

EPFL Applied Machine Learning Days, 28 January 2019

# AI and your data: who owns the data, what if things go wrong, GDPR and AI ?




Michel Jaccard  
[www.idest.pro](http://www.idest.pro)

- ✓ Data rather than AI
- ✓ Data relationship with AI (at least, AML)
  - ✓ Technical and operational dependency (AI raw material for running and improving applications)
  - ✓ Commercial dependency (costs of supply, value | relevancy of outputs)
- ✓ One of the key legal questions with data (in relation with AI) is the ownership issue



No one can own data... thus, it's free (to collect, process, share, use, analyse, combine, inject, twist...)



The image shows a screenshot of the NASA Technology Transfer Program Public Domain website. At the top left, there is a logo for the NASA Technology Transfer Program with the tagline "BRINGING NASA TECHNOLOGY DOWN TO EARTH". To the right is the NASA logo. The main heading is "Public Domain" in large blue letters, followed by "NASA TECHNOLOGIES" in white on a green background. Below this is a search bar with a "Search" button and a "Newly Disclaimed" button. An illustration of an open cardboard box filled with colorful icons representing various technologies is positioned to the right of the search bar. Below the search bar, there is a paragraph of text explaining the purpose of the program and a disclaimer.

NASA TECHNOLOGY  
TRANSFER PROGRAM  
BRINGING NASA TECHNOLOGY DOWN TO EARTH

## Public Domain

### NASA TECHNOLOGIES

To stimulate the innovation economy, NASA makes a portion of its technology portfolio **freely available** for anyone to use.

The technologies in this public domain portfolio do not require a license agreement, and anyone may freely pursue independent product development right away without the need to contact NASA in any way.

- ✓ There is no specific «legal status» protecting data, or data has been explicitly labeled as part of the «public domain».
- ✓ Might be the case for entire sets of open or scientific data, e.g., NASA or CC-0 type licenses.
- ✓ Great for AI research, not so great for many commercial suppliers in the AI ecosystem (data providers) – no protection for «commercially valuable data» means no incentive to invest.



**DON'T TOUCH**

Don't you dare use my data. It's mine !

# Data ownership – an easy legal question ?

- ✓ Implies the acknowledgement by law that data has a legal, protected status, and can be traded, sold, bought, stolen, licensed, and eventually owned...
- ✓ Better perspective for data providers, but not so much for the AI ecosystem as a whole – you need permission to use data in the first place...

- ✓ Clearly the case for personal data, which are owned by the «data subject»... right ?
  - ✓ Ownership right vs right to access, erasure, transparency...
  - ✓ «Data subjects» vs «data controllers» and «data processors».
  - ✓ Consent of data subject is not the only legitimate ground for processing personal data, even under the brand new GDPR...

- ✓ Beyond privacy laws, can IP help me protect «my» data ?
  - ✓ Individuality | Originality issues for copyright protection.
  - ✓ Database rights not universally recognized.
  - ✓ Trade secrets rights promising (recent European Directive) but definition issues and not universally recognized either.





- ✓ No one knows for sure...
- ✓ Absent a clear legal regime, ownership on data generated by AI/AML systems are ... up to grab in contractual negotiations (in hidden or plain text) !

## PAAS AGREEMENT

*CUSTOMER shall be and remain the sole owner of any data produced by the Platform made available to CUSTOMER (DATA).*

*[CUSTOMER shall grant PROVIDER a fully paid up, non exclusive, worldwide license to access, run, modify or otherwise use DATA for the duration of this PaaS Agreement, for the limited purpose of performing hereunder.]*

## SCIENTIFIC RESEARCH COLLABORATION AGREEMENT

*Ownership of Data. All data generated by the activities of the Research Program shall be jointly owned between the Parties.*

*Party 1 Use of Data. Notwithstanding Section [x] above, Party 1 agrees that all data and information provided by Party 2 for use in the Research Program, and all data generated by the activities of the Research Program shall not be used or shared by Party 1 for projects with Third Parties.*

*Party 2 Use of Data. Notwithstanding Section [x] above, Party 2 agrees that all data and information provided by Party 1 for use in the Research Program, and all data generated by the activities of the Research Program, shall not be used or shared by Party 2 with any Party 1 Competitor.*

## GOOGLE ADVERTISING AGREEMENT

**Data Ownership.** *The Parties agree that, as between Company and Google, data or information collected by Company [...] or that is sent by Company to Google in connection with the provision of the Services [\*].*

*[\*] Indicates that certain information in this exhibit has been omitted and filed separately with the Securities and Exchange Commission. Confidential treatment has been requested with respect to the omitted portions.*

- ✓ **The problem with ownership**
  - ✓ Ownership triggers liability if things go wrong (third party claim for infringement, non compliance with data protection laws, cyberbreach,...).
  - ✓ Therefore, what you may seek is rather a non exclusive, mutual right to use the data you have access to, subject only to third party pre-existing rights (such as copyright, trade secret, or personal data).

- ✓ We need to solve the data ownership issue and better describe the rights each participant in the AI/ML ecosystem has on «data» in parallel to the resolution of all other legal questions surrounding AI.
- ✓ Urgent need for standards – soft law or hard regulation.



**Michel Jaccard**

*id est avocats*

**Email**      *Michel.Jaccard@idest.pro*

**URL**          *www.idest.pro*

**Twitter**      *@idestavocats*