



The journey of data for AI – and what the EU is doing to permit more travel

Malte Beyer-Katzenberger – European Commission

«AI firms will soon exhaust most of the internet's data»



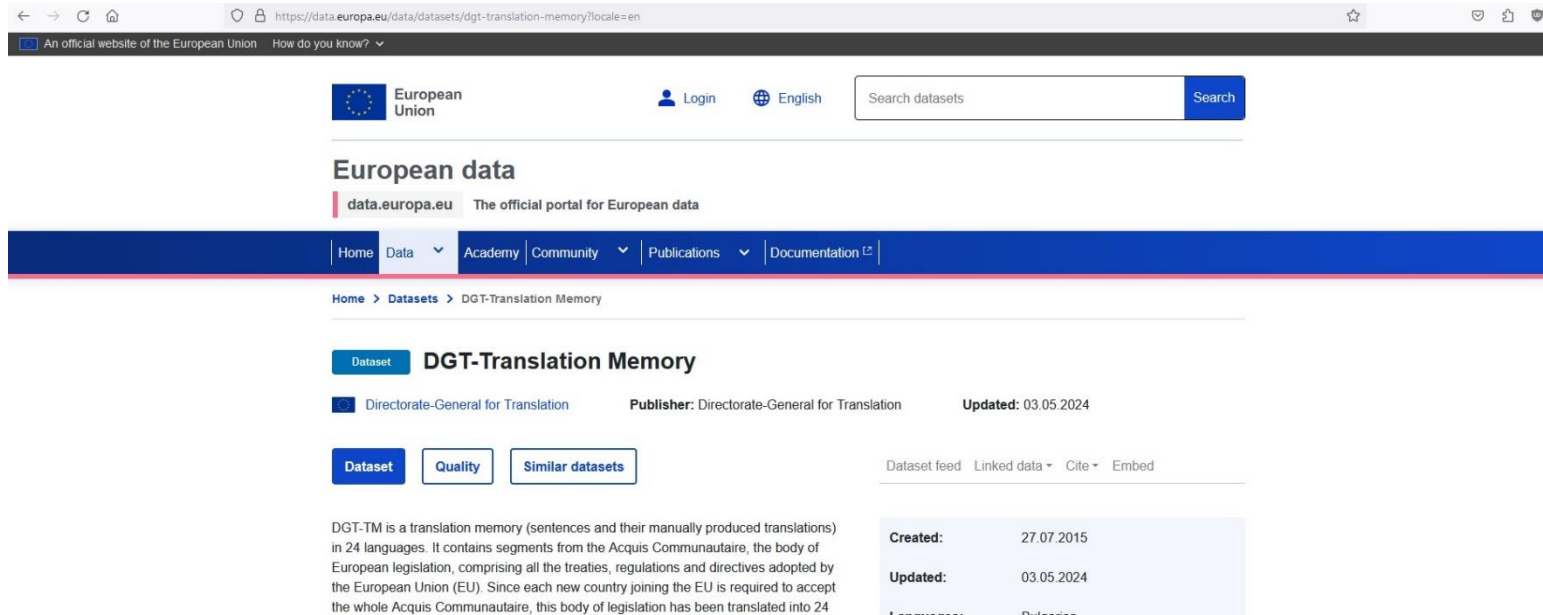
Schools brief | Artificial intelligence

AI firms will soon exhaust most
of the internet's data

Can they create more?

What has the EU done to open more data sources?

1) Providing certain data as „open data“



The screenshot shows the data.europa.eu website. The header includes the European Union logo, a login button, a language selector set to English, and a search bar. Below the header, the page title is "European data" with the subtitle "data.europa.eu The official portal for European data". A navigation bar contains links for Home, Data, Academy, Community, Publications, and Documentation. The main content area displays the "DGT-Translation Memory" dataset, published by the Directorate-General for Translation, with an update date of 03.05.2024. A description of the dataset is provided, and a table shows the creation and update dates.

European data
data.europa.eu The official portal for European data

Home > Datasets > DGT-Translation Memory

Dataset DGT-Translation Memory

Directorate-General for Translation Publisher: Directorate-General for Translation Updated: 03.05.2024

Dataset Quality Similar datasets

Dataset feed Linked data Cite Embed

DGT-TM is a translation memory (sentences and their manually produced translations) in 24 languages. It contains segments from the Acquis Communautaire, the body of European legislation, comprising all the treaties, regulations and directives adopted by the European Union (EU). Since each new country joining the EU is required to accept the whole Acquis Communautaire, this body of legislation has been translated into 24

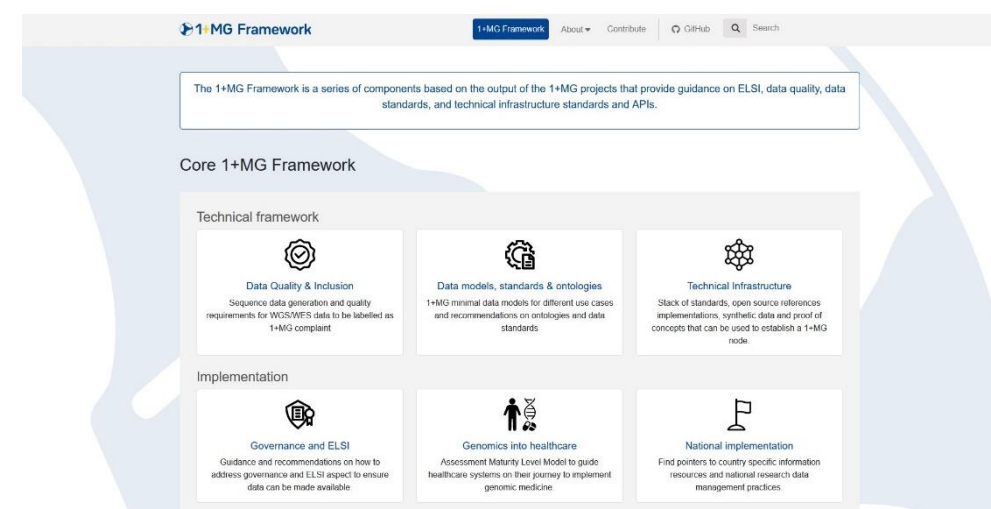
Created:	27.07.2015
Updated:	03.05.2024

What has the EU done to open more data sources?

2) Actively federate selected data repositories



Cancer Image Europe is pioneering a pan-European federated infrastructure for cancer images, fueling AI innovations



What has the EU done to open more data sources?

Working on overcoming the costs of collection where there is in principle the possibility to use certain data:

- 3) Create obligation to have data access and permit bodies in all EU countries for a range of personal health data (and create a right to use): European Health Data Space legislation

PRESS RELEASE | 15 March 2024 | Brussels | 1 min read

Commission welcomes political agreement on European Health Data Space

Page contents

[Top](#)

[Quote\(s\)](#)

[More](#)

The Commission welcomes the political agreement reached today between the European Parliament and the Council of the EU on the European Health Data Space (EHDS) - one of the **central building blocks of a strong European Health Union**. The rules [proposed](#) by the Commission in May 2022 have two main aims:

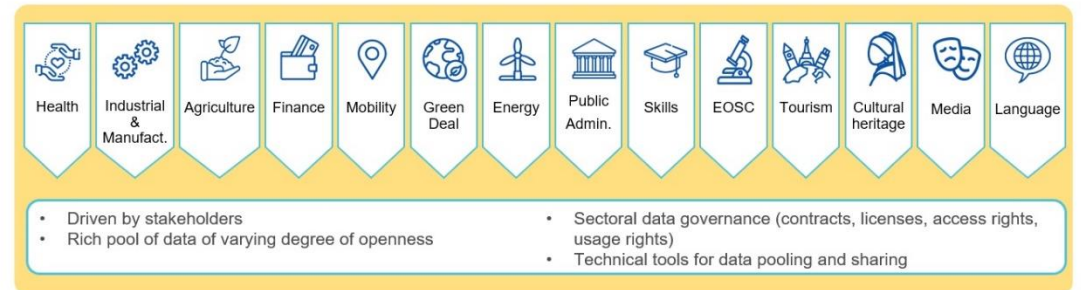
What has the EU done to open more data sources?

Working on overcoming the costs of collection where there is in principle the possibility to use certain data:

4) Support voluntary data sharing by building data sharing ecosystems (data spaces)

Common European Data Spaces

Digital Europe funding focuses on the creation of 14 data spaces, each linked to a strategic topic:



A Common European Data space brings together data infrastructures and governance frameworks

What has the EU done to open more data sources?

But who is gonna do me a „data-as-a-service“ inside such data spaces?

Do we need data fiduciaries/ data trusts

Can we push the legal frontiers on copyright and data protection?

**Should we?
Have we already?**

Can we knit the link between data spaces and AI factories closer?

Who would pay for data-pooling-as-a-service?

Summing up: Fishing for data



Crawl the web



Know your fish dealer
Become a processing
partner



Go to a fish market
(data space)

Thank you



© European Union 2024

Re-use of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license with the exception of the images on slide 2, 4 and 10. The image on slide 2 is © by Mike Hadad for The Economist. Images on slide 4 are screenshots of the websites to which links are provided by clicking on the image.

