Anastasija Nikiforova

Assistant Professor, University of Tartu

Data for AI or AI for data: artificial intelligence as a catalyser for open government ecosystems

19 – 20 MARCH 2025

ABOUT PRESENTER

ASSISTANT PROFESSOR, UNIVERSITY OF TARTU, FACULTY OF SCIENCE AND TECHNOLOGY, INSTITUTE OF COMPUTER SCIENCE, CHAIR OF SOFTWARE ENGINEERING

EUROPEAN OPEN SCIENCE CLOUD TASK FORCE “FAIR METRICS AND DIGITAL OBJECTS”

EDSC AMBASSADOR (EUROPEAN DIGITAL SKILLS CERTIFICATE, AS PART OF ACTION 9 OF THE DIGITAL EDUCATION ACTION PLAN (2021- 2027) – JRC/SVQ/2022/OP/0013)

DIGITAL GOVERNMENT SOCIETY (DGS)

ASSOCIATE MEMBER OF THE LATVIAN OPEN TECHNOLOGY ASSOCIATION

SCIENTIFIC & TECHNICAL ADVISORY BOARD OF SWITCH CONNECTOME

IFIP WG8.5 ON ICT AND PUBLIC ADMINISTRATION MEMBER

Governments worldwide,

at national and supra-national levels, with somehow

less progress at the subnational and local levels, are striving to

improve the FAIRness of their data

AI present unprecedented opportunities

to unlock the full potential of Open Government Data

EU OPEN DATA DAYS

OPEN DATA RHAPSODY

AI WILL ROCK YOU?!

19 – 20 MARCH 2025

DALL·E 3-generated content

Governments worldwide,

at national and supra-national levels, with somehow

less progress at the subnational and local levels, are

striving to improve the FAIRness of their data

HOWEVER,

the current understanding & use of AI are limited

AI present unprecedented opportunities

to specific AI implementations

to unlock the full potential of Open Government Data

in a specific setting of OGD - the portals

SO FAR...

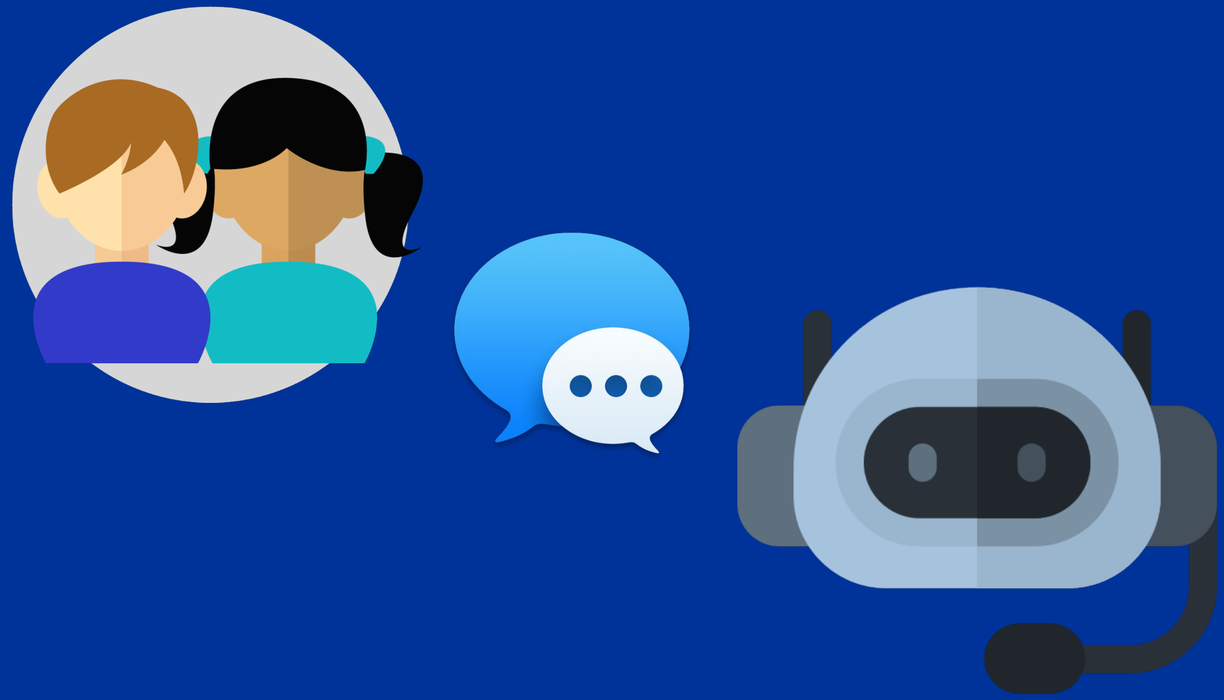


image: Flaticon.com'

SO FAR...

HOWEVER,

AI emerges as a broader avenue for transforming:

governance,

enhancing transparency,

fostering citizen engagement

in the OGD lifecycle and data ecosystem

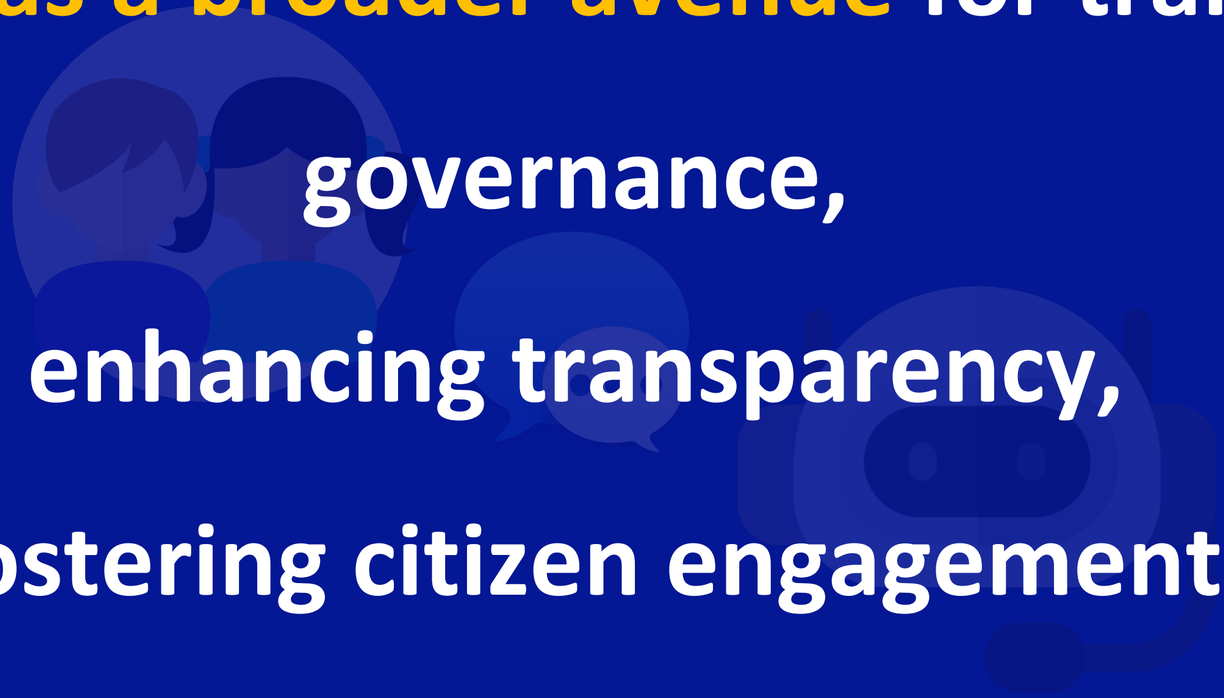


image: Flaticon.com'

OGD LIFECYCLE

Attard, J., Orlandi, F., Scerri, S. and Auer, S., 2015. A systematic review of open government data initiatives. *Government information quarterly*, *32*(4), pp.399-418.
Crusoe, J., Simonofski, A., Clarinval, A. and Gebka, E., 2019, May. The impact of impediments on open government data use: insights from users. In *2019 13th international conference on research challenges in information science (rcis)* (pp. 1-12). IEEE.

OGD LIFECYCLE

PRE-PROCESSING

data creation
data selection
data harmonization
data publishing

EXPLORATION

data interlinking
data discovery
data exploration

TRANSFORMATION

datasets combination
information/ product/ service development based on the aggregated data

MAINTENANCE

data curation to ensure sustainability of the data by raising awareness, cleansing, enriching and updating

OPEN OR PUBLIC DATA ECOSYSTEM

a dynamic network comprising interconnected elements that

enable a range of data-related activities that encompass the entire data lifecycle,

from collection and management to sharing and reuse,

involving diverse stakeholders with varied objectives

CONCEPTUAL MODEL OF A PUBLIC DATA ECOSYSTEM

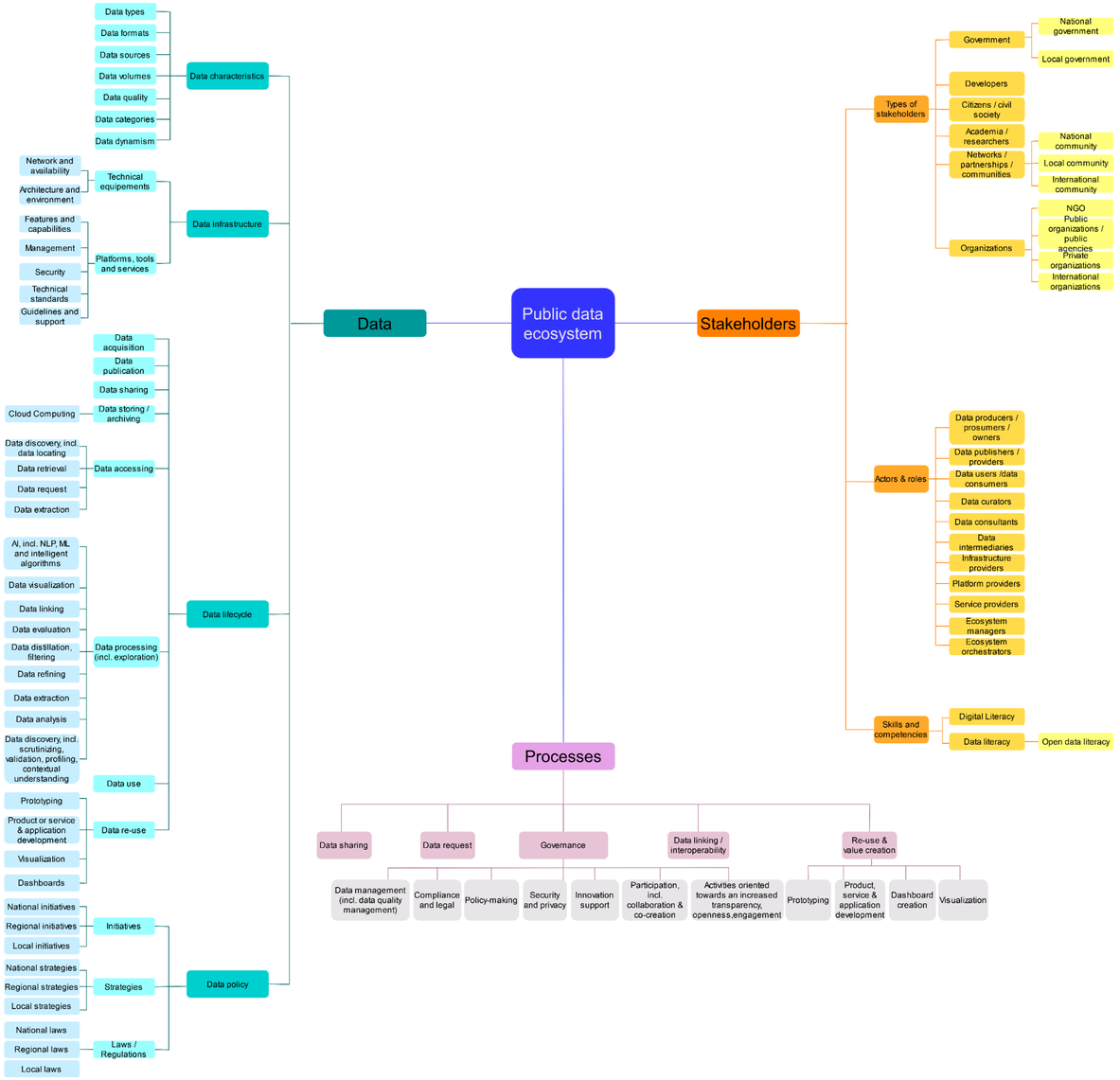
Open Government Data

Open Business Data

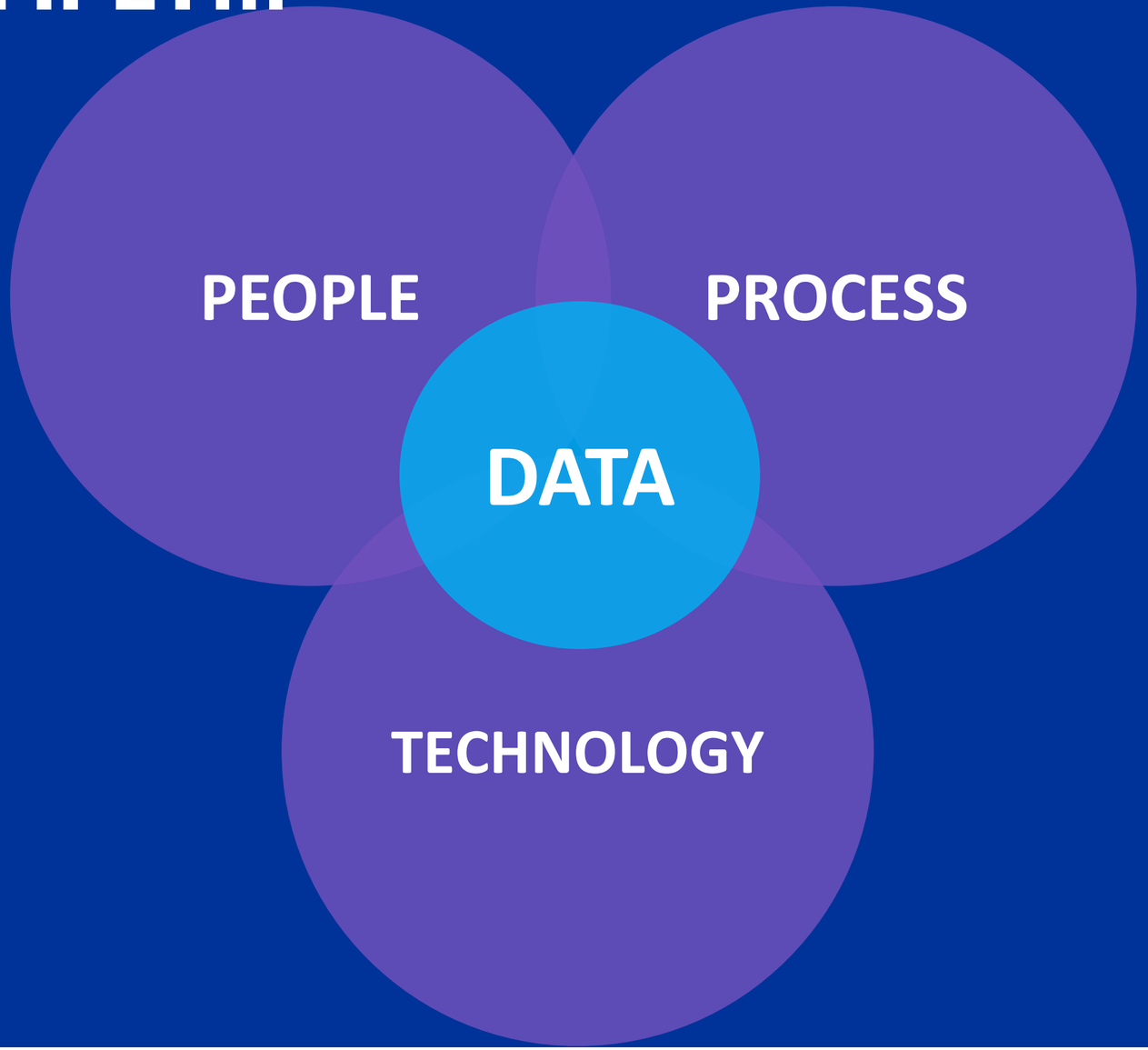
Open Citizen Data

Open Science Data

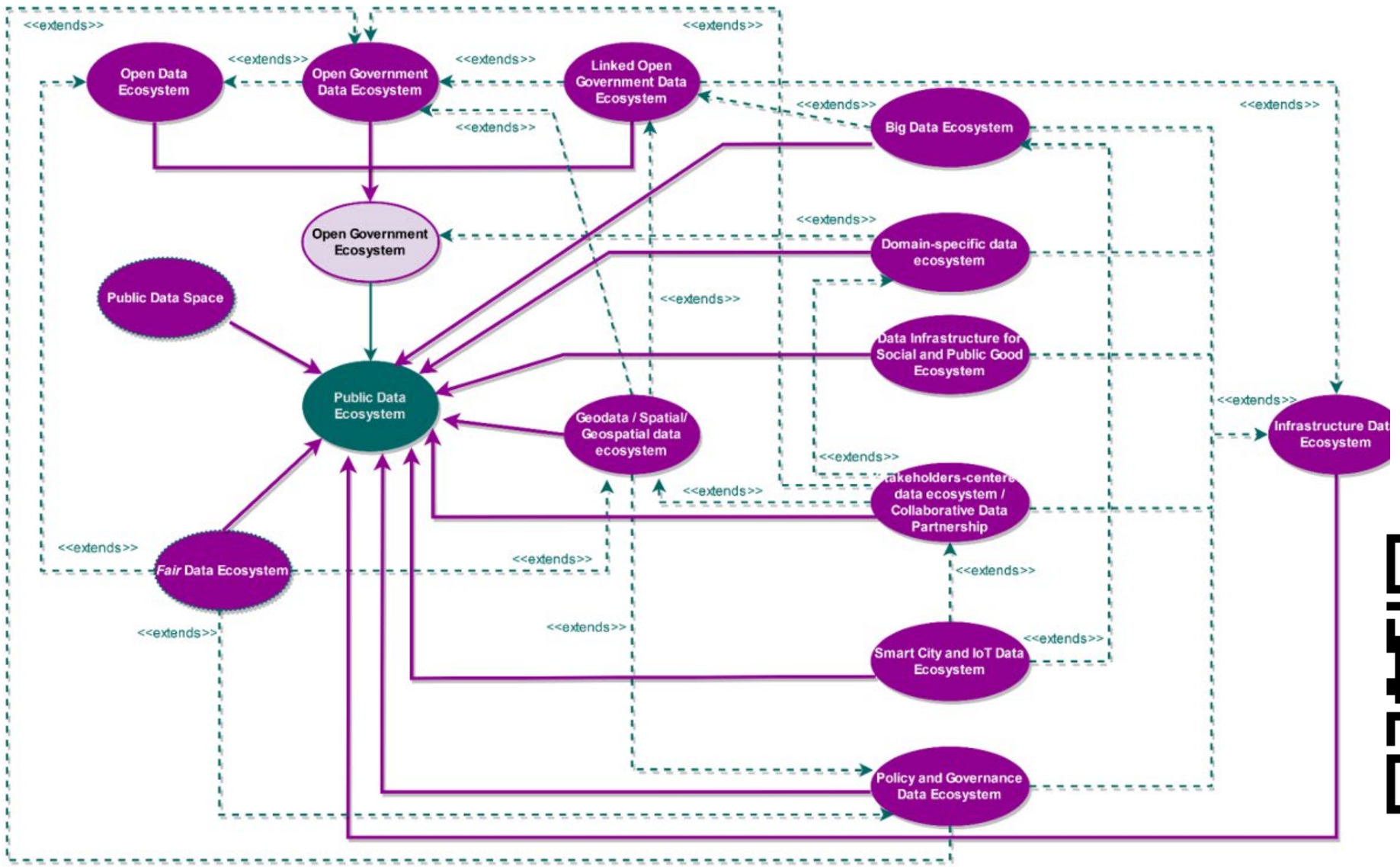
!!! THE EMERGING NEED FOR SUSTAINABILITY OF PUBLIC DATA ECOSYSTEMS



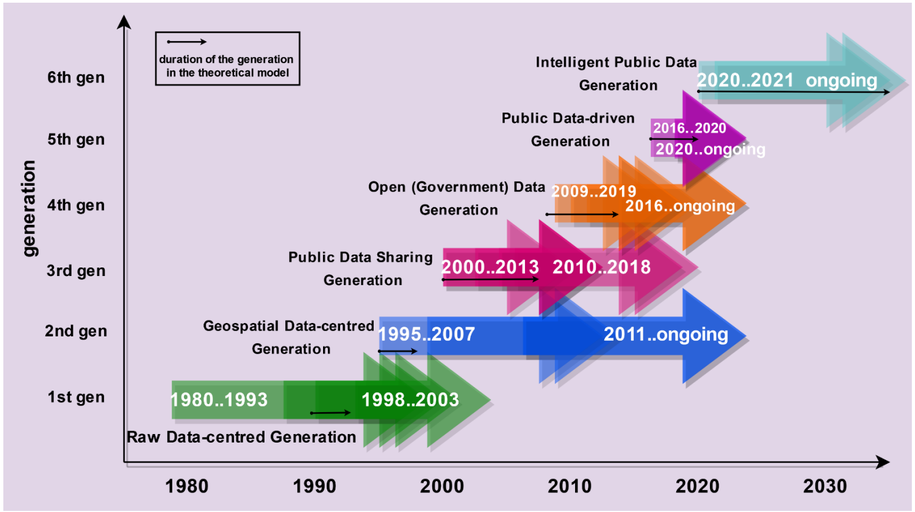
TO PUT IT SIMPLY…



CONCEPTUAL MODEL OF A PUBLIC DATA ECOSYSTEM



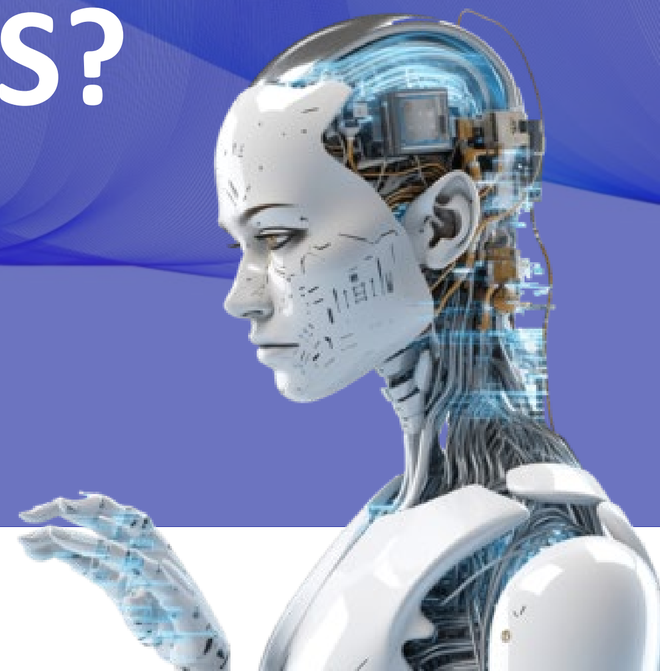
GENERATIONAL PERIODS OF PUBLIC DATA ECOSYSTEMS



	1st generation	2nd generation	3rd generation	4th generation	5th generation	6th generation NEW
Components and relationships	Data infrastructures, Data resources, ICT & data policies	Data policy and other related policies & laws, Data lifecycle management / data governance, (Geo)spatial data infrastructures, GIS and map portals, Data aggregator	Infrastructure and Technologies - Information Systems, platforms at different governmental levels, tools and services, Technical standards and guidelines, data and information policies	Data-related competencies & skills, Infrastructure & Technologies (platforms, tools, services), Data policy, policies & laws & strategies, Data structures, Dynamic data, Dynamics of processes and activities	External pressures – political, economic, environmental, ethical, and legal Data-related competencies and skills, Dynamics of processes and activities, Infrastructure & technologies (platforms, tools, and services)	**Data-related competencies and skills, Participatory mechanisms, 3D models for geospatial data, Big and open linked data, cloud computing, Intelligent algorithms, Artificial Intelligence, Machine Learning, Generative AI, LLM, Natural Language Processing tools, Data spaces, Data pools (lakes)**
Stakeholders	National and local government, Public organizations, Business	National and local government, Public organizations, Citizens, Investors and Entrepreneurs, Creators of Innovations, Researchers / Academia	NGOs, National and local government, Public agencies / Registries, Developers, Citizens, Private organizations - Business, International organizations, Supranational bodies (EU)	National and local government, Public agencies, NGO, Academia, Developers, Networks (regional, international – enthusiast-organized), Private organizations - Business (incl. SME), Citizens, International organizations, Supranational bodies (EU)	Public agencies, National and local government, NGOs, Networks, partnerships etc., Supranational bodies (EU), Private organizations - Business (incl. SME), Academia, Citizens, International stakeholders	**Networks, partnerships, NGO, Artificial Intelligence, Generative AI, National and local government, Research/academia, citizens, Private organizations - Business (incl. SME), International organizations, Supranational bodies (EU)**
Actors and their roles	Data producers, Data owners, Internal data users	Data providers, Data producers, Data prosumers, Data users, incl. residents/citizens, neighborhood activists, investors and entrepreneurs, creators of innovations, researcher	Data producers / Data owners, Data providers / data publishers, Data users, Data intermediaries (NGO, communities) Policies, laws, and rules parties	Data prosumers, Data publishers / providers, Data owners, Data users, Policies, laws, and rules parties, Ecosystem orchestrators, Data intermediaries, Infrastructure providers	Infrastructure providers, Platform providers, Service providers, Data prosumers, Data providers, owners, Data users, incl. systems, Policies, laws, rules parties, Ecosystem orchestrators, International stakeholders, Data curators	**Ecosystem managers, Data curators, Data stewards, Platform providers, Service providers, Data consultants, Data prosumers, Data publishers / providers, Data users, incl. systems, Policies, laws, and rules parties, International stakeholders, Artificial Intelligence, Generative AI**
Data Types	Data (no specific type), Statistical data	(Geo)spatial data, geodetic network,pollution, waste management, protected natural areas, environmental quality data	Public sector data, Open data, Metadata	Open data, Metadata, Linked data	Real-time (stream) data, Big Data, [Regional] geospatial / territory planning data, Priority / high-value data	**Intelligent (smart) data (from enthusiasts, business, networks), Priority - high-value data (HVD), SDG, AI- and Generative AI-ready data**
Processes & activities, data lifecycle phases	Data generation, Data digitization, Data management, Data transfer, Unification, Centralization, Standardization	Data sharing, Data transfer, Data processing, Data visualization, Data quality, Centralization, Unification, Standardization, Innovation, Asynchronous communication and feedback	Data sharing, Data generation, Innovation, Support, Participation, Collaboration, Engagement, Impact,Transparency, Centralization, Data processing, Interoperability (of ICT services)	Openness, Centralization, Unification, Optimisation, Data sharing,Transparency, Interoperability, Participation, Engagement, (Re)use, Innovation, entrepreneurship, Skills and competencies development	Decision-making, (Re)use, Impact, Interoperability, Skills and competencies development, Engagement, Data mashing, Sustainability	**Security, Interoperability, Skills and competencies development, Decision-making, Participation, Innovation, incl. Social innovation, Human-centric innovation, Asynchronous communication and feedback, Storytelling, Resilience**
	Raw Data-centred Generation	**Geospatial Data-centred Generation**	**Public Data Sharing Generation**	**Open (Government) Data Generation**	**Public Data-driven Generation**	**Intelligent Public Data Generation**

WHAT IS THE ROLE OF

ARTIFICIAL INTELLIGENCE IN OPEN DATA ECOSYSTEMS?



AI AS PORTAL **CURATOR**

AI AS DATA ECOSYSTEM **DATA RETRIEVER**

AI AS PORTAL **EXPLORER**

AI AS DATA ECOSYSTEM **CONNECTOR**

AI AS PORTAL **LINKER**

AI AS DATA ECOSYSTEM **VALUE DEVELOPER**

AI AS PORTAL **MONITOR**

AI AS DATA ECOSYSTEM **ENGAGER**

Image source: https://www.datamanagementblog.com/modern-data-ecosystem/

AI AS PORTAL CURATOR

Role:

AI facilitates the publication of OGD on the portal automating tasks such as tagging, categorizing, and data quality checks, incl. anonymizing sensitive data before publishing

Image source: https://www.datamanagementblog.com/modern-data-ecosystem/

AI AS PORTAL EXPLORER

Role:

AI helps users' exploration interfaces through tools such as natural language interfaces to access data and improved search capabilities, incl. AI chatbots

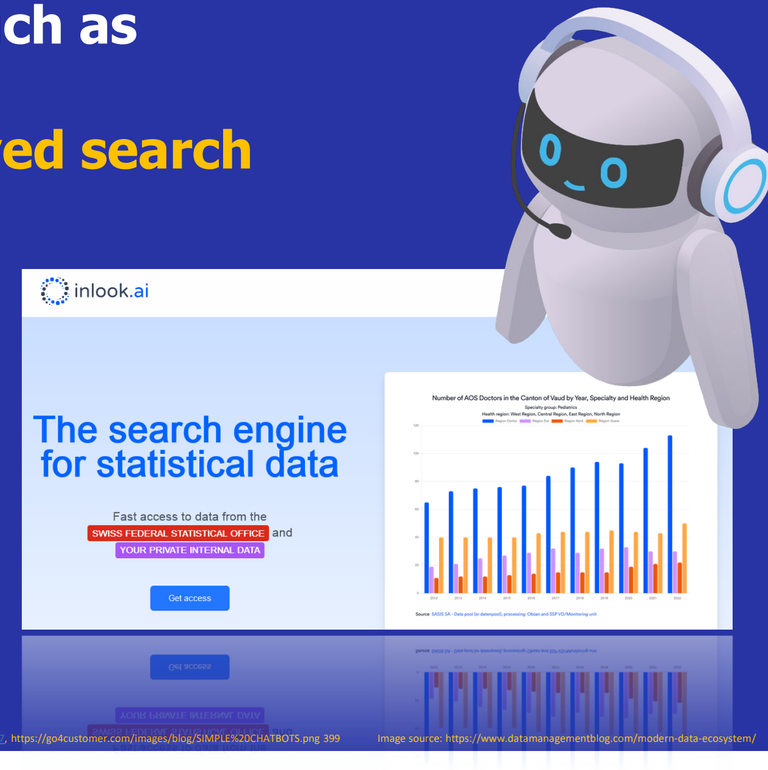


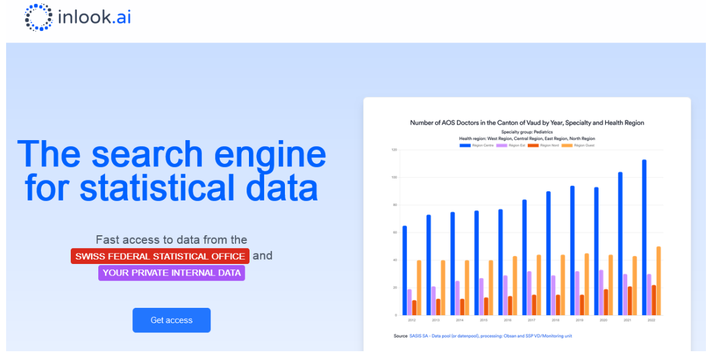
Image source: https://www.flaticon.com/free-icon/data-quality_10397, https://go4customer.com/images/blog/SIMPLE%20CHATBOTS.png 399

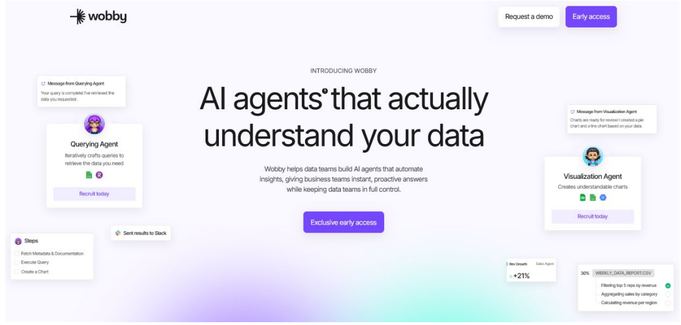
Image source: https://www.datamanagementblog.com/modern-data-ecosystem/

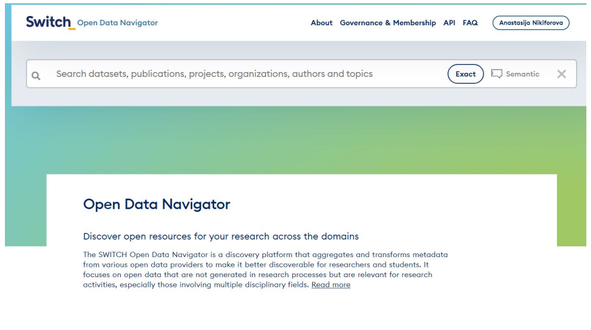
inlook.ai

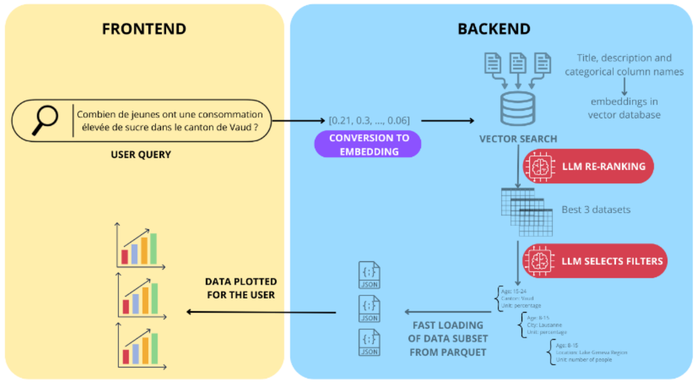
wobby

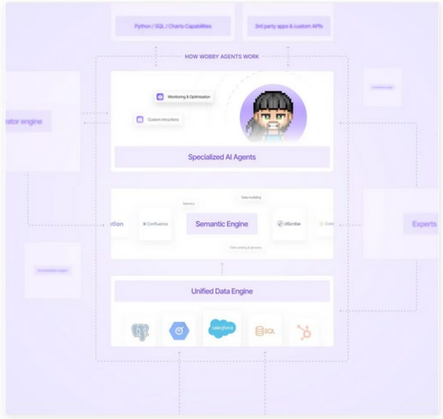
Switch Open Data Navigator

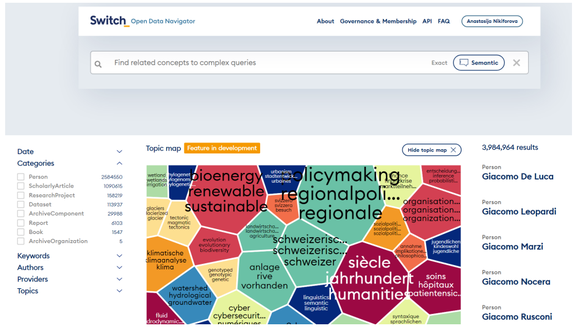












EU OPEN DATA DAYS

DATA FOR AI OR AI FOR DATA: ARTIFICIAL INTELLIGENCE AS A CATALYSER FOR OPEN GOVERNMENT DATA ECOSYSTEMS

AI AS PORTAL LINKER

Role:

AI transforms and aggregates OGD to present it optimally to developers and other users, e.g., transforming data into Linked OGD using ML

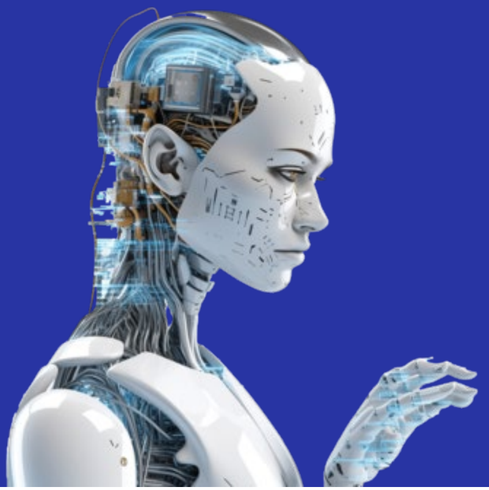


Image source: https://www.freepik.com/icon/data-link_5972054

AI AS PORTAL MONITOR

Role:

AI monitors and ensures portal compliance with standards,

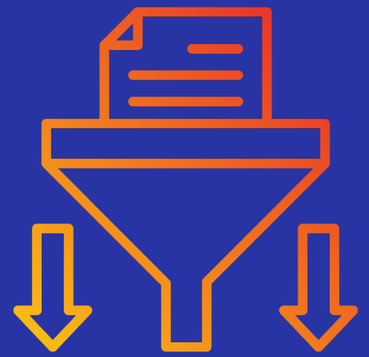
detecting anomalies, and improving metadata quality

Image source: https://www.flaticon.com/free-icon/standard_5576554

AI AS ECOSYSTEM DATA RETRIEVER

Role:

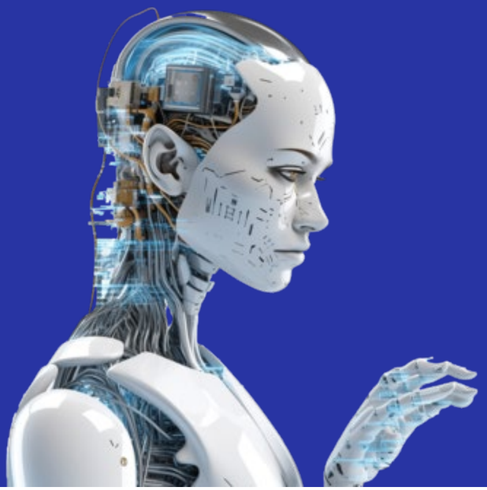
AI identifies and retrieves data from external sources, e.g., retrieving data from legal texts, into the OGD ecosystem, enriching the available data.



AI AS ECOSYSTEM CONNECTOR

Role:

AI connects actors of the ecosystems based on their interests by recommending datasets or other assets of interest.

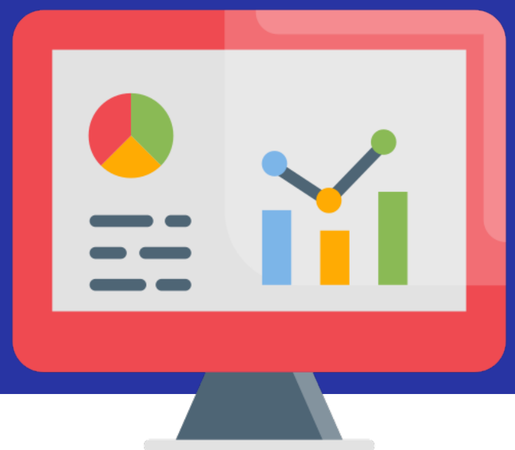


https://www.flaticon.com/free-icon/connection_5500939

AI AS ECOSYSTEM VALUE DEVELOPER

Role:

AI supports value creation by supporting services or product development using OGD, e.g., AI-powered dashboards for visualizing and analyzing OGD or suggesting appropriate analytics technique.



AI AS ECOSYSTEM ENGAGER

Role:

AI engages stakeholders in continued participation in the OGD ecosystem, e.g., through predicting the use of the portal to improve user engagement.

Mage source: https://www.flaticon.com/free-icon/prediction_9304477

PERSPECTIVE/ OGD LIFECYCLE	PRE-PROCESSING	EXPLORATION	TRANSFORMATION	MAINTENANCE
Inward Looking – Portal Perspective	**AI as Portal Curator** AI facilitates the publication of OGD on the portal automating tasks such as tagging, labeling, categorizing, and quality checks, incl. anonymizing sensitive data before publishing using AI-based privacy risk models for publishing data	**AI as Portal Explorer** AI helps users' exploration interfaces to access data through tools such as natural language interfaces to access data and improved search capabilities, as well as AI chatbots in dataset searches on OGD portals	**AI as Portal Linker** AI transforms and aggregates OGD to present it optimally to developers and other users, e.g., transforming data into Linked OGD using ML	**AI as Portal Monitor** AI monitors and ensures portal compliance with standards, detecting anomalies, and improving metadata quality.
Outward Looking – Ecosystem perspective	**AI as Ecosystem Data Retriever** AI identifies and retrieves data from external sources, e.g., retrieving data from legal texts, into the OGD ecosystem, enriching the available data.	**AI as Ecosystem Connector** AI connects actors of the ecosystems based on their interests by recommending datasets or other assets of interest.	**AI as Ecosystem Value Developer** AI supports value creation by supporting services or product development using OGD, e.g., AI-powered dashboards for visualizing and analyzing OGD or suggesting appropriate analytics technique .	**AI as Ecosystem Engager** AI engages stakeholders in continued participation in the OGD ecosystem, e.g., through predicting the use of the portal to improve user engagement.

CURATE

EXPLORE

LINK

MONITOR

RETRIEVE

CONNECT

DEVELOP

ENGAGE

CURATE

AI as Portal Curator

EXPLORE

AI as Portal Explorer

LINK

AI as Portal Linker

MONITOR

AI as Portal Monitor

RETRIEVE

AI as Ecosystem Data Retriever

CONNECT

AI as Ecosystem Connector

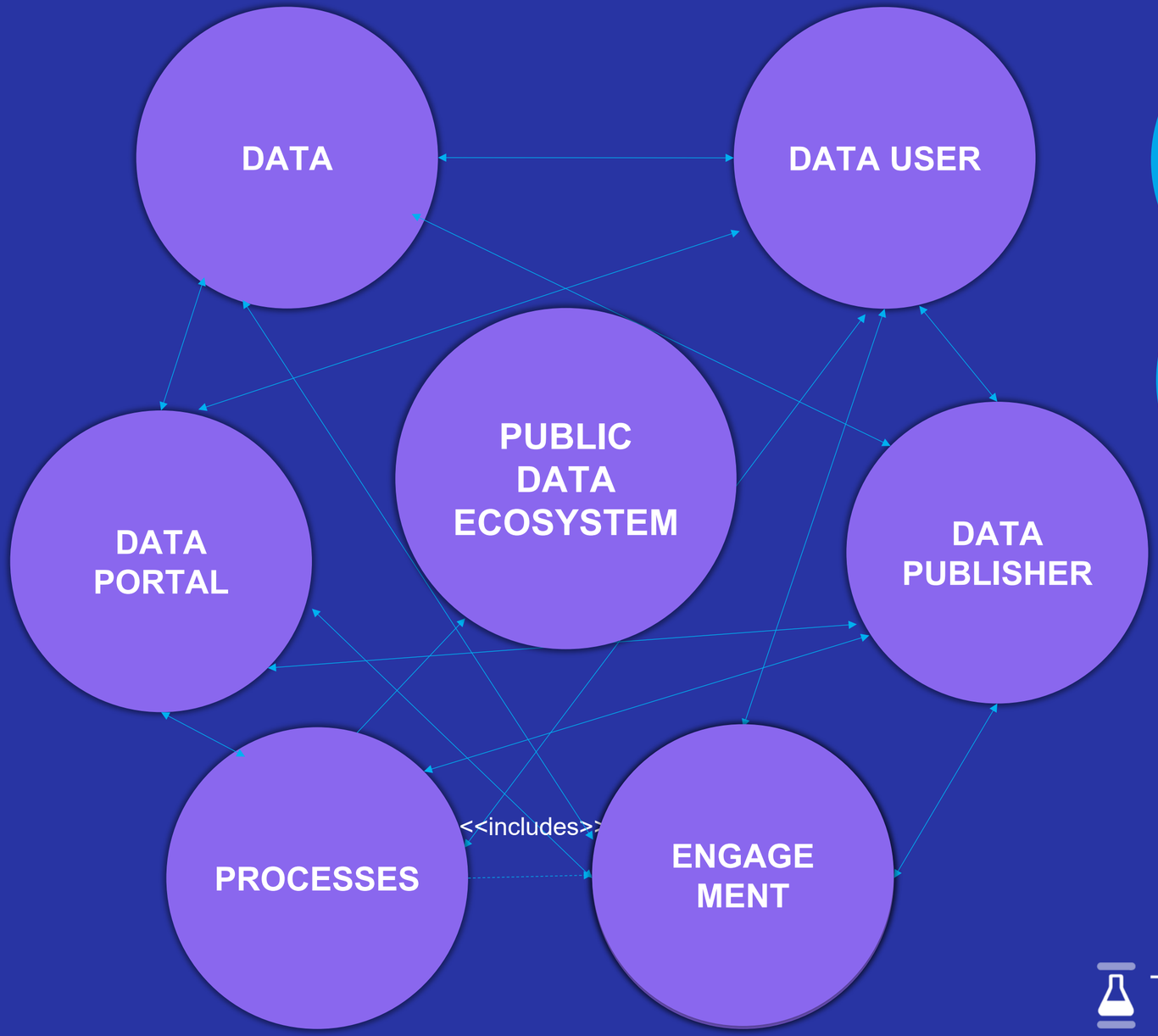
DEVELOP

AI as Ecosystem Value Developer

ENGAGE

AI as Ecosystem Engager

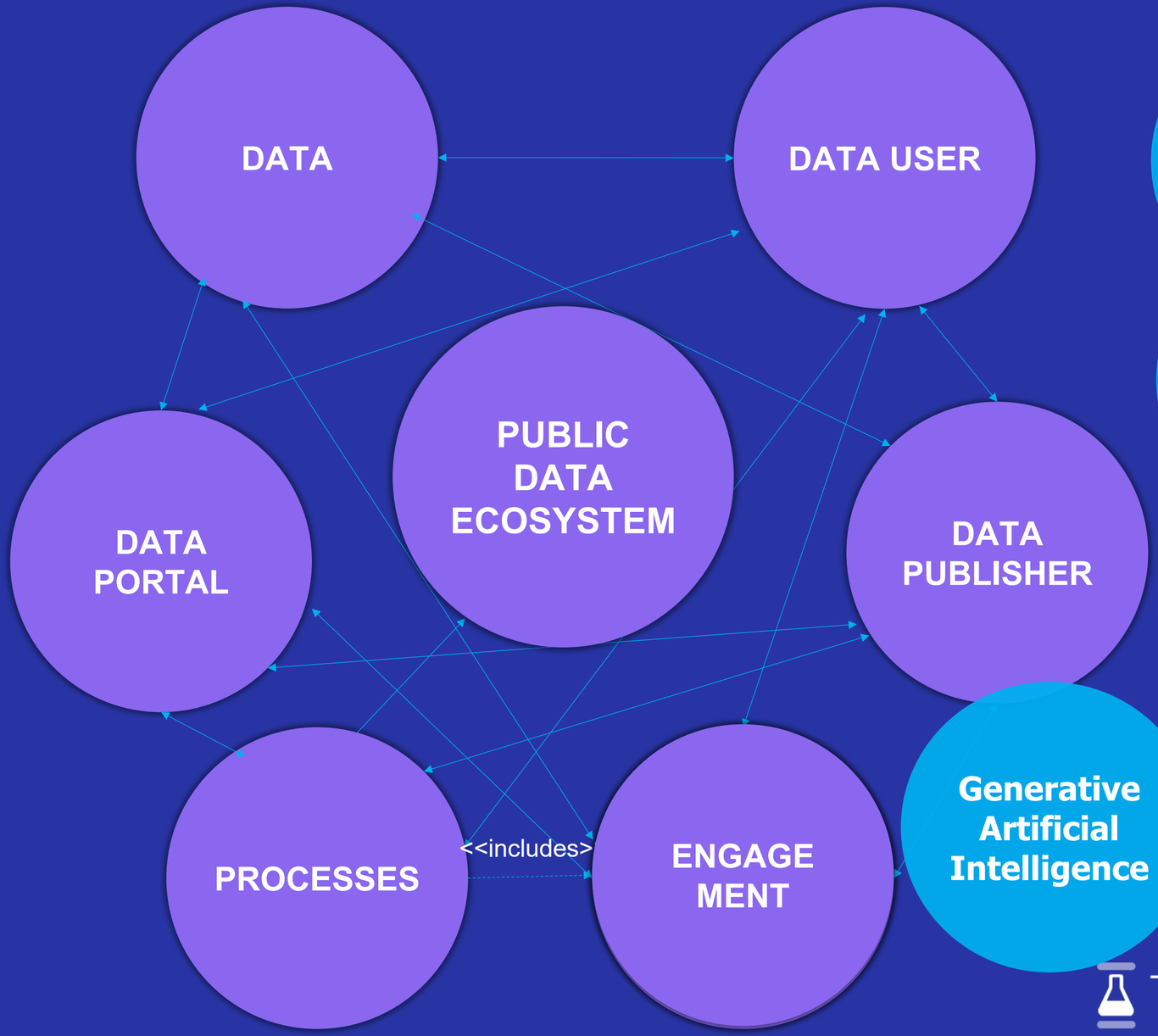
AND EVEN THIS IS OVERSIMPLIFIED IDEA



THEGOVLAB

Verhulst, S., Addo, P.M., Young, A., Zahuranec, A.J., Baumann, D. and McMurren, J., 2021. Emerging uses of technology for development: a new intelligence paradigm.

AND EVEN THIS IS OVERSIMPLIFIED IDEA

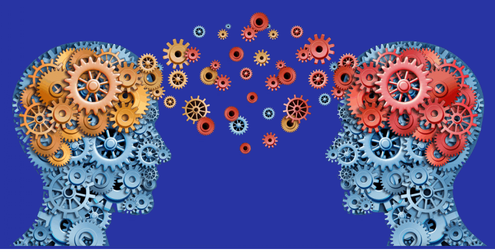


THEGOVLAB

Verhulst, S., Addo, P.M., Young, A., Zahuranec, A.J., Baumann, D. and McMurren, J., 2021. Emerging uses of technology for development: a new intelligence paradigm.

- **These forms of intelligence can help improve the decision-making capacity of development practitioners by enabling them to better understand or communicate relevant insights, when implemented effectively**






THEGOVLAB

Verhulst, S., Addo, P.M., Young, A., Zahuranec, A.J., Baumann, D. and McMurren, J., 2021. Emerging uses of technology for development: a new intelligence paradigm.

NOT TO FORGET ABOUT RISKS

OR WITH *GREAT* DATA COMES GREAT RESPONSIBILITY*

TECHNICAL

data availability & quality

AI-ready data

interoperability & scalability

contextualization

green AI & sustainability

AI governance

ETHICAL

bias & discrimination

privacy & data protection

lack of trust & explainability (FATE)

responsible AI

POLICY

data governance

AI governance & legal liability

absence of global frameworks

regulatory fragmentation

GENERAL SOCIETAL & ADOPTION RISKS

AI literacy & trust

local adaptation

**adaptation of Haruki Murakami "With great knowledge comes great responsibility"*

Nikiforova A., Zuiderwijk A. (2022) Barriers to openly sharing government data: towards an open data-adapted innovation resistance theory, In 15th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2022). Association for Computing Machinery, New York, NY, USA,

UNLOCK

ARTIFICIAL INTELLIGENCE

FOR

PUBLIC DATA EOCSYSTEM!

HOW TO UNLOCK THE SYMBIOTIC RELATIONSHIP OF ARTIFICIAL INTELLIGENCE, DATA INTELLIGENCE, AND COLLABORATIVE INTELLIGENCE FOR INNOVATIVE DECISION-MAKING AND PROBLEM SOLVING?

Image source: https://www.pngfind.com/mpng/TRmwoJi_to-be-continued-png-continued-back-to-the/

EU OPEN DATA DAYS

OPEN DATA RHAPSODY

AI WILL ROCK YOU!

19 – 20 MARCH 2025

DALL·E 3-generated content

EU OPEN DATA DAYS

19 – 20 MARCH 2025

THANK YOU!
AITÄH!
PALDIES!

Contact information:

https://anastasijanikiforova.com/

nikiforova.anastasija@gmail.com

https://www.linkedin.com/in/anastasija-nikiforova-466b99b3/

https://scholar.google.lv/citations?user=1Wj0-A8AAAAJ&hl=lv&oi=ao

https://medium.com/@nikiforova.anastasija