

SURVEY TO SUPPORT AN IMPACT ASSESSMENT FOR THE REVIEW OF THE DATABASE DIRECTIVE

Cover letter

The successful implementation of data-driven business models is of great importance for a growing number of industries with their often-multiple roles as data producer, user, seller, and buyer. In the data economy, machine generated data, including sensor data, are key assets; access to data, exclusive use of data, but also sharing of data and protection of rights address important questions for the competitiveness of European firms.

In this context, the European Commission, DG CNECT is currently preparing an impact assessment on a potential revision of the Database Directive, which was adopted in 1996. The assessment aims to understand if the Directive is still fit for purpose in the data economy, and in particular the relationship of the Directive with machine generated data.

Under EU law, protection of databases against unauthorised use may take two forms with **Copyrights** – which protect the database structure (a creative act) and the **Database rights** (or ‘sui generis’ right). This right *protects the investment* in either obtaining, verifying or presenting the database contents, but not the content as such. **The current survey concerns only the database (sui generis) right.**

On behalf of the European Commission, Copenhagen Economics and Technopolis Group conduct this industry survey. It aims to obtain insights on the use (and usefulness) of the Directive, views on possible revisions (policy options) and their impacts. Given the specificity of the topic, we would expect that senior legal and/or IT managers will be best placed to respond to the questions.

We would appreciate very much your support in improving this legislation to the benefit of the business sector.

Please send any queries to the following email address: databasedirective-survey@technopolis-group.com.

The survey will be open until **xx** June 2021.

In accordance to GDPR requirements, the survey results will only be used in aggregated form (at country or industry level). No personal data is collected.

I. YOUR ORGANISATION AND YOUR PROFILE

Q1

In which country is your organisation located? (country of residence of operating unit)

Drop-down list of EU MS, option: other

Q2

Approximately how many people are currently employed (full-time or part-time) in your organisation, including all branches, divisions and subsidiaries?

10-49

50-249

250-499

More than 500

Don't know

Q3

Which of the following industries best describes your organisation's primary business?

- Agriculture**
- Banking**
- Insurance**
- Business or professional services, excluding IT services**
- IT services**
- Healthcare**
- Process manufacturing**
- Discrete manufacturing**
- Retail trade**
- Wholesale trade**
- Telecommunications**
- Media**
- Transport and logistics**
- Utilities**
- Oil and gas**
- Other, please specify**

Q4

Which of the following best describes your position within your organisation?

- IT expert**
- Senior data engineer**
- Legal expert**
- Senior Legal Officer/ Councilor**
- Other, please specify**

**SINGLE
CHOICE**

**Branching
for IT**

**Branching
for IT**

What do we mean with machine generated data

Machine-generated data (MGD) are data recorded, collected, or generated independent of direct human intervention by:

- sensors processing information received from equipment, software or machinery, whether virtual or real
- computer processes, applications or services.

Please note that MGD includes machine generated data that are the result of observation of (acts of) humans (e.g. which pages within a website a person visits). However, data that are the result of humans consciously providing information/choices are excluded (persons typing in their name, address etc. to create an account)

Q5

Please select what describes you and your organisation best

**[MULTIPLE
Choice/Select Box]**

User of databases including primarily machine-generated data

User of databases including primarily other type of data than machine generated data

Maker of databases including primarily machine-generated data

Maker of databases including primarily other type of data than machine generated data

Don't know / Not applicable

Q6

Which of the categories would further describe your organisation? [Multiple Choice/Select Box]

[conditional
on answer :
Producer/user
of machine
generated
data in Q5

Data holder (organisations competent to decide about the use of data regardless of who collected, stored, processed or disseminated the data, e.g. private sector companies, which are the IoT product/service providers)

Data co-producer (organisations which are IoT product/service users, e.g., transport companies, airlines)

Data re-user (players interested in accessing data and/or re-using the data from a data holder, e.g. data analytics companies, users and repairers of smart devices)

Data intermediary (organisations that enable data holders to share their data, e.g., data marketplaces or industrial data platforms which enable sharing of data)

II. QUESTIONS ON THE TECHNOLOGICAL-COMMERCIAL CONTEXT OF DATA AND DATABASES IN YOUR COMPANY

Q7 How important are the following use cases of machine generated data in your organisation?

[conditional on answer : Producer/user of machine generated data in Q5]

	Very important	Somewhat important	Neutral	Rather unimportant	Not important	Don't know
Generating and using data for internal use only (e.g., optimisation of processes)	<input type="checkbox"/>					
Selling/licensing data to third parties for a fee (commercialising data)	<input type="checkbox"/>					
Providing free services to third parties (e.g., clients, suppliers) based on data	<input type="checkbox"/>					
Providing fee-based services to third parties (e.g., clients, suppliers) based on data analytics	<input type="checkbox"/>					
Obtaining data from third parties through contracts/licensing for own data analytics or optimisation of production processes/devices	<input type="checkbox"/>					
Obtaining data from device manufacturers through contracts/licensing to provide aftermarket services such as repair and maintenance and development of applications	<input type="checkbox"/>					
Sharing data for free on a mutual basis with business partners on a platform	<input type="checkbox"/>					
Using data to design innovative solutions/products	<input type="checkbox"/>					
Using big data for training AI	<input type="checkbox"/>					
Other, please specify.....	<input type="checkbox"/>					

Q8 What is your opinion on the use of machine generated data?

[conditional on answer :
Producer/user
of machine
generated data
in Q5

	Strongly agree	Agree	Dis-agree	Strongly disagree	Don't know
Data from our internal processes are valuable only for organisation-internal deployment	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
My organisation invested substantially in our data infrastructure (IT system, database structure, protection, training, etc.)	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
Selling or licensing our data to third parties provides extra profits	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
Selling or licensing our data to third parties endangers our exclusive data analytics insights	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
Selling or licensing our data to third parties enables wider user/consumer benefits	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
Our organisation benefits from buying third party data	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
Sharing data for free on a mutual basis with business partners/on a platform benefits our organisation	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
Our organisation experienced difficulties when requesting access to other companies' data	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	
Other, please specify	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	

III.GENERATION AND COLLECTION OF DATA AND DEVELOPMENT OF DATABASES

Which type of database for machine generated data do you use from a technological point of view: [Multiple Choice/Select Box] [conditional on answer : Producer/user of machine generated data in question [5]

SQL

NSQL

In-memory

Distributed databases

We are not running databases ourselves but using cloud or other data services

Other

Not applicable

Q9

Are you cooperating with other companies in establishing and running databases?:

[conditional on answer : Producer/user of machine generated data in Q5

Yes, to a wide extent

Yes, in many cases

Yes, only in few cases

No

No, but we plan to

Don't know

Q9a

Are these cooperating partners...

If Yes:
Multiple selection possible)

Firms within the same industry

Firms from other sectors

Suppliers

Customers

Other

Q10

If you collect machine generated data, where is the data stored?

[conditional on answer : Producer/user of machine generated data in Q5

Directly in a database in my company

In a database in a partner company

In the cloud

In an intercompany network

Data remain in IoT device, but can be read from a distance at any moment

Other

Don't know / Not applicable

Q11

If your company provides supplementary data analytics, will the resulting data be stored in: (multiple choice)

[conditional on answer : Producer/user of machine generated data in Q5

The original database

A new database

Don't know / Not applicable

Q12

Is machine generated data collected in your company considered only as a by-product of your primary activity?

[conditional on answer : Producer/user of machine generated data in Q5

Yes, it is central for the good functioning of our operation

Yes, it is secondary to our business

No, but it is becoming increasingly relevant in our operation

No it isn't

Don't know / Not applicable

Q13

If machine generated data is collected by your company, who invests and takes the economic risks of setting up and maintaining the database?

conditional on answer :
Producer/user of machine generated data in Q5

My company

Manufacturer of sensors

Operator of devices (e.g. in tractors, autonomous cars)

Participants in a network (i.e. different stakeholders contributing to database)

Don't know / Not applicable

Other, please specify

Q14 Please provide your best estimate of cost per annum to set up and maintain on average a database in your company: conditional on answer : Producer/user data in Q5

	<500 eur	< 5,000 eur	< 20,000 eur	< 100,000 eur	Above 100,000 eur	I don't know
One-time investment						
Operating costs						

Please elaborate as you wish:

Q15

If the machine generated data is generated/collected by another person or company (controller of the collection device), does the data holder grant you: **[conditional on answer : user of machine generated data in question [5]** Multiple choice

- Right to access/visualise the data
- Right to download/ use the data
- Ownership on data
- Right to re-use/modify the data
- Access was denied
- Data was shared on a mutual basis
- Other
- Don't know / Not applicable

Q16

Would you be able to estimate the costs to establish a contract (e.g., data sharing agreement) to access the data? (contracting costs may include internal and external resources such as legal service, if your organisation contracted in multiple instances please provide an estimation for an average contract, organisation can incur costs even when access was denied):

[conditional on providing answer other than "not applicable" to question 15: "If the machine generated data is generated/collected by another person or company (controller of the collection device), does the data holder grant you"]

- Less than 1,000 EUR
- 1,000 – 5,000 EUR
- 5,000 – 10,000 EUR
- 10,000 – 20,000 EUR
- More than 20,000 EUR
- Other, please specify
- Don't know

IV. SHARING AND ACCESS TO DATA AND DATABASES

Q17

In the last 12 months, how often has your company shared its databases containing machine-generated data or requested access to/use of another company's databases? [Single choice][Select box]
 conditional on answer : Producer/user of machine generated data in question [5]

On a daily basis

Several times

Only a few times

Don't know

Sh- Ac-
ring ces-
da- sing
ta- other
ba- data-
ses ba-
ses

Q18

In the last 12 months, has your company denied a request of access to databases containing machine-generated data? [Single choice][Select box]
 conditional on answer : Producer/user of machine generated data in question [5]

Yes

No

Don't know

Not applicabile

If Yes, for which reason? [Multiple choice][Select box]

Too costly infrastructure investments (allowing APIs, define protocols,...)

Potential loss of competitive advantage from exclusive access to data

Potential loss of revenues of exclusive rights on data

Other, please specify

Q19 Companies may become the sole source of certain data contained in protected databases establishing a de facto monopoly. Please indicate, to which statement you agree or disagree

	Strongly Agree	Agree	Disagree	Strongly disagree	Don't know
For databases containing machine-generated data I cannot obtain this or similarly useful data from other sources	<input type="checkbox"/>				
For other type of databases I cannot obtain this or similarly useful data from other sources	<input type="checkbox"/>				

Q20

When you tried to obtain access to databases containing machine generated data, did you encounter any problems? [Single choice][Select box]

Yes

No

Don't know/ Not applicable

If Yes, for which reason? [Multiple choice][Select box]

No market for the type of database needed

The data in the database was kept secret

The database was legally protected and there was no licensing available

Too high costs to obtain access to the database (licensing costs, infrastructure developments,...)

The database was protected with technical measures

Lack of interoperability

Don't know / Not applicable

Other, please specify

Q21	Do you think there should be mandatory access e.g through a licence, open access, etc. to machine generated data: [Multiple choice][Select box]	for the co-producer	for the re-user	for the intermediary	for none of them	I don't know
	All machine generated data should be accessible...	<input type="checkbox"/>				
	Certain data or databases, depending on their nature (e.g. of specific public interest, such as mobility data needed for decisions on the extension or reduction of the road infrastructure)...	<input type="checkbox"/>				
	Certain data or databases, depending on the business situations (e.g. sole source database, where data is not available from other sources) ...	<input type="checkbox"/>				
	There should not be mandatory access...	<input type="checkbox"/>				

Q21a	If selected "there should not be mandatory access" What are the reasons why a mandatory access should not be granted?	[Multiple choice][Select box]				
	it would affect my business interest	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	it would affect my freedom to contract too much	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Other, please specify	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q21b	If you think there should be access to databases containing machine generated data, how should access be granted? [conditional on answer first three options to q 21: "Do you think there should be mandatory access to machine generated data"]	[Single choice][Select box]
	As a free access	<input type="checkbox"/>
	As a compulsory license on FRAND terms	<input type="checkbox"/>
	Don't know	<input type="checkbox"/>

V. SUI GENERIS AND OTHER TYPES OF DATABASE PROTECTION MEANS

EU legal protection of databases: the *sui generis* right:

The *sui generis* right is a special intellectual property right protecting substantial investment in either obtaining, verifying or presenting the database contents under the EU's Database Directive.

Under this right, substantial extractions and re-use are subject to rightsholder's prior consent.

Under the Directive, some contractual restrictions on the actions of lawful users of a database are ineffective. The *sui generis* right protects against taking of substantial parts as well as insubstantial parts that may cause damage to the investment due to use in competition. However, if a database is not protected by the *sui generis* right (and by copyright), the operator of the database is free to impose whatever contractual restrictions are permitted under the relevant national law. Further protection may arise at national level, mainly via unfair competition. This *sui generis* right is not granted to entities operating outside the EU. So, an investment by a firm based in the United States in production of a database does not give rise to *sui generis* right in the EU (and there is no such right in the US).

Q24

How familiar are you with the Database Directive? [Multiple Choice][Select Box]
[Branching question to identify respondents aware of the DbD able to answer *sui generis* specific questions that follow]

I have been subject to a legal claim due to a database right

I have enforced the database right myself

I am familiar with the Database Directive

I am aware of it but I am not familiar

I do not know the Database Directive

Q25 How would you rate the following alternative means of protecting databases containing machine generated data:

[Consider adding condition to only respondents that are producer/users of MGD]

	Very effective	Partially effective	Not effective	Don't know
Exclusive rights (e.g. sui generis right)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Competition law (enforcement through litigation)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Contracts (e.g., through detailed licensing agreements)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Trade secrets (e.g., through confidentiality obligations)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Technological measures (e.g., access restrictions, encryption)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Smart contracts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other, please specify	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q26 How costly do you assess the following means of protecting databases containing machine generated data for the data holder: on a scale from 1 (no cost) to 5 (very high cost)?:

	1 - no costs	2 – Minor	3 - Considerable	4 - High	5 – Very high	I don't know	
Exclusive rights (e.g. sui generis right)	<input type="checkbox"/>						
Competition law (enforcement through litigation)	<input type="checkbox"/>						
Contracts (e.g., through detailed licensing agreements)	<input type="checkbox"/>						
Trade secrets (e.g., through confidentiality obligations)	<input type="checkbox"/>						
Technological measures (e.g., access restrictions, encryption)	<input type="checkbox"/>						
Smart contracts	<input type="checkbox"/>						
Others	<input type="checkbox"/>						

Q27

In your view, what could be the benefit of an exclusive right that covers databases containing machine generated data and that you can exercise against anybody, not just people with whom you have a contractual relationship? [Multiple Choice][Select Box]

It offers the opportunity to better regulate the relationship with clients, including licences

It offers protection against third-party infringements (i.e. unauthorised use of machine generated data) which is highly needed

It offers protection without triggering unnecessary costs

It would provide an incentive to license or share the data from my database

In my view there is no considerable benefit

Other

Q28 Which means do you use or plan to use to protect the databases containing machine generated data against unauthorised use? [limited to database owners thus Q6, data holders]

	Not using and no plans to use	Already using	Plan to use in the next 12 months	Not aware of	Don't know
Database sui generis right	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Copyright	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Competition law	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Contracts (e.g., through detailed licensing agreements)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Trade secrets (e.g., through confidentiality obligations)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Technological measures (e.g., access restrictions, encryption)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Smart contracts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other, please specify.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

[All remaining questions conditional on NOT answering “I do not know the Database Directive” / I am aware of it but I am not familiar to Q 24: “How familiar are you with the Database Directive?”]

Q29

To what extent is sui generis database right useful for the purposes of your data protection compliance, e.g. to protect personal data from unauthorized access, etc.? [Single choice][Select box] [should be limited to database owners thus Q6, data holders]

Useful

Potentially useful, but we have not used it in practice

Not useful at all

Not applicable / Don't know

VI. POLICY OPTIONS –

Q30 What is your view on the following options for a potential change in the Database Directive specific to machine generated data ? (Please select on a 1-5 scale from 1 – disagree strongly and 5 – agree strongly)

	1 - Disagree strongly	2 - Disagree	3 - Neutral	4 - Agree	5 - Agree strongly	Don't know
Machine-generated data would be explicitly excluded from the scope of application of the sui generis right	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Replacing the sui generis right, particularly as applied to MGD, with an alternative protective mechanism against certain unauthorized uses that could be categorized as abusive or unfair from competition perspective	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Machine-generated data would be included in the scope of application of the sui generis right together with a new data access regime (e.g. new exception or compulsory licences)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The sui generis right would only apply to machine-generated data if the maker of the database takes an additional step, namely taking certain proactive protection measures (such as technical measures) ¹	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

¹ These could be: (i) clearly expressed intention to apply the protection of the sui generis right, e.g. in machine-readable format, (ii) application of reasonable technical protection measures to mark the protected nature of the machine-generated database

Q31 What is your view on the following options for remaining with the status quo or a potential general change in the Database Directive? (Please select on a 1-5 scale from 1 – disagree strongly and 5 – agree strongly)

	1 - Dis- agree strongly	2 - Dis- agree	3 - Neutral	4 - Agree	5 - Agree strongly	Don't know
The Database Directive would remain in its current form	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The sui generis right would be repealed	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Public bodies would be excluded from the sui generis right	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Extension of lawful user rights	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Exceptions to sui generis right would be expanded in line with broader general copyright exceptions like private copying and research	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Broaden the sui generis right to include insubstantial parts (with potential extension to exception or the lawful user right).	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q32 Compared to the current situation, how would you assess the following options for change in terms of costs and benefits for your company? Taking 0 as the current situation, please assign a score from -5 to +5 to the costs and to the benefits of each option:

	Cost	Benefit
The sui generis right would be repealed.		
Excluding machine-generated data from the scope of application of the sui generis database right.		
Replacing the sui generis right, particularly as applied to MGD, with an alternative protective mechanism against certain unauthorized uses that could be categorized as abusive or unfair from competition perspective.		
Machine-generated data would be included explicitly in the scope of application of the sui generis right together with a new data access regime for the benefit of the database users (e.g. new exception or compulsory licences) for the benefit of database users .		
The sui generis right would only apply to machine-generated data if the maker of the database takes an additional step, namely taking certain proactive protection measures (such as technical measures ²).		

Please feel free to elaborate on potential costs and benefits:

² These could be: (i) clearly expressed intention to apply the protection of the sui generis right, e.g. in machine-readable format, (ii) application of reasonable technical protection measures to mark the protected nature of the machine-generated database

Q33 Compared to the current situation, how would you assess the following supplementary options for changing some further aspects of the sui generis right for change in terms of costs and benefits for your company? Taking 0 as the current situation, please assign a score from -5 to +5 to the costs and to the benefits of each option:

	Cost	Benefit
Public bodies would be excluded from the sui generis right		
Extension of lawful user rights		
Exceptions to sui generis right would be expanded in line with broader general copyright exceptions like private copying and research		
Broaden the sui generis right to include insubstantial parts (with potential extension to exception or the lawful user right).		

Please feel free to elaborate on potential costs and benefits:

Q36 What is your view on the following statements:

An exclusion of machine generated data from the database right increases the risk of...

	1 - Dis-agree strongly	2 - Dis-agree	3 - Neutral	4 - Agree	5 - Agree strongly	Don't know
Contractually based restrictions imposed by the database holders	<input type="checkbox"/>					
Decrease of available data due to the decrease of trust in sharing/providing access to databases (Data lock-ins)	<input type="checkbox"/>					
Users to lose access and use regime included in the Directive	<input type="checkbox"/>					
Other, please specify	<input type="checkbox"/>					

Q37 Would you be willing to participate in an in-depth interview on the subject with experts of this study? Please provide your email address