# 2021 SEE USER FORUM

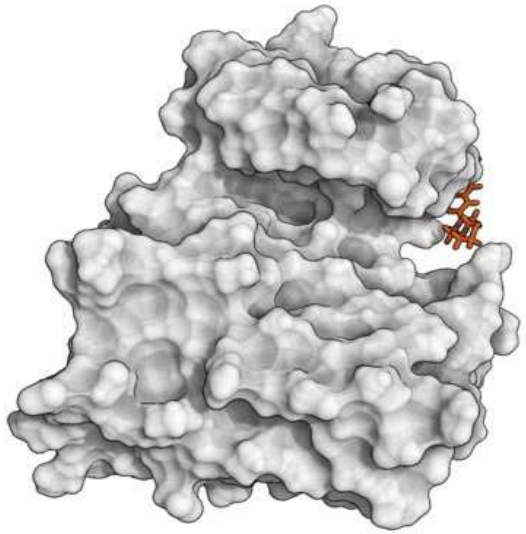
**3-4 November 2021**

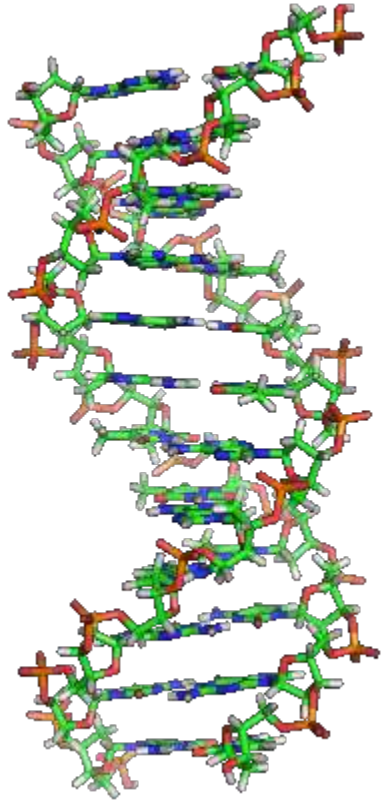
## The added value of EOSC for research

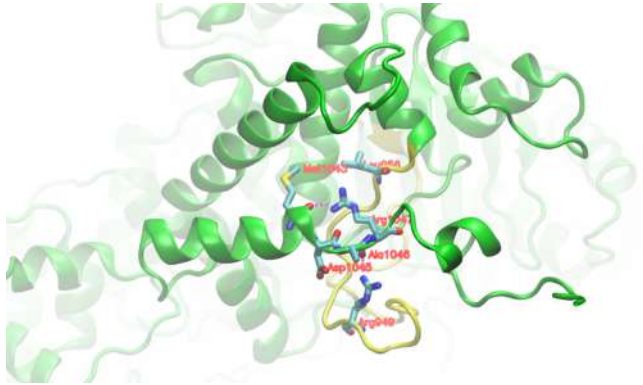
Dr. Zoe Cournia
Senior Researcher, Biomedical Research Foundation, Academy of Athens
Life Science Scientific Community Leader in EOSC project NI4OS-Europe

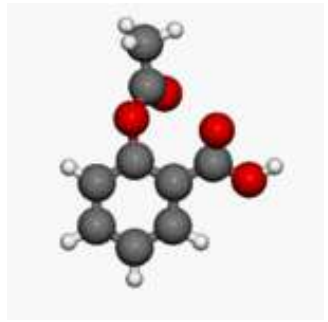
As part of the GÉANT 2020 Framework Partnership Agreement (FPA), the project receives funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 856726 (GN4-3).

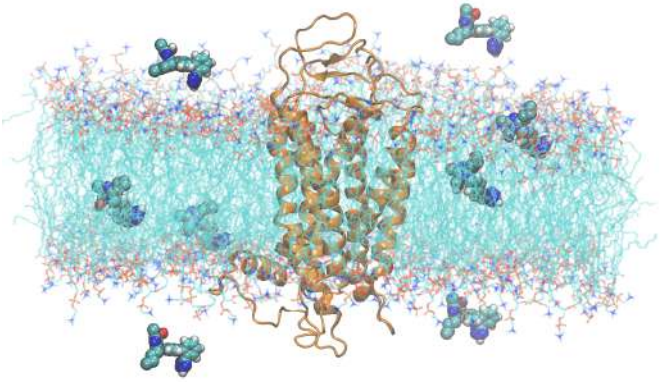
# Cournia Lab – Computational Drug Design



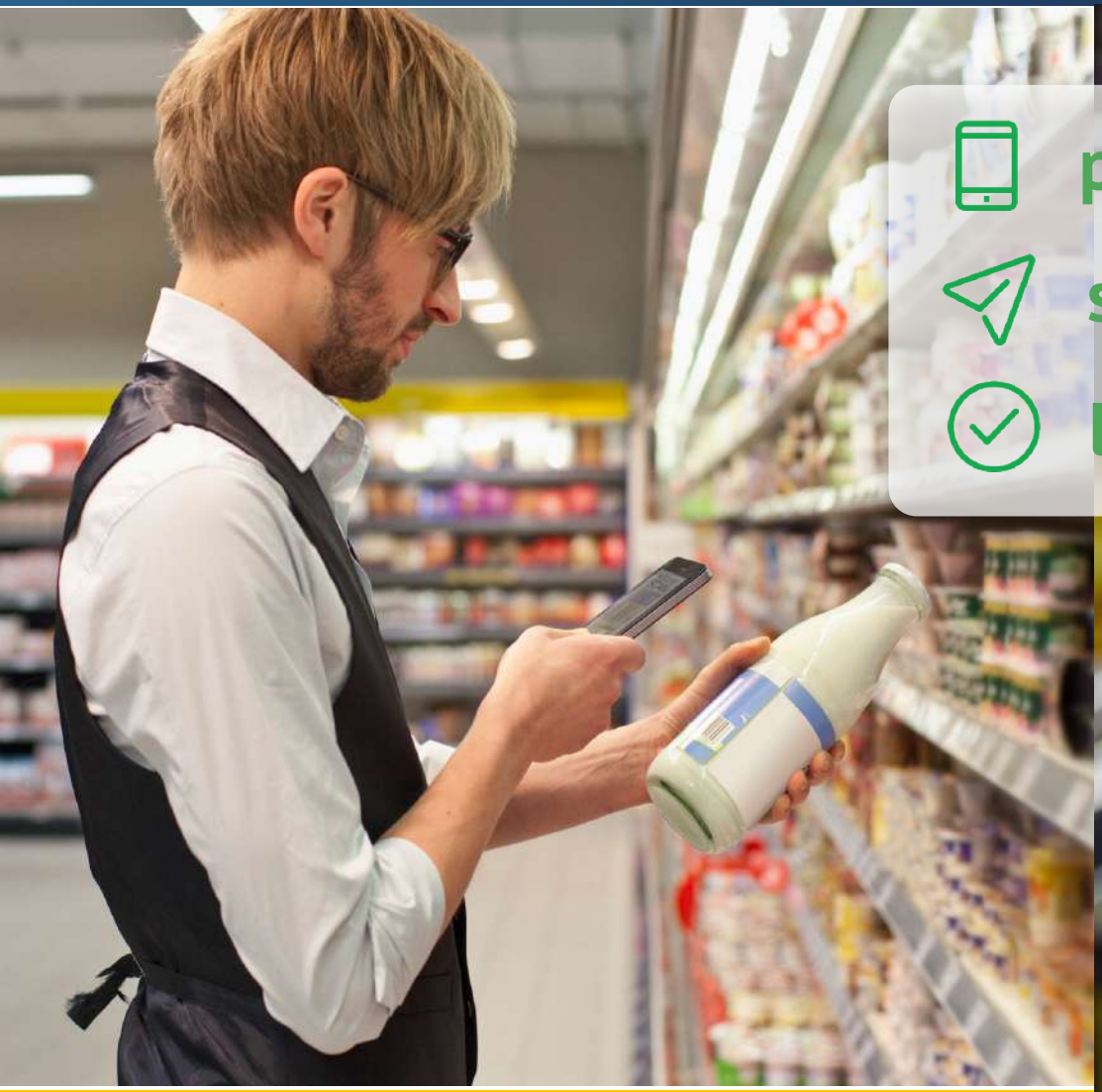








# Ingredio is a tool that makes ingredients easy to understand



photo

send

learn

# Open Data

- ✓ Open
- ✓ Findable
- ✓ Accessible
- ✓ Interoperable
- ✓ Re-usable
- ✓ Reliable

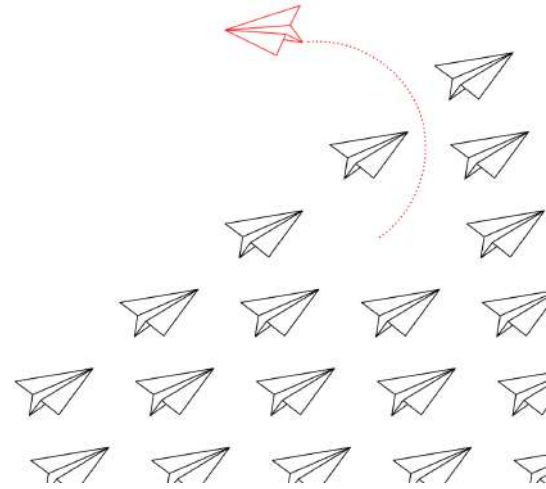
DATA

Key for

- ✓ **Driving Science**
- ✓ **Improving Quality of Life for European citizens**

# Current Challenges for Users

- Generate TB of data every month
- How to store, manage and analyze this data?
- What happens when I publish my data? Can I share datasets in an open repository so that other researchers have access?
- How can I access datasets from other colleagues who have published similar research?
- Current databases mostly fragmented and duplicating research



# Sharing Data

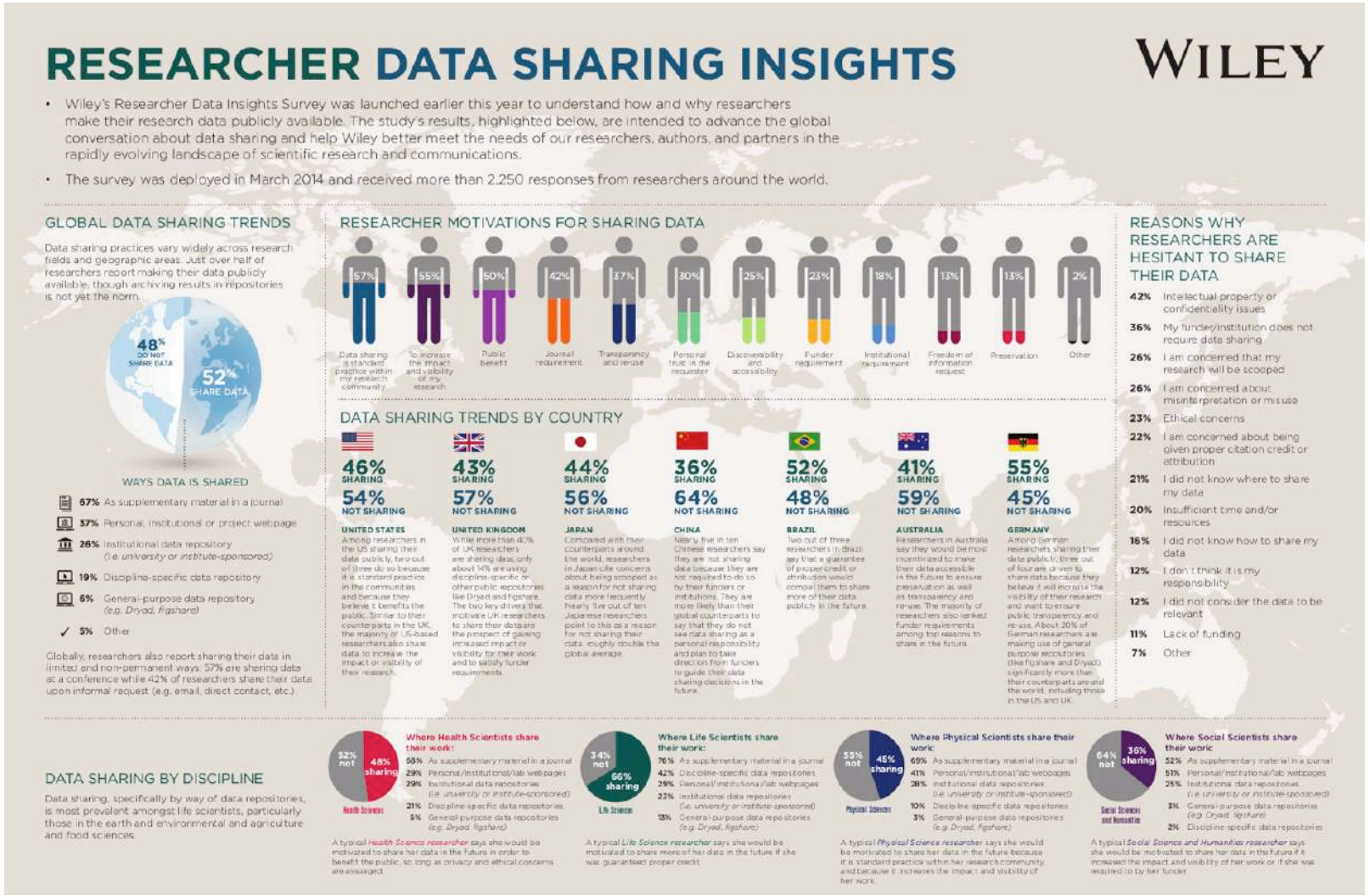
- Progress depends on previous work: New discoveries on old data
- Higher quality science because of reviewed data integrity
- Science is expensive
- Avoid duplication

(EU reportedly spends **€10.2 billion**/year for duplication of research)

- Giving back to community –

responsibility to communicate research results

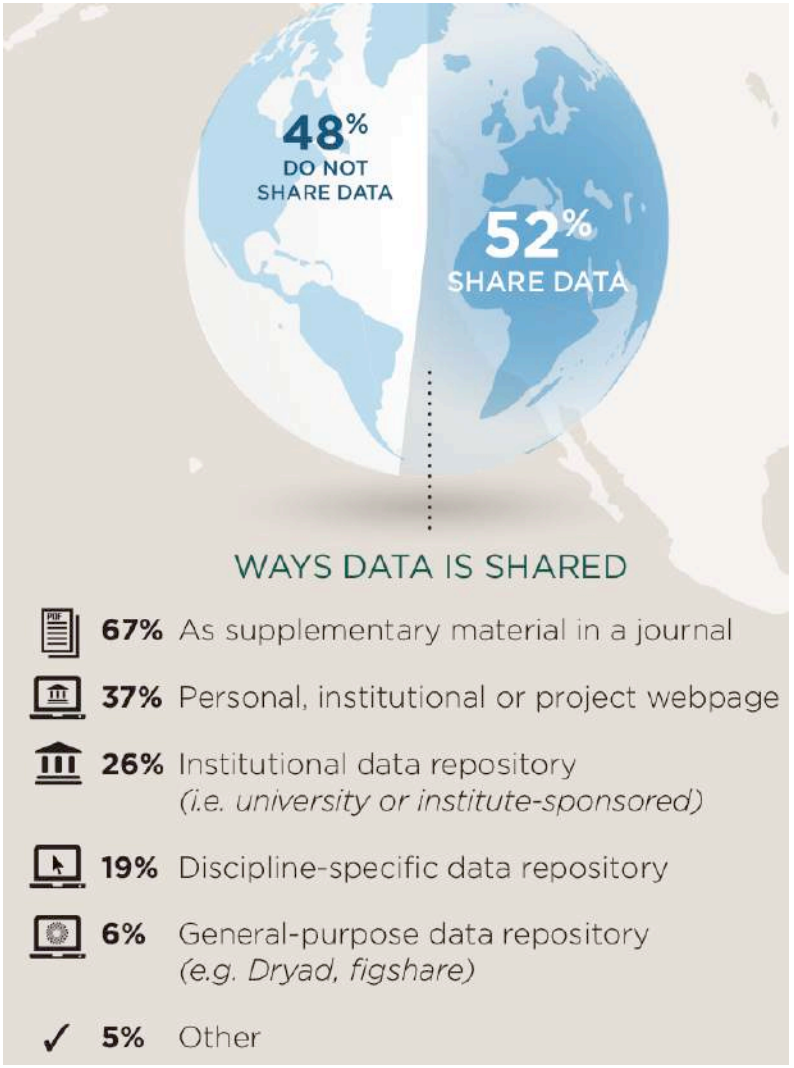
# Is data currently shared and how?






https://www.wiley.com/network/researchers/licensing-and-open-access/how-and-why-researchers-share-data-and-why-they-dont

# Is data currently shared and how?






https://www.wiley.com/network/researchers/licensing-and-open-access/how-and-why-researchers-share-data-and-why-they-dont

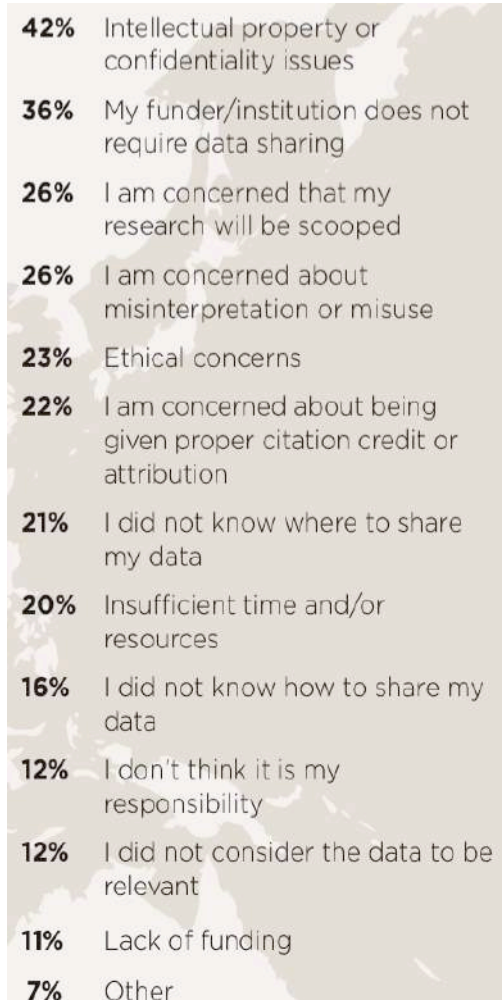
# Why is 52% of data not shared?

| | Responses | Percent |
|---|---|---|
| Insufficient Time | 603 | 53.6% |
| Lack of Funding | 445 | 39.6% |
| Do not Have Rights to Make Data Public | 271 | 24.1% |
| No Place to Put Data | 264 | 23.5% |
| Lack of Standards | 222 | 19.8% |
| Sponsor does not Require | 196 | 17.4% |
| Do not Need Data | 169 | 15.0% |
| Other Reasons For Data Not Available | 164 | 14.6% |
| Should not be Available | 162 | 14.4% |

doi:10.1371/journal.pone.0021101.t012

- ❑ journal policies are not up-to-date
- ❑ scientists are not up-to-date

⇒ no resources to look up **where** to put **what** data






https://www.wiley.com/network/researchers/licensing-and-open-access/how-and-why-researchers-share-data-and-why-they-dont, Slide Credit: Johanna Tiemann

# Findability and Discoverability

- But.... 52% are sharing their data
- Why can‘t I find it?

- no overview where data is deposited
- not all content is visible
- rich metadata (annotations) missing
- machine readable content often not available
- not required

# Interoperability and re-usability

- Now that I ve found it, can I (re)use it?

- Lack of **standardization** of information
- Interoperability and reproducibility require **agreeing on semantics**
- **Community** project
- Cost

- Ideal case: re-run analysis with deposited data & scripts (*reviewing*)

**MolSSI Workshop: Molecular Dynamics Software Interoperability** (New York, November 2019)

# How to remove those barriers?

| |
|---|
| Insufficient Time |
| Lack of Funding |
| Do not Have Rights to Make Data Public |
| No Place to Put Data |
| Lack of Standards |
| Sponsor does not Require |
| Do not Need Data |
| Other Reasons For Data Not Available |
| Should not be Available |

doi:10.1371/journal.pone.0021101.t012

- ❑ Specify & update guidelines & establish standards
- ❑ Increase awareness
- ❑ Publish all datasets (even negative)
- ❑ Look-up repositories for discoverability
- ❑ Standardize information
- ❑ Create Community
- ❑ Train

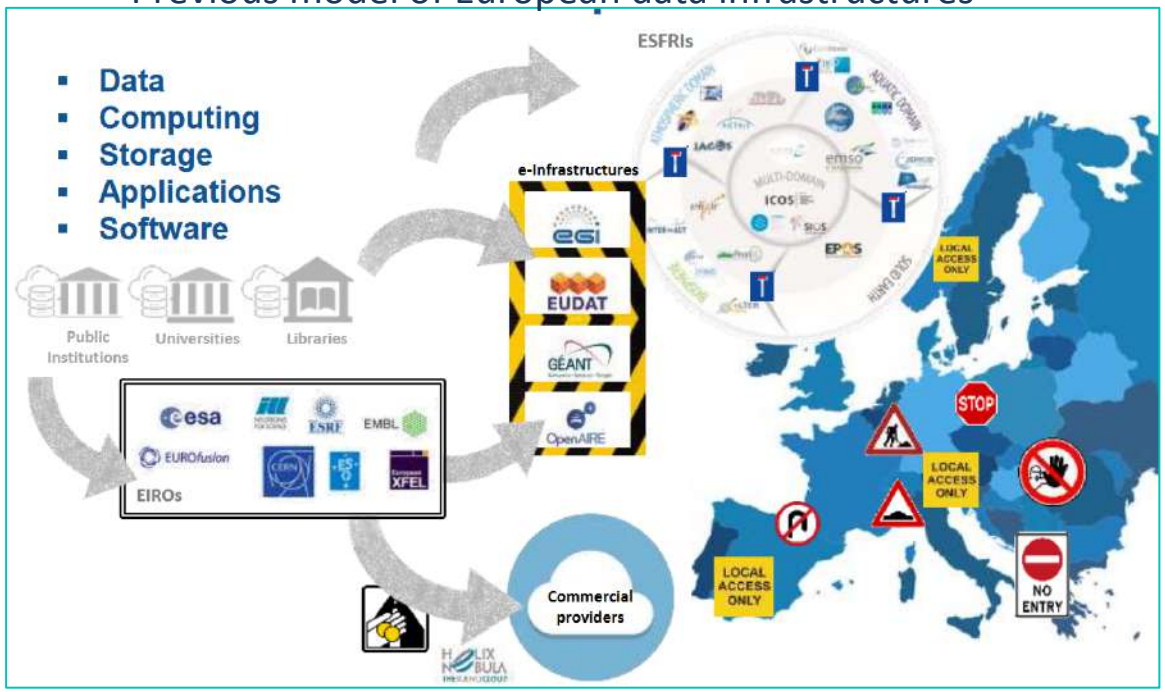
| | |
|---|---|
| 42% | Intellectual property or confidentiality issues |
| 36% | My funder/institution does not require data sharing |
| 26% | I am concerned that my research will be scooped |
| 26% | I am concerned about misinterpretation or misuse |
| 23% | Ethical concerns |
| 22% | I am concerned about being given proper citation credit or attribution |
| 21% | I did not know where to share my data |
| 20% | Insufficient time and/or resources |
| 16% | I did not know how to share my data |
| 12% | I don't think it is my responsibility |
| 12% | I did not consider the data to be relevant |
| 11% | Lack of funding |
| 7% | Other |

EUROPEAN OPEN SCIENCE CLOUD

**Realisation of EC vision for a common & trusted environment for Open Science**

# EOSC Vision

Previous model of European data infrastructures

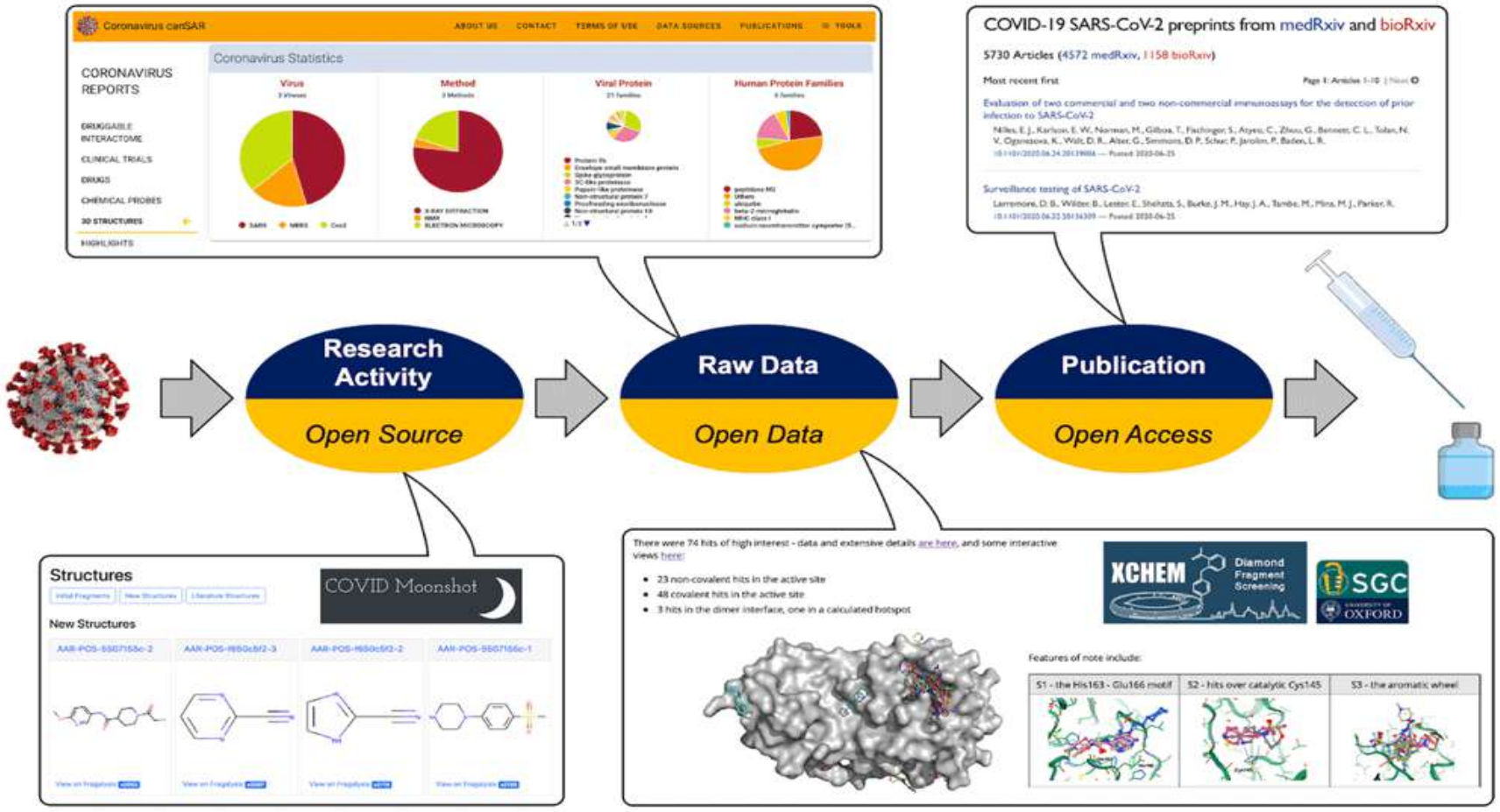



*Source: EOSC Strategic Implementation Roadmap 2018-2020, May 2018, European Commission*

*From fragmentation and uneven access to information to a federated model, where access to data would be universal, building on a strong legacy*


Future EOSC model: federation of data infrastructures

# Application of Open Science for COVID-19 treatment/ vaccine development

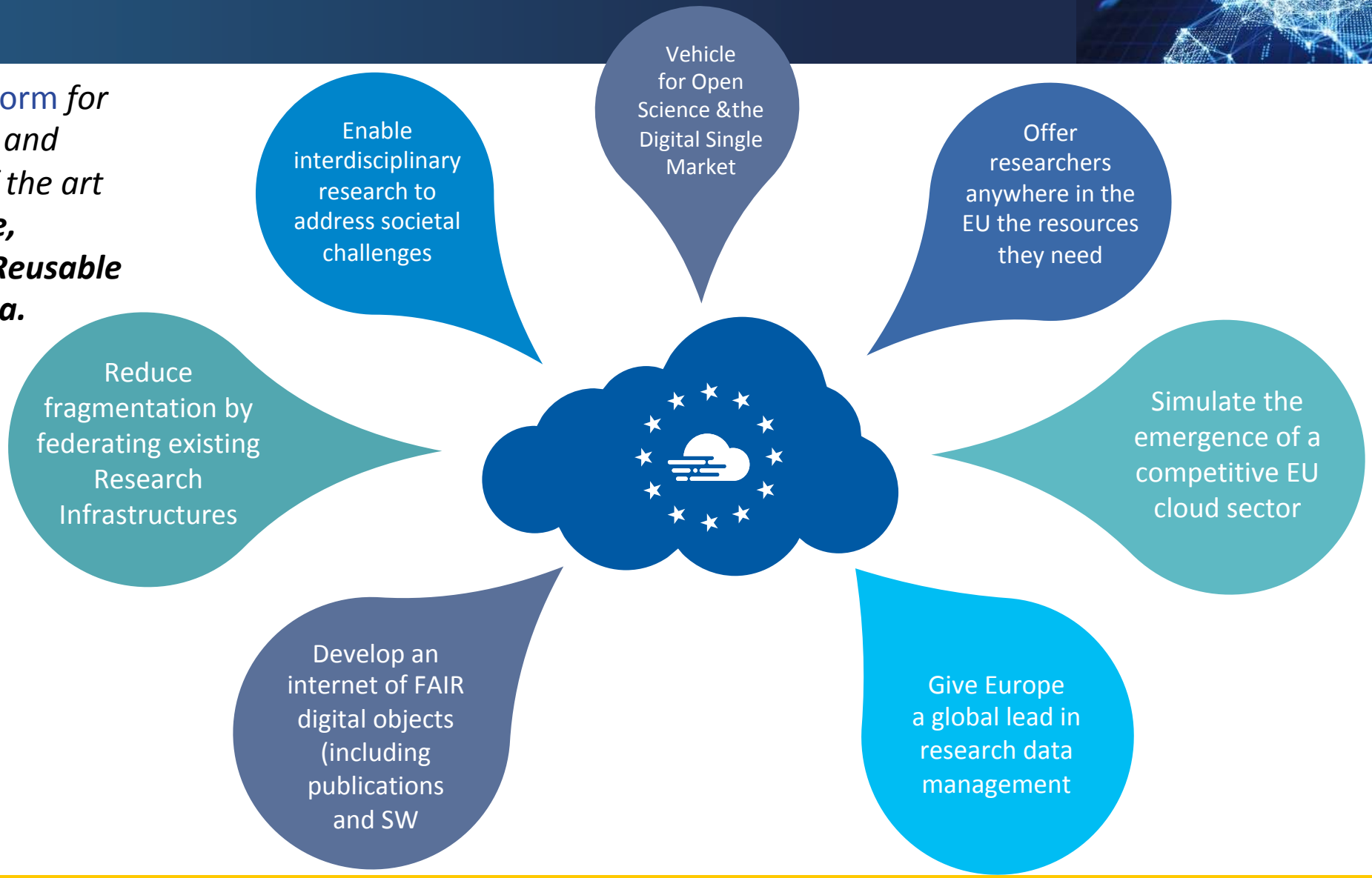





Tse EG, Klug DM and Todd MH. Open science approaches to COVID-19. F1000Research 2020, 9:1043 (doi: 10.12688/f1000research.26084.1)

# EOSC Vision

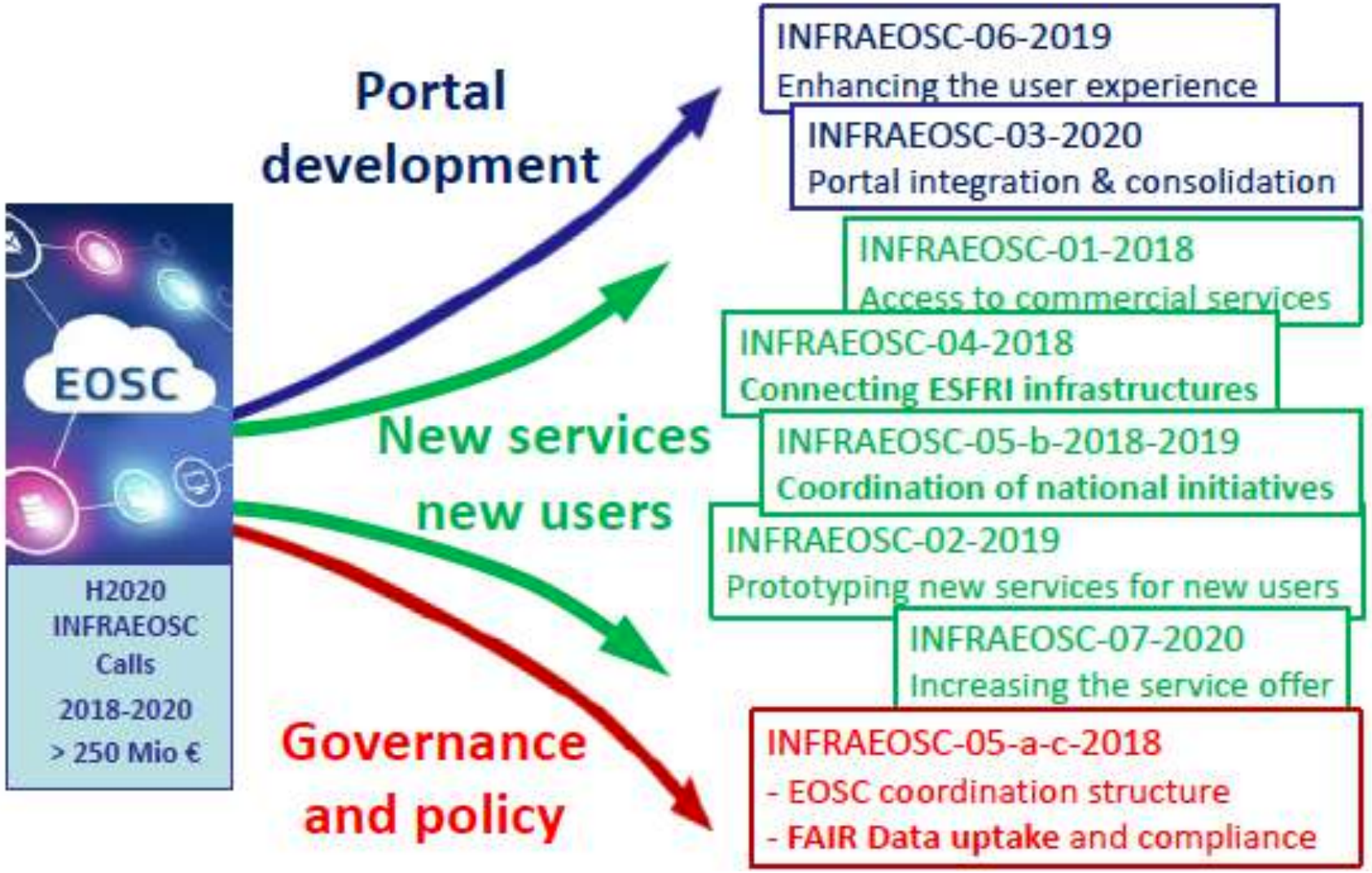
One federated platform *for producing, curating and distributing state of the art* ***Findable, Accessible, Interoperable and Reusable (FAIR) scientific data.***






https://ec.europa.eu/info/sites/default/files/research_and_innovation/funding/documents/ec_rtd_he-partnership-open-science-cloud-eosc.pdf

# H2020 dedicated activities to prototype EOSC




Source: EOSC Webinar

# Current EOSC Projects

Triple

SSHOC

Reliance

eXtreme DataCloud

OCRE Open Clouds for Research Environments

DICE Data Infrastructure Capacity for EOSC

EOSC ENHANCE Connecting Thematic Communities to Advance Open Science

EOSC NORDIC

NI4OS

OpenAIRE

CINECA

panosc photon and neutron open science cloud

INODE

Blue-Cloud Piloting innovative services for Marine Research & the Blue Economy

PRIMAGE Medical imaging Artificial intelligence Childhood cancer research

PROCESS

RDA RESEARCH DATA ALLIANCE EUROPE

EOSCpilot The European Open Science Cloud for Research Pilot Project

ICEDIG.EU

EOSC-Pillar Coordination and Harmonisation of National &Thematic Initiatives to support EOSC

ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures

EOSC FUTURE

ExPaNDS European Open Science Cloud Photon and Neutron Data Services

FAIR4Health

FAIRplus

FNS-Cloud Food Nutrition Security

e-IRG SP6

FAIRSFAIR Fostering Fair Data Practices in Europe

FREYA

GÉANT Networks • Services • People

HELIX NEBULA THESCIENCECLOUD

Digital Health Europe

EGI-ACE

eInfra Central

EOSC-hub

OpenRiskNet

EOSC SYNERGY

EOSCsecretariat.eu Setup and management of the EOSC Secretariat supporting the EOSC Governance

EOSC-Life

ENVRI FAIR

# The role of the INFRAEOSC 05 projects

**EOSC-Pillar** *Coordination and Harmonisation of National Inititiatives, Infrastructures and Data services in Central and Western Europe*

**EOSC-synergy** *European Open Science Cloud - Expanding Capacities by building Capabilities*

**NI4OS-Europe** *National Initiatives for Open Science in Europe*

**EOSC-Nordic** *European Open Science Cloud in the Nordic and Baltic Countries*

## Common objectives

- *Mapping and integration of national and local services to enhance their discoverability/usage*
- *Enhancement of the EOSC with new thematic services*
- *Support of trans-national services integration the EOSC through appropriate technical and policy measures (including legal and business model aspects).*
- *Cross-nation and Region coordination and policy harmonisation.*
- *Scientific community engagement and skill development*
- *FAIR and ORDM promotion*
- *Supporting the Governance*

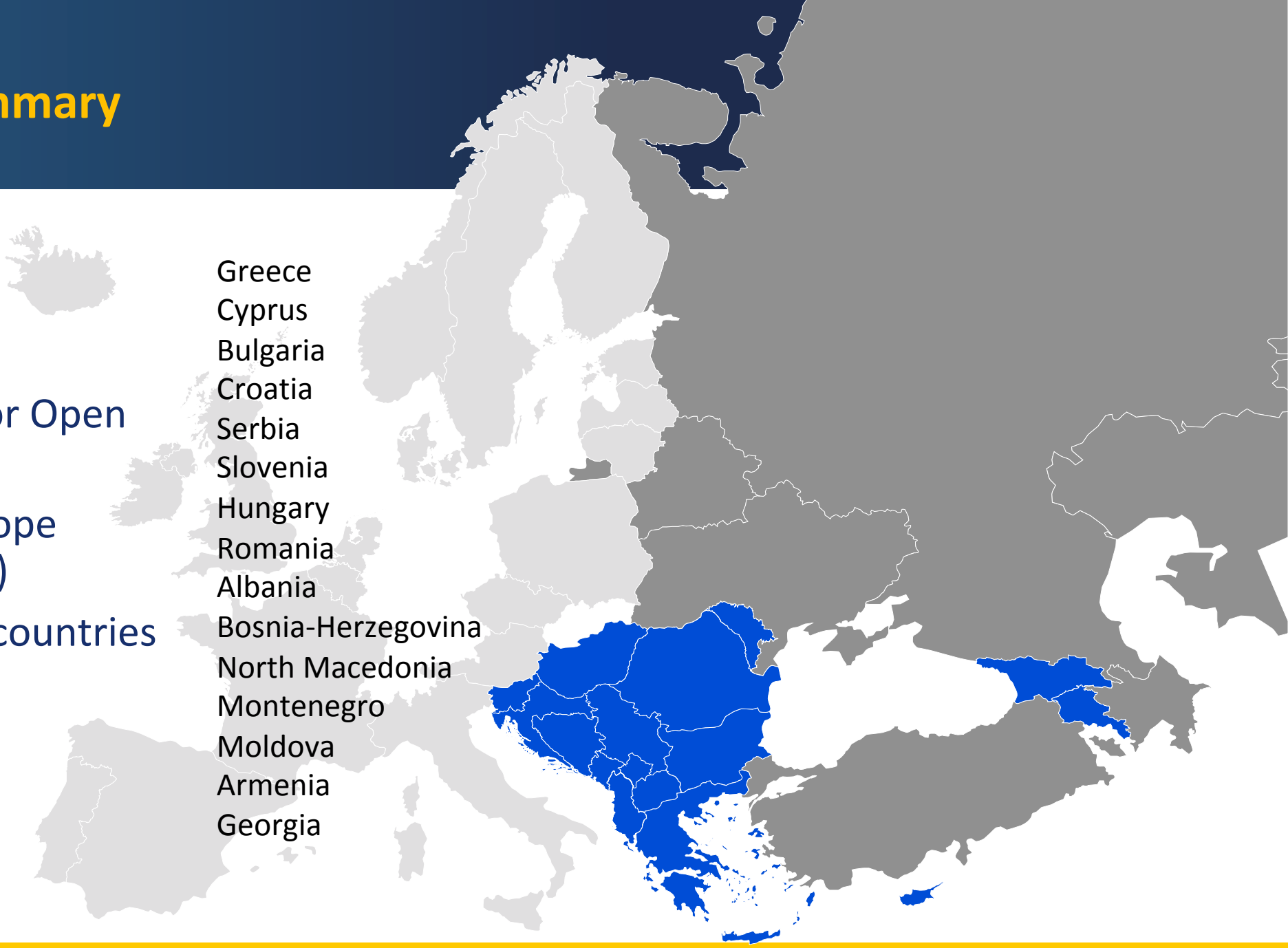


Source: EOSC Webinar

# NI4OS-Europe Summary

- National Initiatives for Open Science in Europe
- Acronym: NI4OS-Europe (pronounced “NIFOS”)
- 22 Partners from 15 countries

Greece
Cyprus
Bulgaria
Croatia
Serbia
Slovenia
Hungary
Romania
Albania
Bosnia-Herzegovina
North Macedonia
Montenegro
Moldova
Armenia
Georgia



# EOSC/NI4OS-Europe: Objectives

**Support** the development and inclusion of the national Open Science Cloud (OSC) initiatives in 15 Member States and Associated Countries in the overall scheme of EOSC governance

**Spread** the EOSC and FAIR principles, ORDM principles, interoperability, federated services, in the community and train it

**Provide** technical and policy support in on-boarding of the existing and future service providers into EOSC

# Use Case Example: Personalized Medicine

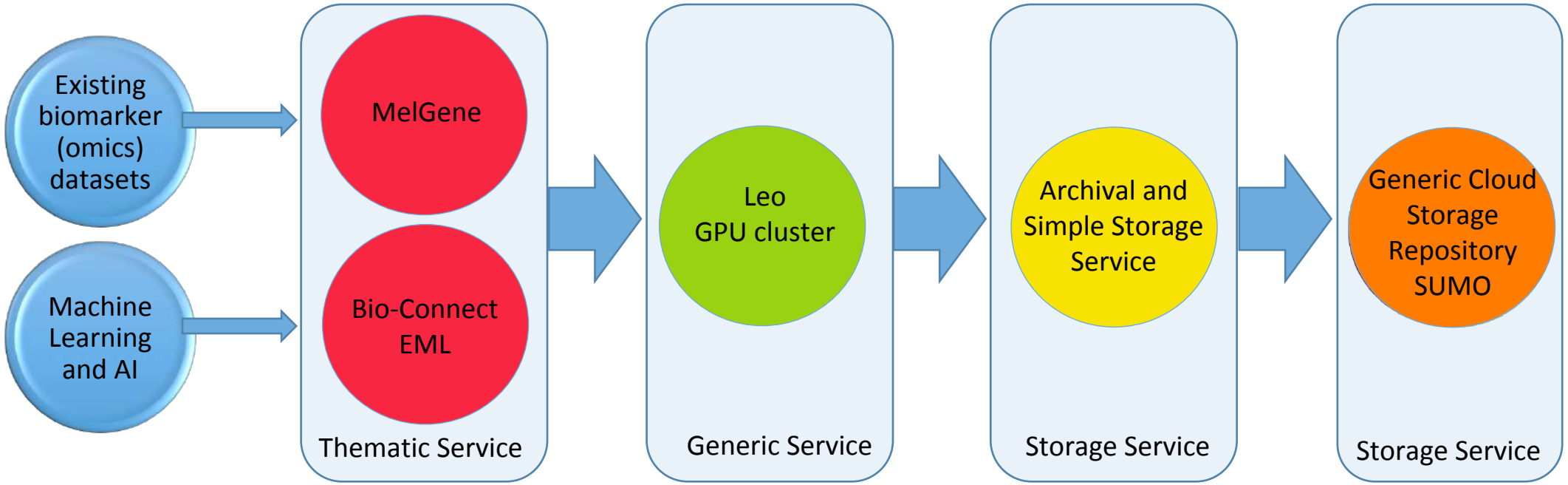
- ✓ The Right Drug
- ✓ To the Right Patient
- ✓ For the Right Disease
- ✓ At the Right Time
- ✓ With the Right Dosage

Genetic & metabolic data
allow drugs to be tailored
to specific patient groups

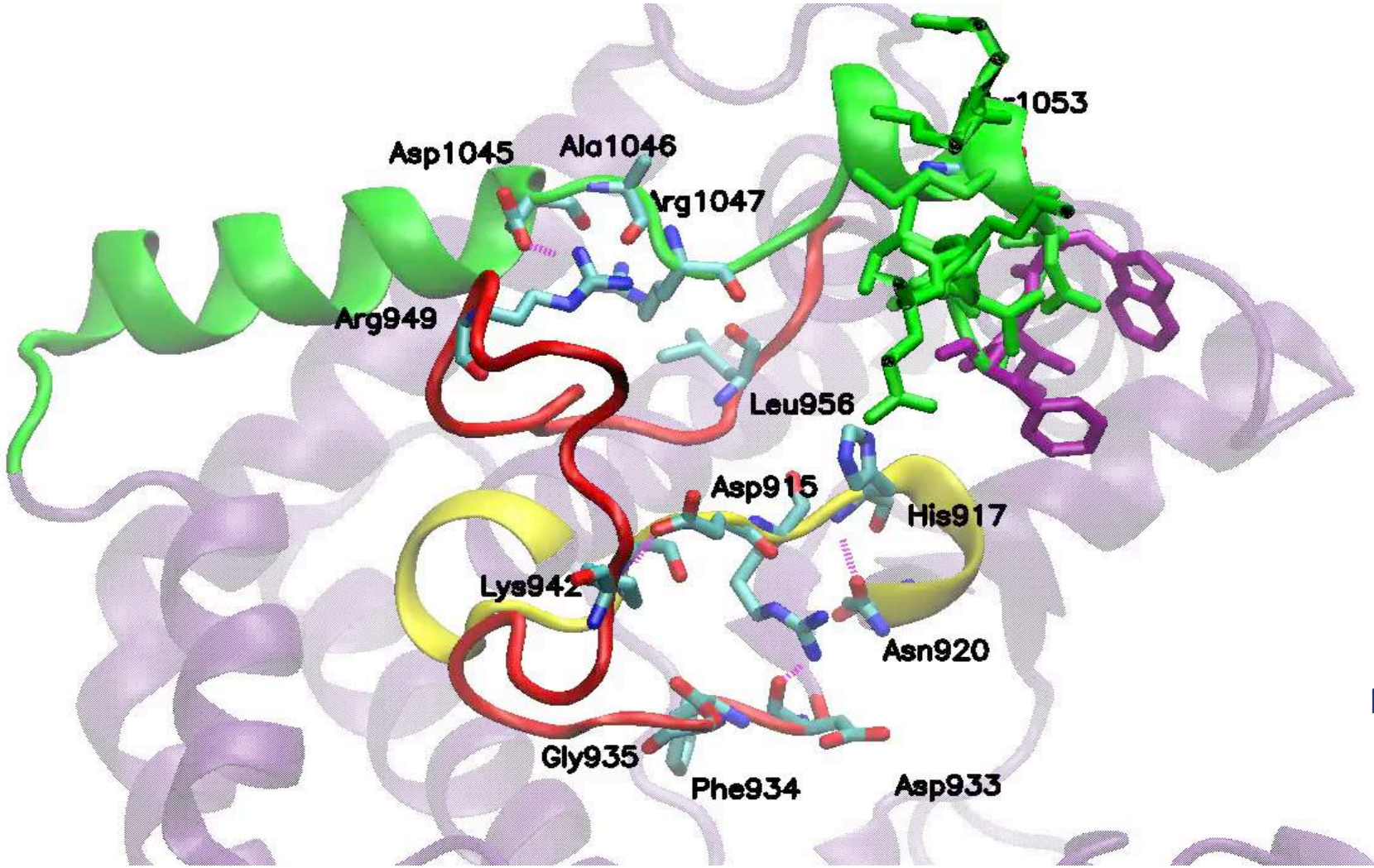
# Use Case Example: Extracting correlations for patient stratification using machine learning

# Extracting correlations for patient stratification using machine learning

- Stratifying patients for melanoma cancer type based on existing, open data



# Molecular Dynamics simulations reveal protein dynamics




Deposited in
NI4OS-Europe resources
Zenodo
Human Brain Project

# Benefits for researchers using EOSC resources

- Store, manage, analyze TB of produced data
- Share datasets to the community upon publication
- Access datasets from colleagues without duplicating data
- More efficient dataset searching based on metadata
- Onboard web-based thematic services in one common, unified portal
- Help with evaluation and certification of datasets (e.g. FAIRsFAIR)
- Training on best practices for Open Science & FAIR principles
- Enable Open Innovation seamlessly
- Opportunities to collaborate with Pharma Industry opened up for my lab / Networking

# Our in-house drug discovery tools: Onboarded on EOSC

NI4OS Europe **EXPLORE** NI4OS-Europe Catalogue

Thematic services

**ChemBioServer** is a publicly available web application for effectively **filtering and clustering chemical compounds** used in drug discovery.

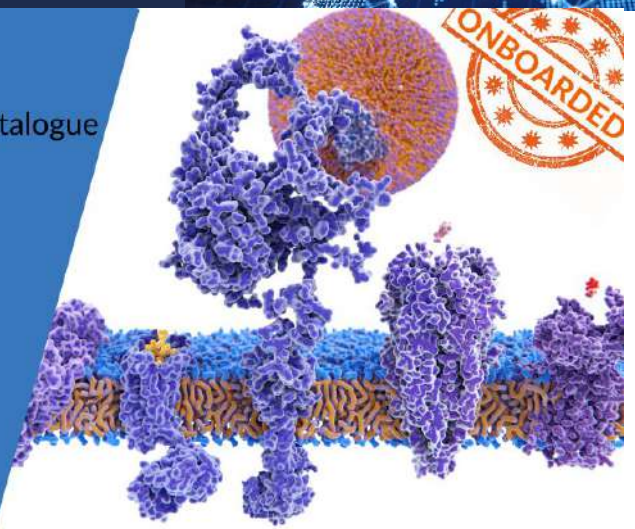
http://chembioserver.vi-seem.eu/


NI4OS Europe **EXPLORE** NI4OS-Europe Catalogue

Thematic services

**DREAMM** is a novel web-based tool that **predicts the protein-membrane interfaces of peripheral membrane proteins** using ensemble machine learning.

http://dreamm-ni4os.eu/

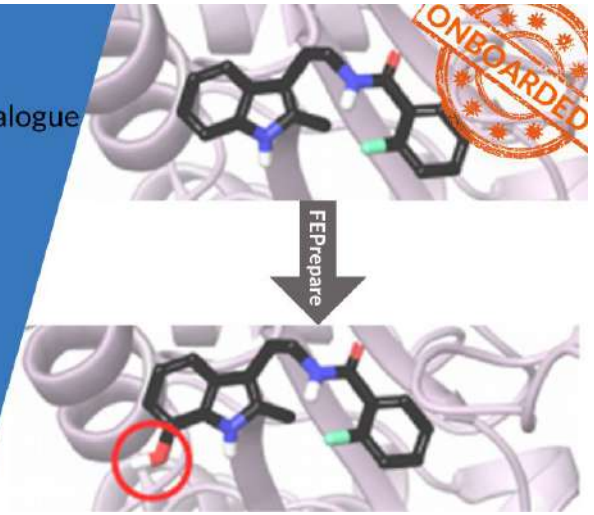



NI4OS Europe **EXPLORE** NI4OS-Europe Catalogue

Thematic services

**FEPrepare** is a webserver, which automates the set-up procedure for performing **NAMD/FEP simulations.**

http://feprepare.vi-seem.eu

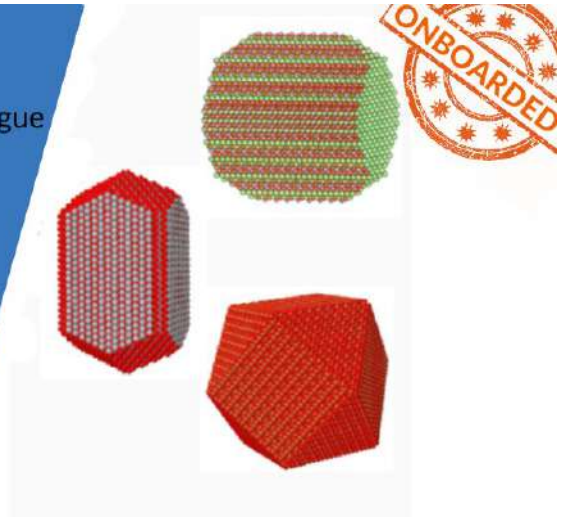



NI4OS Europe **EXPLORE** NI4OS-Europe Catalogue

Thematic services

**NanoCrystal** is a novel web-based **crystallographic tool** for the construction of nanoparticles from any material crystal structure.

http://nanocrystal-seem.eu



# Benefits for service providers using EOSC resources

- Publish, share and advertise services & resources
- Get statistics about access & feedback
- Free online platform to manage service requests
- Interact with users more efficiently & understand needs
- Get support for user authentication
- Open Service to a wider base
- Boosts visibility and discoverability
- Increase computational power of your lab / company with reliable core services

# EOSC works with companies, too

Ingredio

OpenAIRE | ADVANCE

# Enhancing the food & cosmetics OpenAIRE Research Graph for consumer health

# Ingredio – OpenAIRE Advance Collaboration

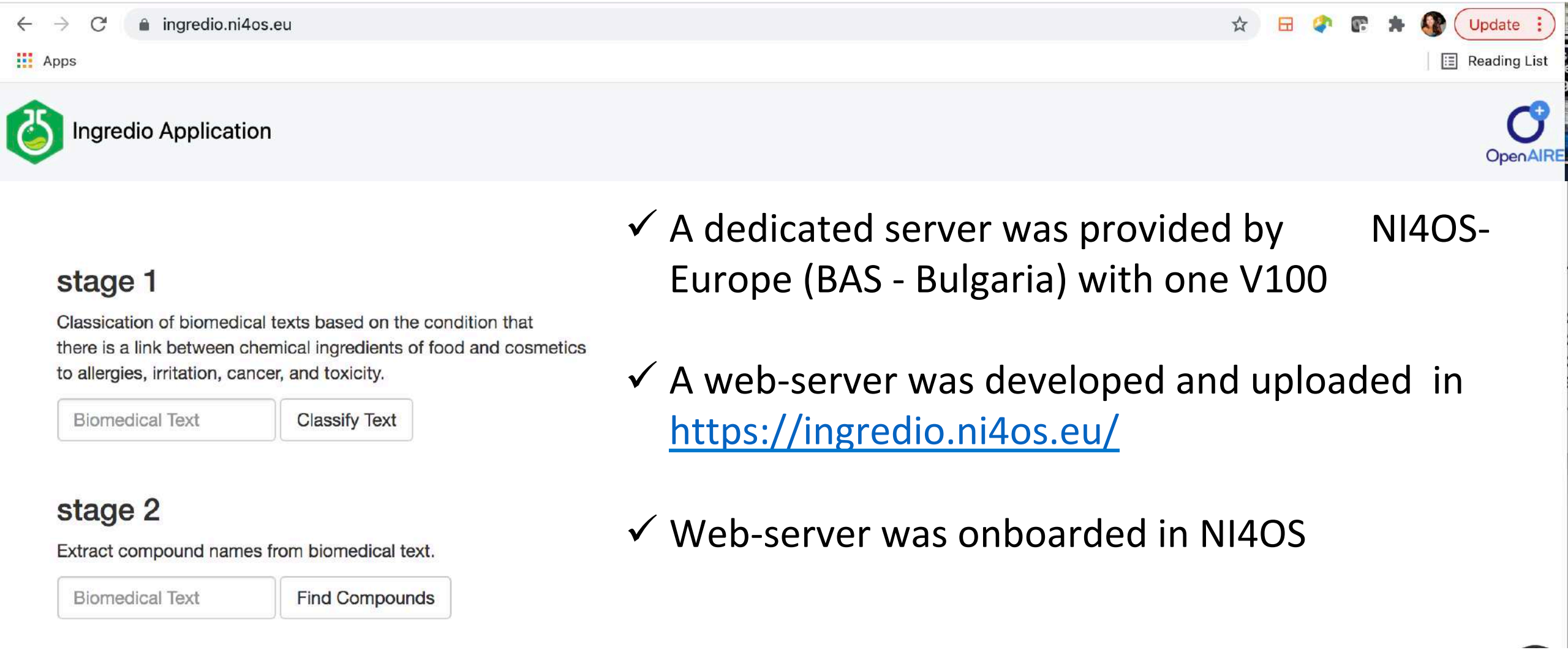
**Goal 1**
Develop text mining and Machine Learning algorithms to extract OpenAIRE data linking chemical ingredients of food &cosmetics to potential health hazards

**Goal 2**
Support the curation of OpenAIRE data with appropriate metadata schemas for efficient integration of information in the OpenAIRE Research Graph

**Goal 3**
Identifying new chemical ingredients from OpenAIRE data to enrich the OpenAIRE research graph & the Ingredio database

**Goal 4**
Identify causal relationships of chemical ingredients with hazards using the collected information

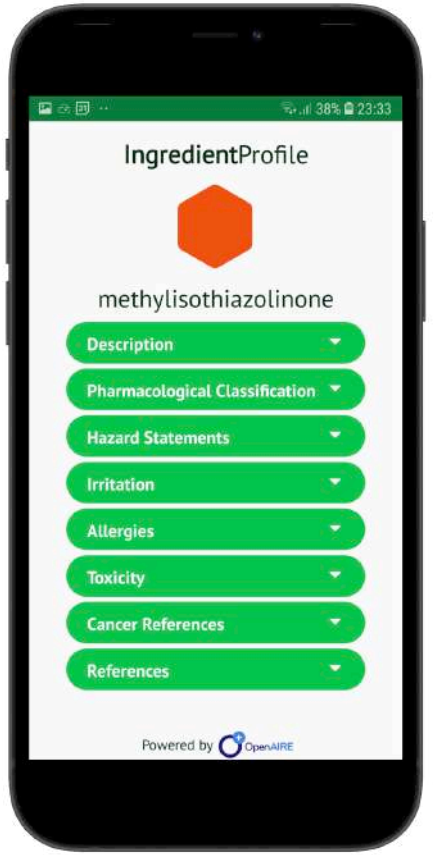
# Onboarding Platform to EOSC (NI4OS-Europe)




- ✓ A dedicated server was provided by NI4OS-Europe (BAS - Bulgaria) with one V100
- ✓ A web-server was developed and uploaded in https://ingredio.ni4os.eu/
- ✓ Web-server was onboarded in NI4OS

# Benefits for SMEs / Community

- Code was provided to OpenAIRE and uploaded to Github
- Beta testing of final products was largely completed
- The service was integrated to EOSC and was onboarded
- The Ingredio database was significantly enriched, which aids our SME competitiveness
- European citizens can now be better informed about potential hazards in food and cosmetics ingredients
- Open science promotes citizen trust in science



# Acknowledgements


**Team Ingredio**

M Kounadis
A Chatzigoulas
D Papakonstantinou
D Trovas

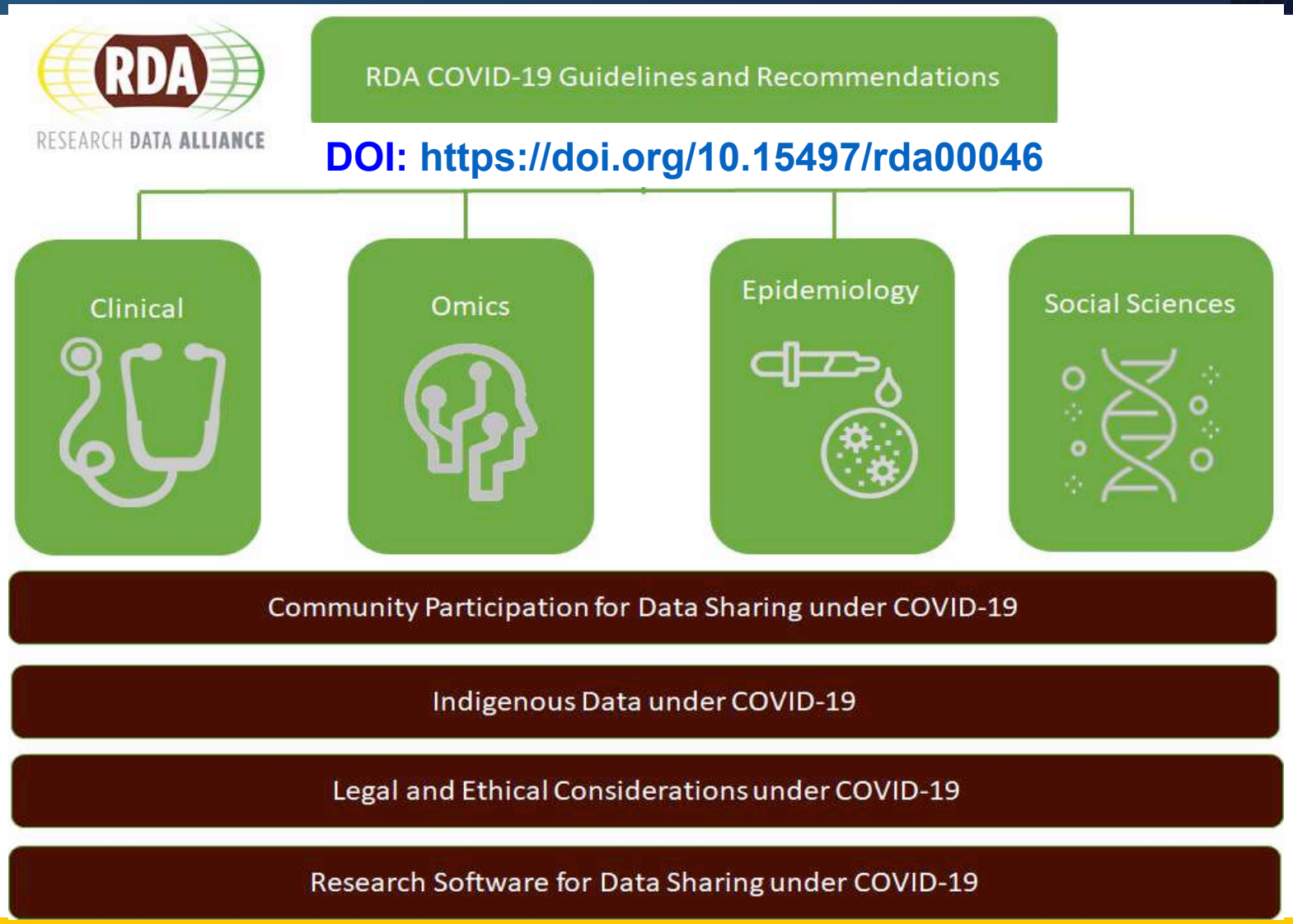
**Team OpenAIRE**

H Dimitropoulos
M Horst
Y Foufoulas
A Manocci
A Bardi
N Berikou

**Team NI4OS-Europe**

D Vudragovic
A Mishev
S Spasov
E Atanassov
M Durchova
K Koumantaros

# Service: Giving back to the community

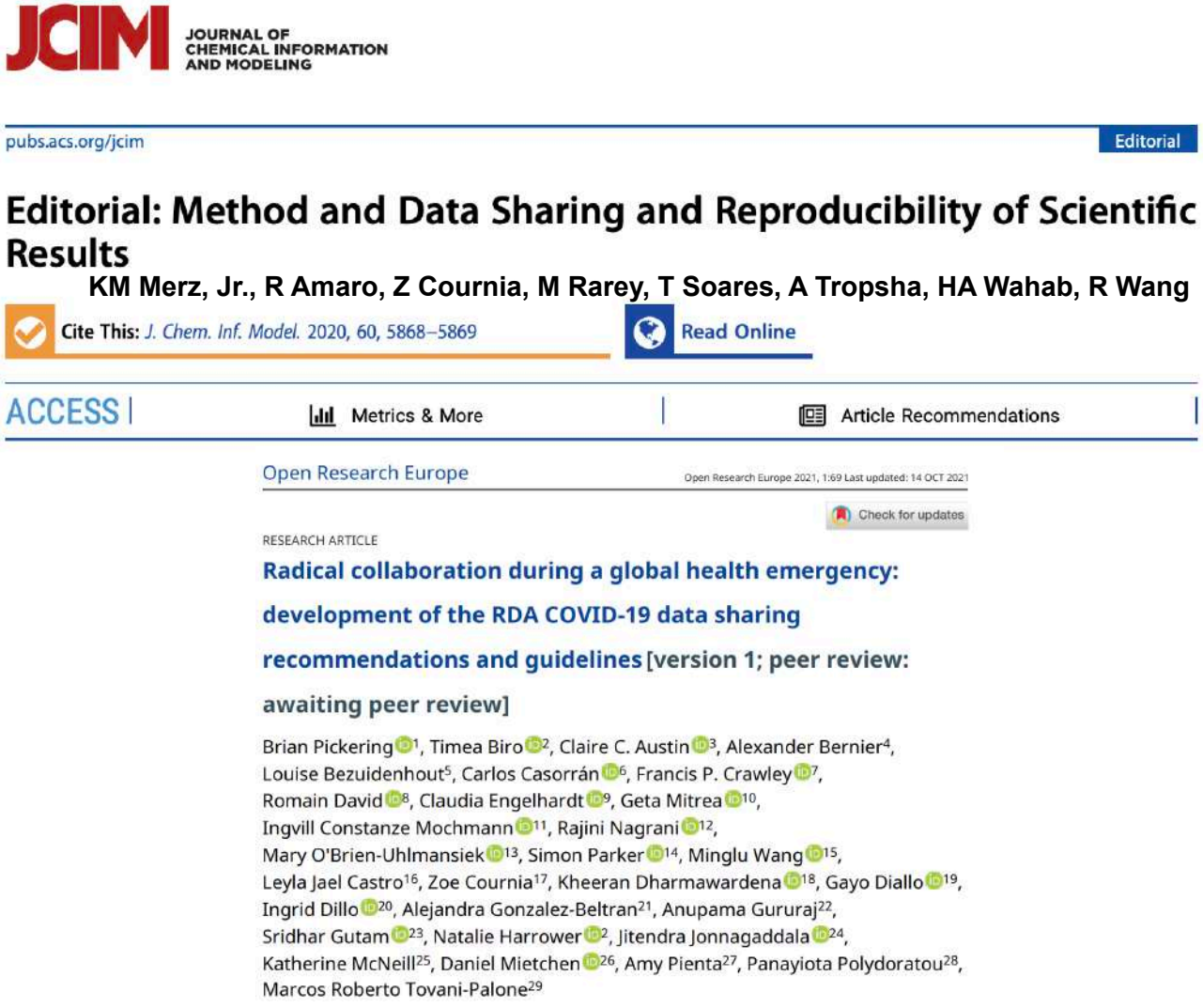


# Drafted Guidelines and recommendations for RDA COVID-19 WG

- Participated in the Research Data Alliance Working Group for COVID-2019 as a representative of NI4OS-Europe
- Joined the RDA COVID19-omics subgroup
- Over 10 meetings between March and May 2020 to:
- Clearly define detailed **guidelines on data sharing under the present COVID-19 circumstances** to help stakeholders follow best practices to maximize the efficiency of their work, and to act as a blueprint for future emergencies;
- Develop **recommendations for policymakers** to maximise timely, quality data sharing and appropriate responses in such health emergencies;
- Address the interests of **researchers, policy makers, funders, publishers**, and providers of **data sharing infrastructures**.


DOI: https://doi.org/10.15497/rda00046

# Service: Giving back to the community



JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Editorial

**Editorial: Method and Data Sharing and Reproducibility of Scientific Results**

KM Merz, Jr., R Amaro, Z Cournia, M Rarey, T Soares, A Tropsha, HA Wahab, R Wang

Cite This: *J. Chem. Inf. Model.* 2020, 60, 5868−5869 Read Online

ACCESS | Metrics & More | Article Recommendations

Open Research Europe Open Research Europe 2021, 1:69 Last updated: 14 OCT 2021

Check for updates

RESEARCH ARTICLE

**Radical collaboration during a global health emergency: development of the RDA COVID-19 data sharing recommendations and guidelines [version 1; peer review: awaiting peer review]**

Brian Pickering[1], Timea Biro[2], Claire C. Austin[3], Alexander Bernier[4], Louise Bezuidenhout[5], Carlos Casorrán[6], Francis P. Crawley[7], Romain David[8], Claudia Engelhardt[9], Geta Mitrea[10], Ingvill Constanze Mochmann[11], Rajini Nagrani[12], Mary O'Brien-Uhlmansiek[13], Simon Parker[14], Minglu Wang[15], Leyla Jael Castro[16], Zoe Cournia[17], Kheeran Dharmawardena[18], Gayo Diallo[19], Ingrid Dillo[20], Alejandra Gonzalez-Beltran[21], Anupama Gururaj[22], Sridhar Gutam[23], Natalie Harrower[2], Jitendra Jonnagaddala[24], Katherine McNeill[25], Daniel Mietchen[26], Amy Pienta[27], Panayiota Polydoratou[28], Marcos Roberto Tovani-Palone[29]

- Regular presentations and article writing for promoting open science
- Member of EOSC promoter group in Southeast Europe
- Member of NI4OS Task force for Metadata and Semantics

OPEN LETTER

REVISED **Fostering global data sharing: highlighting the recommendations of the Research Data Alliance COVID-19 working group [version 2; peer review: 2 approved, 1 approved with reservations]**

Claire C. Austin[1], Alexander Bernier[2], Louise Bezuidenhout[3], Juan Bicarregui[4], Timea Biro[5], Anne Cambon-Thomsen[6], Stephanie Russo Carroll[7], Zoe Cournia[8], Piotr Wojciech Dabrowski[9], Gayo Diallo[10], Thomas Duflot[11], Leyla Garcia[12], Sandra Gesing[13], Alejandra Gonzalez-Beltran[4], Anupama Gururaj[14], Natalie Harrower[5], Dawei Lin[14], Claudia Medeiros[15], Eva Méndez[16], Natalie Meyers[17], Daniel Mietchen[18], Rajini Nagrani[19], Gustav Nilsonne[20], Simon Parker[21], Brian Pickering[22], Amy Pienta[23], Panayiota Polydoratou[24], Fotis Psomopoulos[25], Stephanie Rennes[26], Robyn Rowe[27], Susanna-Assunta Sansone[28], Hugh Shanahan[29], Lina Sitz[30], Joanne Stocks[31], Marcos Roberto Tovani-Palone[32,33], Mary Uhlmansiek[34], Research Data Alliance

# Benefits for the Community

**Allow**
New generation of scholars to find, combine, analyse data & discoveries

**Accelerate**
Transition to Open Science and Open Innovation

**Transform**
Research faster and more cost-efficiently to products that benefit citizens

**Bring**
Science & Research closer to Societal needs

# Vision

Global Open Science as a driver for enabling a new paradigm of transparent, data-driven science as well as accelerating innovation

# Thank you

Any questions?

As part of the GÉANT 2020 Framework Partnership Agreement (FPA), the project receives funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 856726 (GN4-3).