime concern of the "owners" of data is not the protection of the data as such, but rather the protection of the information which is abstractly represented by the data, or rather the knowledge embodied in this information.[39] He states that data can be freely and accessibly transmitted and copied as long as the content is not identifiable, i.e. is encrypted. Hoppen suggests that legislation should focus on unencrypted data portfolios and on the protection of proprietary rights regarding information and knowledge.[40]

3.2 Protection by means of protection goals

The approach advocated by Heymann is similar in its outcome. He argues that an ownership right to protect individual data is explicitly not desirable, and that it would neither be a solution to any of the questions related to the allocation and control rights in data.[41] He proposes to consciously abstain from stipulating any ownership right in data. Instead, he argues that a concept of proper data processing based on protection goals should be aimed for, in order to safeguard objectives such as the confidentiality, integrity, "intervenability" and portability of data.[42] Taking the diversity of data into considerations he rejects a generalised solution.[43]

3.3 Protection by flexibility of private autonomy

Sahl also rejects a generalised, "static" solution to the allocation of data rights.[44] In its place, he recommends individual data usage contracts. He concedes that this approach has its weaknesses, especially in preventing third party infringements,[45] but he suggests that such a system would suit the "dynamic development" of digital markets and business models and the requirements of the individual case better because of its greater flexibility.[46] He considers that a general statutory provision is not necessary if some fundamental elements are maintained in every data use contract.[47] He queries as a fundamental issue and concern is in whose favor such a "one size fits all" solution would be determined.[48] In his view, a legislative solution would automatically benefit one of the parties, and that this is unlikely to do justice to the large number of different cases and interests which can exist in various constellations.[49]

Ernsthaler doubts that such a solution would be of any benefit. He believes that a contractual agreement does not answer the question of who the data are originally allocated to, i.e. who they belong to.[50] He believes that the exchange of data will only work by virtue of a person giving something in consideration for receiving something that did not belong to him before. He therefore considers that the question of the allocation of rights should be answered apart from the contractual options.

3.4 Transferable exclusivity right of the economically responsible data creator

Specht/Rohmer advocate a right of exclusivity in data that is based on the investment protection under Sections 87a et. seq. of the Copyright Act (UrhG).[51] They suggest that although individual data are not protected by those provisions, the allocation of rights in data is still based on the principle that the party which significantly invests in the procurement etc. of the affected data is the party which has or deserves a right of exclusivity.[52] In this connection, Specht/Rohmer want to distinguish between personal and non-personal data, but they admit that a distinction can only be made with difficulty in the individual cases.[53]

3.5 Extension of the concept of ownership

A similar approach is taken by Schwartmann/Hentsch, who also regard the Copyright Act as a model for new data usage rights.[54] To this end, their first goal is to categorise data so that a graded scale of protective concepts can be applied. They also suggest that the legislator should extend the concept of property in Article 14 (1) of the German Constitution to include "virtual" property.[55]

3.6 Right of erasure versus the right of ownership

As discussed above,[56] there is also a problem under data protection law that adds a striking argument against assuming ownership rights in individual datasets which would pose a significant challenge in aligning with suggested ownership rights in data. Wherever individual datasets also have the quality of personal data, or can be related to identifiable persons by other means, e.g. in combination with other data records, the data subjects can claim data deletion ("right of erasure") at any time, as well as object to any forms of processing that are not justified in law. These rights are anchored constitutionally in the data subjects' personality rights and the right of informational self-determination. How could an ownership right in individual datasets exist — even if it is anchored as a fundamental right under Article 14 of the German Constitution — if it is in constant pending conflict with a possible right of deletion that a data subject can action at any time against the "data owner", based on the data subject's fundamental right under Article 2 (1) of the German Constitution (right of informational self-determination) and the related right of erasure under data protection law. Given the "absolute" nature of exclusivity rights and property law, which normally persist in perpetuity, any such ownership right in data would be deprived of those key characteristics from the start - and limiting such ownership right to "non-personal data" does not appear viable given the fluidity of what might turn out to be personal data after all.[57]

III. Rights to data collections — rights of the database maker

In the view of the present author, therefore, it is not possible or desirable to derive ownership protection for individual data records from the provisions of civil law, so the key to the structure of data transactions should be found in the rights of the database authors. Therefore, the following section will look at the individual components of this right in more detail and point out where the law could be supplemented.

1. Requirements in Sections 87a et. seq. of the Copyright Act (UrhG)

1.1 Definitions

Section 87a (1) of the Copyright Act defines a database as a collection of works, data or other independent elements arranged in a systematic or methodical way and individually accessible by electronic or other means and whose obtaining, verification or presentation requires a substantial qualitative or quantitative investment. However, according to settled case law, economic expenses to procure the data (e.g. installation, development or operation of sensor equipment) are not recognised as investments under Section 87a (1) of the Copyright Act.[58] Only direct investments in the database are covered.[59] The maker of a database is deemed to be the party which has made this investment. This party does not need to have participated directly in the production of the database, the decisive question is who bears the economic risk which is associated with the creation and maintenance of a database.[60] If the activities of the database creator do not fall under the investment concept even though the database producer has invested much in the procurement of the data but relatively little in the systematic structure of the database as such,[61] protection of related rights under competition law may apply on the basis of the subsidiary relationship of unfair competition law to the right of the database maker (RDB) in this case.[62]

1.2 Scope of the protection

The legal definition underlines that the concept of the database can be understood very broadly. By contrast with the situation before the introduction of the database directive 96/9/EC, which was implemented in Sections 87a et. seq. of the German Copyright Act (UrhG), legal protection does not depend on a fixed (e.g. electronic) form,[63] nor does it require a certain number of data items or elements.[64] This means that protection is not only afforded to "creative" databases in which the selection or arrangement of the content in the database represents an independent intellectual creative act by the author,[65] rather there is a protective right for databases in their own right (sui generis database right) which protects significant investments in the procurement, review or presentation of the database content. In its core, therefore, the right of the database maker (RDBR) describes the eligibility of the investment in an organisational structure for the electronic selection of datasets for protection, but not the eligibility for protection of the individual data items as such.[66] The eligibility of this organisational structure for protection therefore focuses on the contextuality of datasets, but not necessarily on the content of the individual data item. Naturally, there are or can be points of contact, overlap-ping and similarities with the content of individual data items or the entirety of individual data items that can be parameterized. Consequently, the protection under Sections 87a et. seq. of the Copyright Act is limited to investments in existing data and their collection or classification.[67] The protection excludes investments which are made to generate the data as such, as a means to make up the content of a database. Data which are yet to be generated are thus not covered by these provisions.

At this point it is again clear that database rights do not protect the data which are collected and appropriately categorised.[68] Sections 87a et. seq. of the Copyright Act do not apply to the content — i.e. the data entered in the database — and do not constitute (new) information protection rights.[69] Under present legislation, a final answer to the central question about data sovereignty cannot be given on the basis of the criteria applying to database law. However, the consultation process announced by the EU Commission for the end of 2017 and running in 2018 to review the database directive offers the opportunity to discuss how the scope of the protection could be reasonably extended.

1.2.1 Application to meta data

Two topics are especially important when we define the concept of a database. On the one hand, it must be considered whether the sui generis protection for databases should be extended to meta data as such by way of an appropriate interpretation or extension of the applicable criteria. In the light of sub-section II.1.3, such an understanding could be a useful extension (especially where meta data also constitute or contain company and trade secrets).

1.2.2 Application to semi-structured and pre-structured data

On the other hand, it deserves being discussed whether the right of the database producer should apply to semi-structured or pre-structured data. Specifically, the question is at which point the data collection should be deemed as systematically or methodically arranged, and where, by contrast, it merely constitutes a “pile of data”. The distinguishing feature cited by the European Court of Justice is whether the collection includes any technical or other means (e.g. an index or organising structure) which enables each independent element contained in the collection to be localised.[70] In other words, a database is deemed to exist if each of its parts can be located by such means, whereas a “pile of data” does not have such a feature.

1.3 Contractual partners outside the EU

Where a contractual party makes an agreement for the allocation of data usage rights outside the EU, it must be taken into account whether the foreign contractual partner is at all entitled to claim database rights be database rights under Section 87a of the Copyright Act (UrhG),[71] respectively which difficulties may arise in enforcing any such rights.

1.4 European database rights as a role model beyond?

As much as the database rights, complimented by contractual solutions may appear incomplete, we do see an opportunity to further discuss and consider changes made at the European level to given a leading example and function as a “role model”.[72] The crucial step proposed to achieve this is that any new ancillary rights should be positioned a strict subsidiary manner in relation to the other “instruments of supplementary protection under competition law”, and, further, that the exemptions and limitations to such rights should be aligned with European copyright law.[73] This comes as an opportunity in light of the challenges of with database rights, as currently perceived worldwide .[74]

2. Company and trade secrets

Zech[75] points out that an exclusive right to data can be achieved not only by means of ownership, but also by factual exclusivity if the data are kept secret (protection of factual exclusivity). But if the secret is disclosed, this exclusivity ends, and by contrast with a genuine right of exclusivity there is then no (legally granted) exclusivity.

In addition to this important pillar to represent and implement the legal means to perform transactions with of data collections through suitable instruments (database licence agreements, transfers of rights to databases, etc.), the protection of trade secrets (Section 17 of the Act on Unfair Competition/UWG) and the - less clearly defined - general protection of know-how can be seen as the second pillar in the context of data collections and, where applicable, the protection of individual data items. This may result in quasi-proprietary rights both to data collections and, where applicable, to individual data items. However, in essence the statutory framework is actually designed and limited as rights to protect data against unauthorised access, interference and exploitation under civil and criminal law.

The European directive on the protection of confidential know-how and business information (trade secrets) which came into force in July 2016 will largely harmonise the different levels of protection in the member states and create minimum standards throughout the EU. For the German legal framework — which so far has been lacking an aligned approach for the protection of trade secrets in Germany (cf. for example Section 611 of the Civil Code, Sections 17 and 18 of the Act on Unfair Competition (UWG), Section 823 of the Civil Code) — this will bring changes in the systematic legal structure and especially in content. In future, for example, one criterion for assessing a trade secret is whether the relevant information is of commercial value and whether the owner of the secret has taken reasonable measures to protect its confidentiality (Article 2 No. 1 of the directive). Especially the first aspect is problematical in view of the fact that a "data value", as has been shown, can only arise from a certain volume of data. Here, we see that the directive is based on Article 39 (2) of the TRIPS Agreement, and has not been adapted to the circumstances of today. In terms of legal consequences, the directive comes closer to an industrial property law in that the owner of the secrets can now assert rights of recall and destruction (Article 12 of the directive). Similarly, the protection of secrets in court proceedings (Article 9) and the protection of whistleblowers are strengthened.

IV. Initiative "Creation of a European data economy"

In its position document of January 2017,[76] and most recently in a proposed regulation for free transactions with non-personal data of September 2017,[77] the European Commission has emphasised the high economic importance of data and data services and outlined further key points for the creation of a European data economy. In 2015 the value of the EU data economy reached 272 billion euros, which already came to 1.87% of the gross domestic product of the EU.[78] It is estimated that this will rise to 3.17% by 2020. On the other hand, only about 4% of all data are actually stored in EU countries.[79] The Commission is taking this development as a reason to expand the legal framework for a digital single market in order to exploit the full data potential within the EU in future.[80]

As a result of technical developments, especially the possibility of data connectivity, there are now new ways to access data. The traditional physical access to data is increasingly being replaced by remote access.[81] To harness these opportunities, the Commission wishes to remove unjustified hindrances to free data transactions and to remedy the legal uncertainty that prevails in many areas. Such "digital border controls"[82] especially consist of public authority regulations concerning the place of data storage and the processing of data and take the form of legal provisions, administrative regulations, or administrative procedures.

The proposal for a regulation on non-personal data supplements the European provisions for the protection of personal data, especially under the General Data Protection Regulation (GDPR). This underlines the Commission's goal of creating a more competition-based integrated single market for data processing services and activities.[83] It especially aims to achieve this by removing obstacles within the single market (i.e. geo-blocking) and by simplifying the transmission of data.

In its consideration of the creation of free data transactions in the position document of January 2017, the Commission also touched on the question of data ownership, but without presenting a final solution. Instead, it gave indications of how a right of the data creator to non-personal data could be structured.[84] It suggested that the owner or long-term user of a device (i.e. the party in possession of it) could have the right to use the data or permit others to use the data. It was suggested that this would give the data creator greater freedom to decide what could happen with the data generated by its machine, and that it would also help to avoid exclusive access to the data.[85] At the same time, it was suggested that exceptions should be created, for example for transport management or environmental reasons.

However, there are several reasons which lead us to reject a right of the data creator.[86] First of all, the existing provisions of civil and criminal law already offer sufficient instruments to protect the data.[87] Secondly, especially for machine-generated data there is the difficult question of who should be regarded as the data creator. According to the Commission, such data are "generated by machines without any direct human intervention in the course of computer processes, applications and services or by sensors which receive information from virtual or real devices or machines or from a software program."[88] Therefore, a number of different legal subjects could be regarded as the data creator, for example the manufacturer of the device or the software, their owner, their user, the party which invested in the development of the device or the party which operates and paid for the device.[89]

V. Conclusion

The discussion about the expediency of creating data ownership rights and the associated question of data sovereignty is in a state of flux. However, the analysis of the different arguments and approaches shows that individual data items themselves are generally not worth and adapt for an allocation of eligible for ownership rights. In addition to the difficulties in providing technical implementation that would safeguarding related systems and controls, this is especially due to overall economic reasons. A generalised assignment of exclusivity rights in data, without at the same time creating rights of access and participation to mitigate such legal position, would entail a high risk of creating obstacles for innovation and preventing the desired "free flow of data" even before it begins. On the other hand, the protection of individual data items and data collections by contractual agreements or by way of protecting databases in their own right (sui generis database rights) is likely to enable dynamic developments and allow for addressing a multitude of specific situations on an individual basis, leaving it to the parties concerned and participants in the market to address and regulate their commercial, information society and other potential interest in sharing and exchanging data in appropriate ways.

As a result and given the nature and potential value of data put in context, emphasis should rather be on considering the protection of data collections rather than individual data items. This can be seen, for example, in meta data which only have a value as a result of bringing together and correlating of different data items and types of data. The statutory framework lies in Sections 87a et. seq. of the German Copyright Act with its sui generis protection of databases and which is not dependent on a specific level of creativity. These provisions are supported by the protection of know-how (especially under Sections 17 and 18 of the Unfair Competition Act/UWG under current law), which will be enhanced by a more uniform standard of protection when the European Trade Secrets Directive is implemented into national law.

Data collection and creation of a European data economy

Data are often contained in texts, images or databases, and to generate data value chains they must first be extracted using text and data mining technology. But in the discussion about the creation of “data ownership” we have already seen that copyright and intellectual property rights can also create undesirable practical obstacles to access.

The following article deals with this problem from the perspective of a start-up and discusses current developments in relation to text and data mining (TDM) and the extension of the ancillary copyright for press publishers at the EU level. A simple TDM permission based on the fair use doctrine which is compatible with international copyright agreements is proposed as an appropriate option for a solution.

Text and data mining in the context of smart data — an economic perspective

Patrick Bunk, Ubermetrics Technologies GmbH

In the course of the smart data research projects, data value chains should be established to enable large, small and medium-sized companies to access key technology in the process of digitisation and thus participate in the budding of a European data economy. This perspective is not supported by the current position of the German Government on the European copyright law reform and the German regulations for text and data mining (TDM) from the perspective of copyright law.

At the European level it is currently being discussed whether permission for TDM should be created, but limited to research institutes and only for research purposes.[90] As a consequence of this proposal, it must be feared that the use of TDM by private entities or for commercial purposes will always require the approval of the copyright holder.[91] At the same time, the reform aims to extend protection to the products of press publishers without any requirements for the level of creativity of the data and without any exceptions even for extremely short text extracts.[92] This would mean that normal texts would also be protected as soon as they are included in a press publication, even if they only consist of single words, and thus they would be removed from public use for 20 years.

In Germany, the planned Act on Copyright in the Knowledge Society[93] also aims to allow TDM only for non-commercial research purposes (Section 60d of the Knowledge Society Copyright Act/UrhWissG-E). The use of TDM technology by business enterprises and start-ups, which is becoming increasingly important in practice, could therefore be hindered by the rather impractical need to obtain licences, and this would impair the competitiveness of such companies on international markets.

To understand the effects of this proposed reform on data value chains, it is helpful to start by considering the concept of “text and data mining”, which is a new idea in the legal context.

1. Text and data mining

Text and data mining (TDM) can be defined as a process by which high quality information or connections are extracted from texts or data. These processes have been a regular element of computer science for decades and can be found in many different fields of application.

A classic example of TDM is the search function in a Windows, Mac or Linux operating system. This function analyses all documents on the PC, automatically makes copies of all sentences in all documents and stores them in structured form in a database. As soon as the user searches for a document by entering any part of the document in a search mask, the information is retrieved in a fraction of a second, but it is retrieved from the database, not the original document.

There are many other examples of TDM:

- Spellchecking and grammar checking, and many other types of machine-based analysis of human language
- Pattern recognition processes which make it possible, for example, to click on phone numbers contained in e-mails on a smartphone
- Trend detection and analysis
- Spam detection
- Internet-search engines such as Bing, Google, Qwant or Cliqz

Most processes based on artificial intelligence (AI) are TDM technologies by definition because they “learn” high quality information in the form of rules from data or test input. Recent breakthroughs in the last 5 years have shown that AI technology is able to deal with monotonous information extraction and classification tasks. For example, there are applications which can sort texts by the language they are written in, recognise a bank account number on an invoice or identify animals on pictures.

2. How do today‘s artificial intelligence algorithms which are based on deep learning actually work?

AI technology creates a simplified version of a neuron, a human brain cell, and simulates it. In practice, hundreds of such artificial neurons are created and then connected to each other. This neural network receives certain input data, such as animal images or sentences, together with a classification which is then used to derive rules by using algorithms. To this end, the relevant algorithms must be shown differently classified objects for a long enough time, so that some of the simulated neurons learn some of the aspects of the problem. As soon as this point is reached, the algorithm can then perform tasks which could previously only be performed by a human.

But this requires a very large volume of data as a starting point so that the patterns can be learned. Normally this means hundreds of millions or even billions of texts and images. But small and medium-sized enterprises in particular do not have their own independently generated data collections. These only arise in very few major IT companies such as Google and Facebook. All other users in the fields of research and business therefore use the freely accessible public data collections on the Internet such as Wikipedia as a basis with which AI technologies can learn the structures in the relevant data.

3. The current legal position

It is currently in dispute whether copyright permission must be obtained for TDM, because as the government draft of the Knowledge Society Copyright Act states: “automated evaluation itself, the core of so-called text and data mining, is not an activity which is relevant to copyright”.

But this statement is not reflected in the practical use of TDM. The development, evaluation or improvement of TDM processes requires a constant data corpus which must be as large as possible in order to measure the quality of the algorithm. This pre-structured corpus always consists of a large number of validated documents or data streams which are representative of the specific problem. They are at least temporary copies, and they are naturally subject to copyright. In theory, a use of transient and incidental copies is possible if they are needed as an integral and essential part of a technical process, if their sole purpose is a lawful use and if these copies do not have any independent economic significance (cf. Section 44a of the Copyright Act/UrhG). But up to now, this exception can only apply to TDM processes in which the corpus is deleted immediately after the extraction of the information.[94] In practice, though, this would prevent any new development, evaluation or improvement of TDM-algorithms in Europe. Furthermore, in the field of TDM algorithms it is often necessary, from a user perspective or due to data protection requirements, to present the analysis results that have been achieved in a verifiable form by disclosing the original sources.[95]

This problem presents companies which use TDM technology with two major challenges:

1. Regarding millions of texts, they must determine whether they are protected by copyright. This requires an individual evaluation of each text to judge whether it has the necessary individual level of intellectual creativity, whereby this protection may also apply not only to creative works with their own distinctive character but even to works with a low creative value.[96]
2. To obtain licences for the use of such works, the companies must identify who are the authors or the holders of the exploitation rights of each work.

If we take into account the fact that the goal of using TDM is generally only to extract information — a process which is free if carried out by a human — it must be asked whether copyright protection is really appropriate in this area. The original intention of copyright protection is to protect the creative intellectual activity, not to protect the information itself.[97] According to the goal of the existing protection system, the information as such should be free for public use, otherwise this would hinder social interaction and progress.[98] As long as the exploitation of the original work is not affected by TDM, any exception to TDM should merely ensure that TDM is not used as a back door to other types of use which are relevant to copyright.[99]

The draft by the EU Commission states: "Text and data mining may also be carried out in relation to mere facts or data which are not protected by copyright, and in such cases no approval is necessary."[100] If this argument is inverted, as is confirmed by the interpretation presented by the German Federal Council, this means that in all future TDM processes in Europe where there could be a risk of using copyright-protected data, either the explicit approval of each author must be obtained, or the relevant TDM process should not be carried out at all.

4. Minimisation of liability by filtering out protected works as an alternative approach?

As a result of this proposed regulation, companies will ask themselves whether it is possible to avoid the risk of liability by first filtering out all works which could potentially be protected by copyright. Economically, this would be advisable because of the risks involved. However, even in an optimistic estimate, we would have to assume that 1% of one per cent of all documents in a corpus could contain parts which are copyright protected, and that by carrying out a TDM process with this corpus we would be committing a copyright infringement for the relevant documents. In a corpus consisting of 10 billion texts, this means that in a TDM process we would have to assume the risk of a copyright infringement in 1 million cases. In view of the customary compensation sums in the commercial sector, this would mean a liability risk of several million euros in each TDM process. With economic risks of this magnitude, TDM processes in Europe would not even be feasible in corporate research departments of large companies.

Therefore, the proposed EU regulation would require the creation of a complete and comprehensive copyright filter infrastructure. But this is not currently possible.

To determine whether a text contains a sufficiently large copy of another third party text which is protected by copyright, the filter would need to contain all texts by all copyright holders and then compare them with the relevant corpus text — and under the present interpretation of European law, even the creation of such a filter would itself be an infringement of copyright.

Optimists may object that computer science will advance to a point where such a solution can be found. But the absurdity of the proposed regulation is seen in the fact that even if the creation of such an algorithm were possible, the act of determining whether an article is copyrighted is by definition a TDM process in its own right. This means that every time this process identifies a part of a document that is worthy of copyright, the very act of checking the document would constitute a copyright infringement which could lead to a liability for compensation, simply by analysing the part of the text that is protected by copyright.

Therefore it must be concluded that filtering copyright material is not an option which could reduce the risk. As a practical consequence, this means that there would be inestimable disadvantages for the competitive development of TDM and AI technology in Europe, and that the proposed regulation is diametrically opposed to the postulated development of a European data economy.

5. Competitive consequences

A further consequence of this proposal would be a permanent change in the competitive structure of the AI technology sector.

As has been shown, it is impossible to prove that the training sets used for AI technology are free from copyright material. This means that licences are needed for very large volumes of data in order to develop AI technology. For companies without a dominant market position such as small and medium-sized enterprises and start-ups, it is a practical impossibility to negotiate contracts with every copyright holder in Europe. An exception for start-ups would not solve this problem because even highly successful start-ups hardly ever achieve a dominant market position. For all small and medium-sized enterprises and start-ups this means that the transaction costs for negotiations with every copyright holder in Europe are prohibitively high, and that the asymmetric market structure in negotiations means that any market solution is socially inefficient.

What happens with the large US corporations such as Google?

Google crawls the internet and trains its algorithms under the fair use doctrine[101] in the USA. The fair use doctrine permits the use of copyright material for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research (cf. 17 U.S.C. section 107), but the activities of commercial search engines can also fall under this provision.[102] To decide whether the use of a work constitutes fair use in individual cases, the following decisive criteria are used
(4 factor test):

- Purpose and character of the use, including whether such use is of a commercial nature or is for non-profit educational purposes,
- the nature of the copyrighted work,
- the amount and substantiality of the portion used in relation to the copyrighted work as a whole and
- the effect of the use upon the potential market for or value of the copyrighted work.

The jurisdiction of the Supreme Court place special importance on the question of whether the use of the work is transformative, i.e. whether it adds something new, pursues a new purpose or whether the work appears in a new context because of the use.[103]

The proposal by the EU Commission in itself has no effect on this circumstances due to the territorial scope of copyright law.

Google is a very large search engine with many users in Europe. Being found on the Internet is still very important for all companies and authors, and especially for publishers. For this reason, standard contracts with companies which are already powerful or dominant in the market are feasible.[104] Because this is linked to the search function, copyright holders will automatically be very willing to enter into contractual commitments. But companies with less market power, especially start-ups, will hardly ever be able to use such advantages. In future technology sectors such as AI, this could therefore cement monopolies and create market entrance barriers for European companies, even though the powerful or dominant companies in the market have not themselves engaged in any anti-competitive behaviour.

A comparable situation has already been seen in connection with the German ancillary copyright for press publishers. In this case, all German publishers granted Google a free licence to display text snippets. The competition authorities did not see this as an infringement of competition law.[105] At the same time, many publishers demanded licence fees from all smaller service providers.[106] Consequently, there is a risk that small providers are placed at a disadvantage, and that they will be squeezed out of the market in the long term.

Furthermore, publishers could shift the competition situation from their primary market to the secondary market of data refiners and information intermediaries. For example, if a search engine provider which is subsidised by publisher A does not receive any licences from publisher B, and in return the search engine subsidised by B has no access to the texts of publisher A, the results of any analysis would only reflect part of the reality, and their quality would suffer. If these providers are not able to establish themselves in the market, in the long term this means that the market could become concentrated on just a few dominant providers, which would reduce the supply of services in the market. But preserving the diversity of media and opinions in the search machine sector is an important social goal.[107]

Establishing licence models in which successful companies which use or develop TDM or AI technology pay part of their profit to copyright holders is only feasible if these companies also have a profitable business model in an international context. But to remain competitive, companies which use TDM or AI technology could move their registered place of business to countries with a legal situation which is comparable with the fair use doctrine. That would mean that this profit, and these jobs with high value creation potential for AI developers, would not arise in Europe.

As a result, the European economy will be dependent on AI systems from non-European providers. The proposal of the EU Commission therefore ignores the technical conditions for the development of AI and effectively consolidates the natural monopoly of the existing powerful non-European market participants.

6. How could this problem be solved?

Some draft statements by parliamentary committees envisage a broader TDM exception,[108] but they must first be accepted by the relevant committee and then prevail in the trialogue process between the Commission, the Council and the Parliament.

A simple TDM exception could be based on the model of the fair use doctrine, which is compatible with international copyright treaties, and could at the same time maintain the interests of the authors:

> "Uses which are necessary for TDM are to be permitted without the approval of the author, for both research institutes and private providers and for both non-profit and commercial purposes, on the following conditions:

- That access to the original sources is lawful or they are publicly accessible,
- That the commercial exploitation of the original source is not impaired by the analysis or dissemination of the analysis results, and especially that there is no substitute exploitation,
- That the work is only used for the extraction of information or for any other use permitted by the author."

By means of these restrictions, it could be ensured that the authors and originators are not disadvantaged in the exploitation of their work by the use of TDM or by the dissemination of the analysis results.

7. Conclusion

In the transition to a European data economy, public information should continue to be accessible to all so that everyone can find it. Apart from very limited exceptions, copyright law should still only protect the artistic expression, but not the facts and information which are processed in a work of art.

The proposed TDM regulation unnecessarily frustrates the development of a viable legal structure for a European data economy by introducing unavoidable risks of copyright liability. If a public document is legally accessed, there should not be any distinction based on whether the information contained in the document is processed by a human or a machine. Computers and algorithms get no enjoyment from the artistic value of a work that is protected by copyright. At least not yet. This proposed directive, which effectively merely attempts to subsidise the business models of European publishers by imposing new copyright fees for search engine technologies, will actually be a further massive boost to the dominance of a small number of giant U.S. technology corporations and will thus strengthen their negotiating position in relation to the publishers. As has been shown, it can already be anticipated that this will happen at the expense of the competitiveness and innovative power of European companies.

In view of this project, the economic benefit of the existing competence in key technologies such as data analysis and artificial intelligence systems in Europe can only be harnessed by a small number of American technology companies.

The need to finance journalism is understandable. The problem is economic. But the proposed regulation will achieve the opposite of what is intended because of its effect on the conditions for the development of the European data economy. In the medium term, classical publishing business models will actually be harmed, and only a handful of Internet companies with a strong market presence will benefit.

Data as an economic asset: Brief overview of the legal framework

The large number of legal disciplines and concepts which arise in a consideration of data as an economic asset are often confusing for lay persons, and even for legal experts, and this makes sensible detailed solutions difficult. The following section therefore aims to present a general overview of the different concepts and legal instruments. This condensed presentation is designed to help the reader to understand the concepts of the complex legal questions involved.

The following questions are especially relevant in connection with data as an economic asset: When does copyright protection apply? What does the concept of sui generis database rights mean? How can companies protect their trade secrets? A new element is the ancillary copyright for press publishers. And if third parties gain unauthorised access to data, the provisions of criminal law may also be relevant.

Copyright Law

Requirements for copyright protection
The protected works of literature, science and art include for example speech and text works, computer programs, music, pantomime, dance, works of art, photographs, films, scientific or technical presentations such as drawings, plans, maps, sketches, tables and sculptures (cf. Section 2 of the Copyright Act/UrhG). Official works such as laws or official announcements are not protected by copyright law (Section 5 of the Copyright Act).

Only personal intellectual creations qualify as works under the Copyright Act. Even very short text passages may qualify as works.[109] However, simple descriptions or reproductions of pure information in everyday language are not sufficient, so especially when user generated content is involved it is necessary to assess the creative quality of each individual case.[110]

Exclusive rights
In Principle, the author has the exclusive right to exploit the work, i.e. to use it, permit others to use it or exclude others from the use of the work. These uses especially include the right of editing, reproduction, dissemination, exhibition and public performance, and this in turn includes the right to make the work publicly accessible (cf. Sections 15 ff. of the Copyright Act). Special provisions for computer programs can be found in Sections 69a ff. of the Copyright Act.

The author may grant third parties the rights of use for individual types of use or all types of use. A distinction is made between simple and exclusive rights of use which can be granted limited to location, time or content (cf. Sections 31 ff. of the Copyright Act).

Digital copies generally fall under the right of reproduction, and this in principle applies to every form by which a copyrighted work is transmitted to another storage medium, irrespective of whether the copies are privately or publicly made or whether they are transient, permanent or made in another format.[111] Another relevant issue in an online context is the right to make the work publicly accessible. This may also apply to works which have already been published online if a work is reproduced by a different technical process which differs from the previously used process, or if it is reproduced for a new audience. The audience is deemed to be new if the author did not direct the original public reproduction to this audience, for example in cases of limited retrievability or access control.[112]

Exceptions and Limitations to copyright
The copyright is limited by exceptions which define the conditions under which the works can be used without permission. In the context of smart data, for example, temporary acts of reproduction are relevant under Section 44a of the Copyright Act, which states that transient copies which are technically indispensable for lawful use still can be permissible. In addition, other exceptions for science and teaching are planned in the Knowledge Society Copyright Act[113] and in the draft of a text and data mining exception in the draft directive on copyright in the digital single market.[114]

Open source
Works under an open source licence can generally be used free of charge, but in some cases they are also subject to restrictions in their licence provisions. One example of this are “Copyleft” clauses which stipulate that all further developments based on the work must be freely accessible under the same licence conditions.

Consequences of copyright infringements
If there are infringements of copyright, authors are entitled to injunctive relief and damages and to insist that all reproduced copies must be destroyed (Sections 97 and 98 of the Copyright Act). In addition, any deliberate unauthorised exploitation of copyright works is liable to prosecution (Section 106 of the Copyright Act).

Database works

(Section 4 of the Copyright Act)

If the necessary level of creativity is fulfilled in the arrangement and design of a database, the database can also constitute a copyright-protected work under Section 4 of the Copyright Act. The level of creativity of a work will generally not be deemed to apply if the arrangement or presentation automatically arises from the nature of the subject or is prescribed by the laws of expediency, by logic or by necessities and if there is not sufficient scope for an independent intellectual design of the form.[115] Electronic databases need to have an output format which enables the data to be accessible in a systematically and methodically ordered manner.[116] The decisive factor for recognition as a work is the originality of the links and the query options.[117]

Sui generis database rights **(Section 87a of the Copyright Act)**

Database manufacturers can assert a right to protection under Section 87a of the Copyright Act if the data are systematically or methodically arranged and individually accessible by electronic means or in any other way, and if their procurement, review or presentation requires an investment of a significant manner or extent.[119] There is no protection for mere “piles of data” due to the lack of a systematic or methodical arrangement of the individual elements, i.e. for raw data that have not yet been especially structured, even if the procurement of the raw data required a significant investment.[120] But an unstructured internal data storage system may be protected if the query system creates a systematic or methodical order.[121] The decisive factor is the connection between the volume of data and a query system which permits targeted searches for individual elements in the data.[122] The Higher Regional Court (OLG) of Cologne[123] ruled that this does not depend on whether the individual information items recorded in the database have been processed.[124]

Another requirement is the significant investment. This may be financial, or it could consist of an investment of time, work and energy.[125] It could also apply, for example, to investments in processing the data, designing the links and developing the query options, but not the means used to generate the data themselves.[126] This means that resources are covered which are used to identify and compile data which previously exist, but not the resources used to generate the data elements themselves.[127] However, checks for correctness and reliability have been recognised as allowable investments in case law rulings,[128] so it can be assumed that the database-related investment costs for data mining processes which analyse and identify hidden connections in existing data will also be recognised.[129] Therefore it must be checked whether investments or work should be directed towards the necessary structuring and processing of the (existing) data or towards generating the raw data or “new” data.

Sections 87a ff. of the Copyright Act do not protect the information contained in the database.[130] The sui generis protection of the database author is not designed to create a new right to the individual elements collected in the database as such.[131] “The protection is afforded not to the individual information items entered into the database, but to the database as the overall total of the content which has been collected, organised and made individually accessible by means of a significant investment, as an intangible asset.” [132] The database creator can only prohibit the reproduction, dissemination and public presentation of the database as a whole, or of parts which are significant in their type/extent. According to Section 87b (1) sentence 2 of the Copyright Act, the same protection is granted to the repeated and systematic use of minor parts as long as these activities run counter to a normal evaluation of the database or unreasonably impair the justified interests of the database creator.

Ancillary copyright for press publishers

The introduction of this ancillary right grants the publisher of a press publication the exclusive right to make the press publication or parts thereof publicly available for commercial purposes for one year from the date of publication (Sections 87f and 87g (2) of the Copyright Act). Press publications include:

> Editorial fixation of a collection of journalistic contributions, within a periodical publication under a single title on any medium, which - regarding the overall circumstances - is characteristic for publishing purposes and not mainly used for self-advertising purposes. Journalistic contributions especially consist of articles and illustrations which serve for information, opinion forming or entertainment.

Even blogs can constitute press publications if they can be deemed to be an editorially selected collection of journalistic contributions.[133]

The protection does not apply in relation to all possible readers, only in relation to commercial providers of search engines and commercial providers of services which process content accordingly (Section 87g (4) of the Copyright Act). The press publisher only has the exclusive right to make the original publicly available for commercial purposes, so there is explicitly no restriction for reproductions.[134] However, as soon as press publications reach the relevant level of creativity, they can enjoy copyright protection.

But linking remains possible[135] because the new protection right does not extend to individual words and extremely short text extracts.[136] There is controversy about the permissible length of these “short text extracts”, especially in connection with the display of “snippets”. According to a recent case law ruling by the Higher Regional Court (OLG) of Munich, text extracts with a length of at least 25 words cannot be regarded as extremely short text extracts under Section 87f (1) sentence 1 of the Copyright Act.[137]

The ancillary copyright proposed by the EU Commission in its draft for a directive on copyright in the digital single market[138] would go even further because the draft does not contain any limitation in the specific obliged parties or text lengths, and the right of reproduction “for digital use” is to be reserved for the press publishers. In addition, the duration of the protection would be extended from 1 year to 20 years if the draft directive comes into force.

Trade and Business secrets

By analysing sensor data, for example from machines, conclusions can be drawn about the production of machines and products or the use of these machines in the business company. This means that the trade secrets of both the machine manufacturer and the machine operator can be inferred, and that company-related knowledge can be obtained.

Up to now, the provisions for the protection of company and trade secrets in Sections 17 and 18 of the Act on Unfair Competition (UWG) entitled to injunctive relief and damages if facts, circumstances or transactions related to a company, which are only known to a limited group of persons and not publicly known, are revealed by an unauthorised act although the holder of the protection rights has a justified interest in the preservation of the secrecy of this information.[139]

An upcoming revision according to of Article 2 No. 1 (a)-(c) of the Directive (EU) 2016/943 will define trade secrets as follows:

- Secret (not generally known or not readily accessible),
- Of commercial value because it is secret, and
- A subject to reasonable steps to keep it secret regarding the individual circumstances

Under the future legal position, if affected parties wish to protect data as trade secrets it will not simply be assumed that they have an interest in maintaining secrecy, instead they must actively take "appropriate" steps to maintain secrecy. This could lead to a necessity to establish technical and organisational measures similar to that in data protection law. If access to the data is not made so difficult by technical means or organisational steps that data access would require a disproportionate amount of effort, there will probably be no legal protection. To prevent competitors from obtaining access to company-specific data, one option would be to use the known mechanisms under data protection law such as data separation, access control and anonymisation (removal of any reference to the company).

Punishment for data interception and espionage

Consequences under criminal law are available for persons who carry out the following unauthorised actions:

- Gain access to data which is not designated for them and is specifically protected from unauthorised access, by overcoming the access protection feature, or
- Use technical means to obtain data not designated for them from a non-public data transmission or from the electromagnetic radiation from data processing equipment (Sections 202a and 202b of the Criminal Code/StGB).

But the criminal law provisions only apply to data which are stored or transmitted electronically, magnetically or in any other not directly visible manner.

Entitled access to the data

The person or organisation which stores the data is normally entitled to access the data, and it is not decisive whether this person is also the owner of the data medium. The authorisation may be transferred, e.g. by assignment for use, if the authorisation to use the program data is assigned at the same time.[140]

Overcoming the access protection

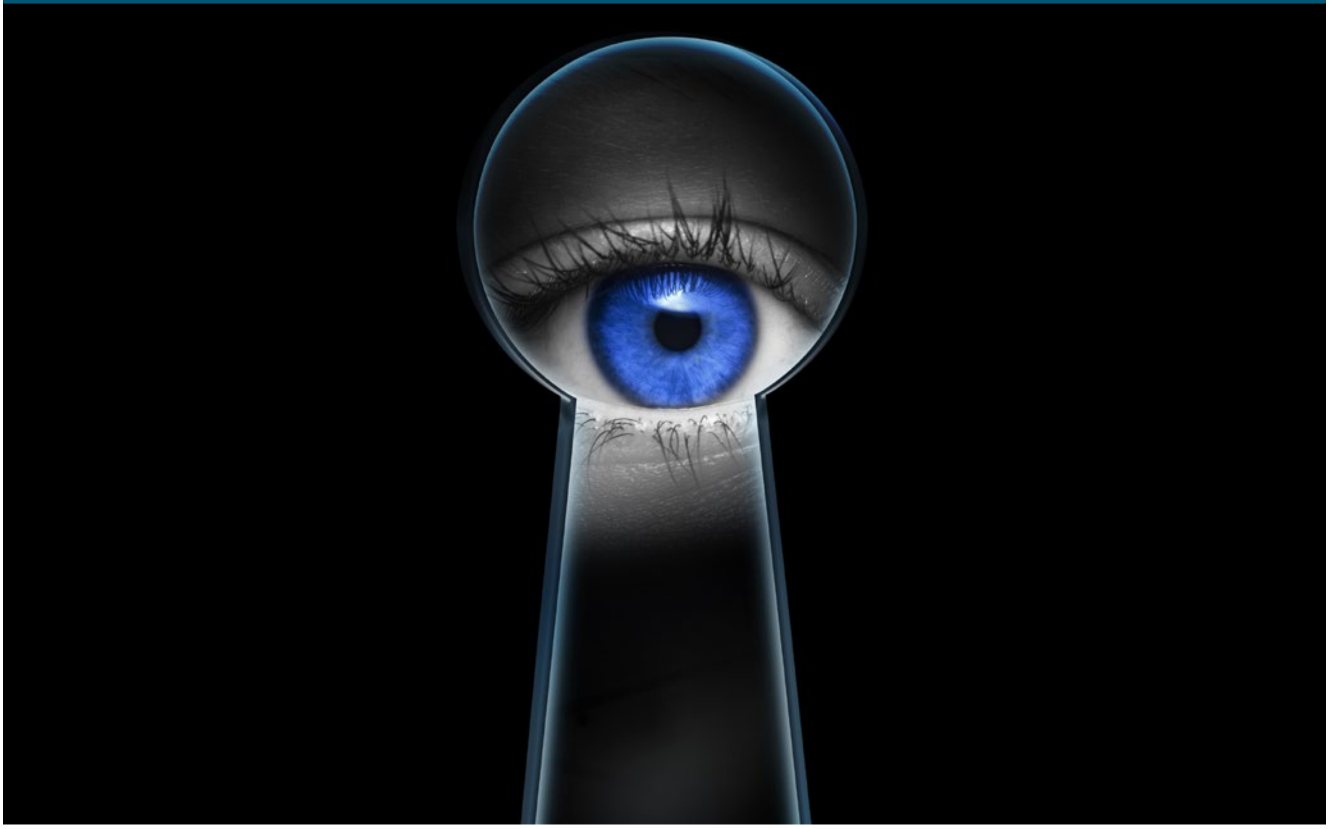
The special protection must have the purpose of preventing access.[141] It is also demanded that overcoming the access protection must be not readily possible, and must require a significant amount of time or technical effort.[142]

Unauthorised

So-called penetration tests which serve to identify security flaws in the IT system are not liable to prosecution if they are permitted by the person entitled to the data.

Wilful intent

Perpetrators must at least willingly take the risk of unauthorised access to access-protected third party data.



About the authors

Authors

DR. ALEXANDER DUISBERG

... is a partner of Bird & Bird in Munich, handling commercial, digital transformation and data protection issues. He is a recognized leader and government advisor to the German government think tank projects “Trusted Cloud”, “Smart Data” and “Plattform Industrie 4.0”.

PATRICK BUNK

... the founder and CEO of ubermetrics Technologies GmbH. He studied Economics in Berlin and at the Northwestern University in the USA:. He is an expert in quantitative methods to analyse the dissemination of information and its effects on the market.

Publishing editors

PD DR. OLIVER RAABE

... leads the legal expert group within the Smart Data project. He is founder of the research group “Information law for technical systems and legal informatics - ITR)“ at KIT (Karlsruhe Institute of Technology) and director at FZI Forschungszentrum Informatik.

MANUELA WAGNER

... is research associate at the Center for Applied Legal Studies at KIT (Karlsruhe Institute of Technology) and member of the research group ITR. She supports re-search projects in the field of data protection, IT-security and energy law.

Members of the legal framework specialist group

Amon, Peter
Siemens AG, Virtuose-DE project

Bremert, Benjamin
Independent Center for Privacy Protection Schleswig-Holstein (ULD), iTesa project

Bretfeld, Jürgen
Advaneo GmbH

Bretthauer, Dr. Sebastian
Johann Wolfgang Goethe University of Frankfurt am Main, Smart Regio project

Bunk, Patrick
Ubermetrics GmbH, Smart Data Web project

Drepper, Dr. Johannes
TMF e. V., SAH RA project

Duisberg, Dr. Alexander
Bird & Bird

Eckhardt, Dr. Jens
Derra, Meyer und Partner Rechtsanwälte PartGmbB

Elteste, Thomas
DB-Systel, SD4M project

Fasching, Peter
UK Erlangen, KDI project

Freitag, Gerald
DB-Systel, SD4M project

Friederici, Florian
Fraunhofer FOKUS, Virtuose-DE project

Fröhlich, Sven
Technical University of Dresden, ExCELL project

Gläß, Valérie LL.M.
TMF e. V., SAH RA project

Gül, Serhan
Fraunhofer HHI, Virtuose-DE project

Guzman, Liliana
Fraunhofer IESE, PRO-OPT project

Hilber, Dr. Marc LL.M.
Oppenhoff & Partner

Janneck, Kai
Independent Center for Privacy Protection Schleswig-Holstein (ULD), iTesa project

Jeske, Henning
Technical University of Dresden, ExCELL project

Klein, Achim
University of Hohenheim, InnOplan project

Lenk, Dr. Alexander
BMW Group

Maier, Florian
Fraunhofer IAO, Smart Energy Hub project

Meiers, Thomas
Fraunhofer HHI, sd-Kama project

Oppermann, Henrik
USU Software AG, SAKE project

Premm, Marc
University of Hohenheim, InnOplan project

Runde, Dr. Detlef
Fraunhofer HHI, sd-Kama project

Schallaböck, Jan
iRights.Law Rechtsanwälte, Smart Data Web project

Schmidt, Martin
Cautus Service GmbH

Spiecker genannt Döhmann, Prof. Dr. Indra LL.M.
Johann Wolfgang Goethe University of Frankfurt am Main, Smart Regio project

Stecher, Björn
Initiative D 21

Steinmann, Jonas
TMF e. V., SAHRA project

Steffen, Dr. Matthias
Bayer AG, Sidap project

Troemel, Marc
Vico Research & Consulting GmbH, Smart Data Web project

Ursinus, Sven
BITMi Bundesverband IT-Mittelstand e. V.

von Grafenstein, Maximilian LL.M.
Alexander von Humboldt Institute for Internet and Society

Wachovius, Juliane
Hof University of Applied Sciences, Institute of Information Systems at Hof University (iisys)

Wacker, Richard
YellowMap AG

Weber, Prof. Dr. Beatrix MLE
Hof University of Applied Sciences, sd-Kama project

Weichert, Dr. Thilo
Netzwerk Datenschutzexpertise

Willkomm, Dr. Marlene
Deputy head of the Cologne Flood Control Centre, Cologne municipal sewage enterprise, sd-Kama project

Wimmer, Max
Hof University of Applied Sciences, sd-Kama project

Xu, PD Dr. habil. Feiyu
DFKI German Research Center for Artificial Intelligence, SD4M project

Zwingelberg, Harald
Independent Center for Privacy Protection Schleswig-Holstein (ULD), iTesa project iTesa project

Footnotes

[*] We especially wish to thank Dr. Benedict Vogel, LLM, Bird & Bird LLP, for valuable support in the preparation of this report.

[1] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions of 17 Jan. 2017, “Building a European data economy”, COM(2017) 9 final.

[2] “Open data in Germany”, seven demands by the working group “Economic potential and social acceptance” of the Smart Data Research Department, available under: http://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/smart-data-brosch%C3%BCre_open_data_deutschland.pdf?__blob=publicationFile&v=9.

[3] European Commission, proposal for a directive of the European Parliament and the Council on copyright in the digital single market of 14 Sept. 2016, COM(2016) 593 final.

[4] Wiebe, CR 2017, 87 (91) also speaks of the “principle of the common use of information” in this connection.

[5] Cf. Recital 3 of the PSI Directive.

[6] On interoperability, cf. especially Kerber/Schweitzer, Interoperability in the Digital Economy, in: Journal of Intellectual Property, Information Technology and Electronic Commerce Law (JIPITEC) 8(1), 2017, 39.

[7] The term meta data denotes data which contain information about other data (cf. Harrach, Risiko-Assessments für Datenqualität (Risk assessments for data quality), P. 21) and enable the user to understand and transmit data collections (cf. ISO/IEC specification 11179 [IS099]).

[8] Zech, Information als Schutzgegenstand (Information as a protected object, 2012), P. 37 categorises information as semantic (meaning) and syntactical (represented by symbols).

[9] Naturally, here contextuality arises and is added by the meta data included in or added to the measurement (e.g. place and time of the measurement, measurement parameters etc.).

[10] Heymann, CR 2016, 650 (657).

[11] One of many: European Commission, Legal study on Ownership and Access to Data — Final report 2016, P. 61 f., available under https://book-shop.europa.eu/en/legal-study-on-ownership-and-access-to-data-pbKK0416811/ incl. further documentation

[12] Zech, CR 2015, 137 (144); cf. also Dorner, CR 2014, 617 (618) incl. further documentation

[13] Cf. Heun/Assion, CR 2015, 812 (818). A critical view of contract drafting options: Ernsthaler, NJW 2016, 3473 (3474); see also Chapter 3.3 below

[14] Cf. Peschel/Rockstroh, MMR 2014, 571 (572).

[15] Consistent case law since Federal Constitutional Court ruling of 15 Dec. 1983, BVerfGE 65, 1 (43f.).

[16] Ensthaler, NJW 2016, 3473 (3475).

[17] Ensthaler, NJW 2016, 3473 (3475).

[18] Hoeren, MMR 2013, 486.

[19] Boesche/Rataj Zivil- und datenschutzrechtliche Zuordnung von Daten vernetzter Elektrokraftfahrzeuge (Classification of data from networked electric vehicles under civil and data protection law), P. 42, available under http://schau-fenster-elektromobilitaet.org/media/media/documents/dokumente_der_begleit_und_wirkungsforschung/EP21_Zivil-_und_datenschutzrechtliche_Zuordnung.pdf

[20] Heun/Assion, CR 2015, 812 (818).

[21] Heun/Assion, CR 2015, 812 (814).

[22] Grosskopf, IPRB 2011, 259.

[23] Grosskopf, IPRB 2011, 259.

[24] Specht, CR 2016, 288 (292).

[25] Weichert, NJW 2001, 1463 (1476) with reference to Ladeur, DuD 2000, 12 (18); Kilian, Gedächtnisschrift for Wilhelm Steinmüller (Memorial Volume for Wilhelm Steinmüller), 195 (207 ff.).

[26] Federal Ministry of Transport and Digital Infrastructure, Wir brauchen ein Datengesetz in Deutschland (We need a Data Act in Germany), 20 Mar. 2017, available under http://www.bmvi.de/SharedDocs/DE/Artikel/DG/daten-gesetz.html.

[27] Cf. also the documentation in chapters 2.1.2 and 2.3.

[28] Federal Ministry of Transport and Digital Infrastructure, op cit

[29] Federal Ministry of Transport and Digital Infrastructure, op cit

[30] Cf. Federal Ministry of Transport and Digital Infrastructure, op cit

[31] E.g. Specht, CR 2016, 288 (296); Drexl/Hilty/ Desaunettes/Greiner/Kim/Richter/Surblyt6/Wie-demann, GRUR Int. 2016, 914 f.; Grützmacher, CR 2016, 485 (495); Schefzig, K&R 2015, No. 9, supplement, 3, 3 (6). Cf. also Dorner, CR 2014, 617 (626), who especially points out constantly rising investments in big data applications which do not indicate any deficiency in the protection.

[32] For a detailed consideration of non-personal data cf. Kerber, A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis, in: Gewerblicher Rechtsschutz und Urheberrecht (Intellectual Property Law and Copyright). International part (GRUR INT), 2016, 989, chap. VIII.

[33] Dorner, CR 2014, 617 /626).

[34] Federal Constitutional Court, ruling of 15 Dec. 1983, BVerfGE 65, 1 (44). Recently, however, the ruling has sometimes been interpreted to mean that the individual “cannot be granted unrestricted dominion over the data which apply to him, but that this does not countermand a less far-reaching entitlement which is limited by appropriate barrier provisions”; e.g. Specht, CR 2016, 288 (293); cf. also Specht/Rohmer, PinG 2016, 127.

[35] See above, footnote 11.

[36] Kerber, op cit (footnote 32), chap. IV.

[37] In summary: Kerber, op cit (footnote 32), chap. VIII.

[38] Hoppen, CR 2015, 802.

[39] For a similar approach, Heymann, CR 2015, 807 (808).

[40] Hoppen, CR 2015, 802 (806); also Heymann, CR 2015, 807 (811).

[41] Heymann, CR 2015, 807 (809).

[42] Heymann, CR 2015, 807 (810f.).

[43] Heymann, CR 2015, 807 (810).

[44] Sahl, PinG 04.16, 146 (150).

[45] Sahl, PinG 04.16, 146 (149); also Becker, GRUR Newsletter 01/2016, 7.

[46] Sahl, PinG 04.16, 146 (150).

[47] Assion/Mackert, PinG 04.16, 161 (161); Sahl, PinG 04.16, 146 (150).

[48] Sahl, PinG 04.16, 146 (149).

[49] Cf. also Becker, GRUR Newsletter 01/2016, 7.

[50] Ensthaler, NJW 2016, 3473 (3474).

[51] Specht/Rohmer, PinG 04.16, 127 (131).

[52] Specht/Rohmer, PinG 04.16, 127 (131).

[53] Specht/Rohmer, PinG 04.16, 127 (129).

[54] Schwartmann/Hentsch, PinG 04.16, 117 (120).

[55] Schwartmann/Hentsch, PinG 04.16, 117 (120f.).

[56] See above, II. 1.6.

[57] On the tension between “data ownership”and data protection cf. Härting, CR 2016, 646

[58] European Court of Justice, MMR 2005, 29; Federal Court of Justice (BGH), MMR 2005, 754.

[59] For a general consideration see Zdanowiecki, in Bräutigam/Kindt (eds.), Digitalisierte Wirtschaft/Industrie 4.0 (Digitised economy/Industrie 4.0, 2015), P. 19 ., available under: http://www.bdi.eu/Gutachten_Digitalisierte-Wirtschaft_Industrie-40.pdf.

[60] Recital 41 of the Directive 96/9/EC.

[61] Cf. Sahl, PinG 04.16, 146 (148).

[62] Cf. Leistner, Der Rechtsschutz von Datenbanken im deutschen und europäischen Recht (Legal protection of databases in German and European law, 2010), P. 343

[63] Cf. Article 1 (1) of the Database Directive 96/9/EC, implemented by Sections 87a ff. of the Copyright Act; which covers databases “in any form”.

[64] Cf. the proposal of the European Parliament in its statement of 23 Jun. 1993 on the proposal for the Database Directive, ABIEG No. C 194, P. 144.

[65] Cf. Götz, ZD 2014, 563 (564) incl. further documentation; . Leistner, Der Rechtsschutz von Datenbanken im deutschen und europäischen Recht (Legal protection of databases in German and European law, 2010), P. 343 ff.

[66] Dorner, CR 2014, 617 (622); Zech, CR 2015, 137 (143).

[67] European Court of Justice, ruling of 9 Nov. 2004, ref. C-203/02, EuZW 2004, 757 (759 & 760).

[68] Cf. also Specht, CR 2016, 288 (293 f.); Drexl/Hilty/Desaunettes/Greiner/Kim/Richter/Surblyté/Wiedemann, GRUR Int. 2016, 914 (915). (648 f.).

[69] Ensthaler, NJW 2016, 3473 (3474).

[70] European Court of Justice, GRUR 2005, 254 (255), marginal note. 30f.

[71] Assion/Mackert, PinG 04.16, 161 (161f.).

[72] Leistner, in Handwörterbuch des Europäischen Privatrechts (Dictionary of European private law) Vol 1, 2009, 298 (301).

[73] Leistner, in Handwörterbuch des Europäischen Privatrechts (Dictionary of European private law) Vol 1, 2009, 298 (301).

[74] Leistner, in Handwörterbuch des Europäischen Privatrechts (Dictionary of European private law) Vol 1, 2009, 298 (301).

[75] Zech, CR 2015, 137 (140).

[76] European Commission, Communication v Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions of 10 Jan. 2017, “Creation of a European data economy”, COM(2017) 9 final.

[77] European Commission, proposal for a regulation of the European Parliament and the Council on a framework for the free movement of non-personal data in the European Union of 13 Sept. 2017, COM(2017) 495 final.

[78] European Commission, op cit (footnote 77), P. 2.

[79] Cf. Wiebe, CR 2017, 87.

[80] European Commission, op cit (footnote 77), P. 8.

[81] European Commission, op cit (footnote 77), P. 4.

[82] European Commission, op cit (footnote 77), P. 5.

[83] European Commission, op cit (footnote 78), P. 2.

[84] In detail, cf. also Wiebe, CR 2017, 87 ff..

[85] European Commission, op cit (footnote 77), P. 14.

[86] See above, chap. 2.1.1.

[87] Van Asbroeck/Debussche/César, Data Ownership: A new EU right in data, available under https://sites-twobirds.vuture.net/52/1373/uploads/sup-plementary-paper-on-data-ownership.pdf.

[88] European Commission, op cit (footnote 71), P. 10.

[89] Wiebe, CR 2017, 87 (90). On the same problem cf. also Specht, CR 2016, 288 (295).

[90] Article 3 of the proposal by the EU Commission: Proposal for a Directive of the Parliament and the Council on Copyright in the Digital Single Market, COM(2016) 593 final, of 14.9.2016.

[91] Bundesrats-Drucksache (Federal Council documents) 535/16, P. 7.

[92] Article 11 COM(2016) 593 final.

[93] Draft law to align copyright law with the current requirements of the knowledge society (Knowledge Society Copyright Act — UrhWissG-E)

[94] On the individual requirements, cf.: Triail-le/de Meeüs d'Argenteuil/de Francquen, Study on the legal framework of text and data mining (TDM), funded by the European Commission, March 2014.

[95] Federal Council printed document 535/16, P. 7.

[96] Federal Parliament printed document 1V/270, 38.

[97] Raue GRUR 2017, 11 (13).

[98] Raue GRUR 2017, 11 (13).

[99] Schack ZUM 2016, 266 (269).

[100] Draft law of the Federal Government align copyright law with the current requirements of the knowledge society, P. 44.

[101] Cf. for example the case Authors Guild versus Google, Inc. Court of Appeals for the Second Circuit New York, ruling of 16 October 2015 — 13-4829-cv.

[102] KELLY versus ARRIBA SOFT CORPORATION, United States Court of Appeals, Ninth Circuit, ruling of 6 February 2002 — No. 00-55521.

[103] Campbell versus Acuff-Rose Music, Inc., Supreme Court of the United States, 510 U.S. 569, 579, 114 S.Ct. 1164, 127 L.Ed.2d 500, ruling of 7 March 1994.

[104] The request for approval of the display of text excerpts free of charge by a press publisher which operates a website was not considered by the Federal Cartel Office (Bundeskartellamt) to be an abuse of the dominant position of a search engine operator, Federal Cartel Office Bonn, ruling of 8 Sept. 2015 — B 6 126/14.

[105] Federal Cartel Office Bonn, ruling of 8. Sept. 2015 — B 6 - 126/14.

[106] Higher Regional Court (OLG) Munich, ruling of 14 July 2016 — 29 U 953/16.

[107] Paal ZRP 2015, 34.

[108] Committee on Industry, Research and Energy, draft opinion, of 2 Mar. 2017, available under: http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2f%2fEP%2f%2fNONS-GML%2b-COMPARL%2bPE-592.363%2b01%2b-DOC%2b-PDF%2bV0%2f%2fEN; Committee on the Internal Market and Consumer Protection, draft opinion of 20 Feb. 2017, available under http:// www.europarl.europa.eu/sides/getDoc.do?pub-Ref=-%2f%2fEP%2f%2fNONSGML%2b-COMPAR-L%2bPE-599.682%2b01%2bDOC%2b-PDF%2b-VO%2f%2fEN

[109] Cf. European Court of Justice, ruling of 16 July 2009 — C-5/08; Higher Regional Court (OLG) Munich, ruling of 14 July 2016 — 29 U 953/16 marginal note. 63 f.

[110] Wandtke/Bullinger/Bullinger UrhG (Commentary on the Copyright Act) Section 2 marginal notes 156, 159.

[111] Cf. Dreier/Schulze/Schulze (UrhG) Section 16, marginal note 4; Wandtke/Bullinger/Heerma (UrhG) Section 16 marginal note 5, 16-17.

[112] Higher Regional Court (OLG) Munich, ruling of 14 July 2016 — 29 U 953/16. 953/16 —, marginal note 72, with reference to European Court of Justice, decision of 21 October 2014 — C-348/13 —, Federal Court of Justice (BGH), ruling of 9 July 2015 — I ZR 46/12.

[113] Draft law to align copyright law with the current requirements of the knowledge society (Knowledge Society Copyright Act — UrhWissG-E)

[114] European Commission, proposal for a directive of the European Parliament and the Council on copyright in the digital single market of 14 Sept. 2016, COM(2016) 593 final.

[115] On distinguishing works with a lower level of creativity, cf. for example Bisges GRUR 2015, 540.

[116] Dreier/Schulze/Schulze (UrhG) Section 4, marginal note 17

[117] Dreier/Schulze/Schulze (UrhG) Section 4, marginal note 19

[118] The concept of "sui generis" stands for a separate type of legal right, because the right of database creators is not a form of copyright. The protection is independent of any copyright protection for database works under Section 4 of the Copyright Act (UrhG).

[119] Even if a database is assumed to be a work under Section 4 of the Copyright Act, protection for the database creator can fundamentally apply in addition (Dreier/ Schulze, part 2, Verwandte Schutzrechte (Related protection rights), P. 6, Schutz des Datenbankherstellers (Protection of the database creator), preliminary remarks, marginal note 8).

[120] Dreier/Schulze/Schulze Section 87a, marginal note 7.

[121] Higher Regional Court (OLG) Cologne MMR 2007, 443.

[122] Wandtke/Bullinger/Thum/Hermes Section 87a, marginal note 19.

[123] Higher Regional Court (OLG) Cologne MMR 2007, 443.

[124] A. A. Dreier/Schulze/Schulze Section 87a, marginal note 7: only the added information value of the data processing should be protected.

[125] Recital 40 of the Database Directive 96/9/EC.

[126] European Court of Justice, ruling of 9 Nov. 2004 — C-203/02.

[127] Dreier/Schulze/Dreier UrhG Section 87a marginal note 13

[128] European Court of Justice, ruling of 9 Nov. 2004 — C-203/02.

[129] BeckOK Koch (UrhG) Section 87a marginal note 21.

[130] Wandtke/Bullinger/Thum/Hermes UrhG Section 87a, marginal note 5.

[131] Recital 46 of the Directive 96/9/EC.

[132] Higher Regional Court (OLG) Cologne MMR 2007, 443.

[133] Bundestags-Drucksache (German parliament documents) 17/11470, 8.

[134] Cf. Bundestags-Drucksache (German parliament documents) 17/11470, 7.

[135] Wandtke/Bullinger/Jani UrhG Section 87f marginal note 12; Bundestags-Drucksache (German parliament documents) 17/11470, P. 6.

[136] Dreier/Schulze/Schulze UrhG Section 87f, marginal note 3.

[137] Higher Regional Court (OLG) Munich, ruling of 14 July 2016 — 29 U 953/16.

[138] European Commission, proposal for a directive of the European Parliament and the Council on copyright in the digital single market of 14 Sept. 2016, COM(2016) 593 final.

[139] Cf. Federal Constitutional Court, ruling of 14 March 2006 —1 BvR 2087/03, 1 Be 2111/03 —, BVerfGE 115, 205-259.

[140] Fischer, StGB (Commentary on the Criminal Code), 63th edition 2016, Section 202a marginal note 7a; Schönke/Schröder/Eisele/Lenckner (StGB) Section 202a marginal note 9; MüKoStGB/Graf StGB Section 202a marginal note 19.

[141] Bundestags-Drucksache (German parliament documents) 16/3656, P. 9, 10.

[142] MüKoStGB/Graf StGB Section 202a marginal note 26.

www.smart-data-programm.de