Leiden University Medical Center

RD Connect

BBMRI Biobanking and Biomolecular Resources Research Infrastructure

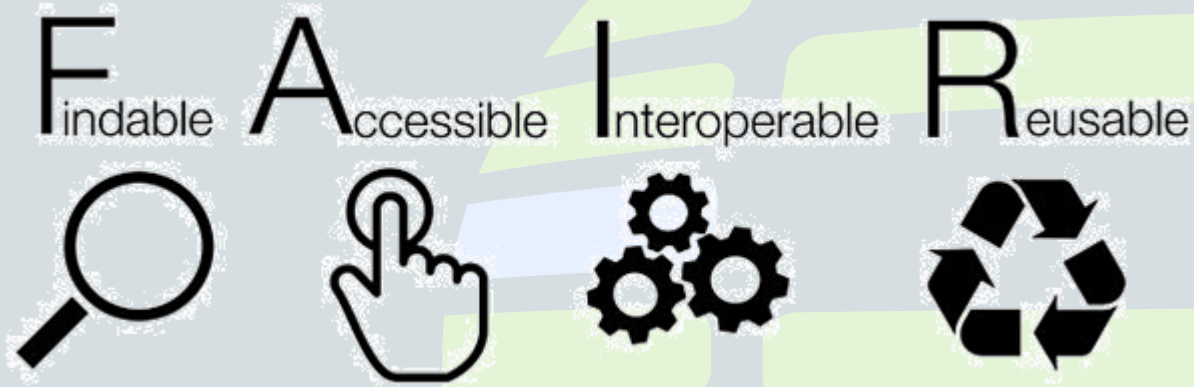
GO FAIR

EUDAT

DTL DUTCH TECHCENTRE FOR LIFE SCIENCES

elixir

excelerate



VASCERN

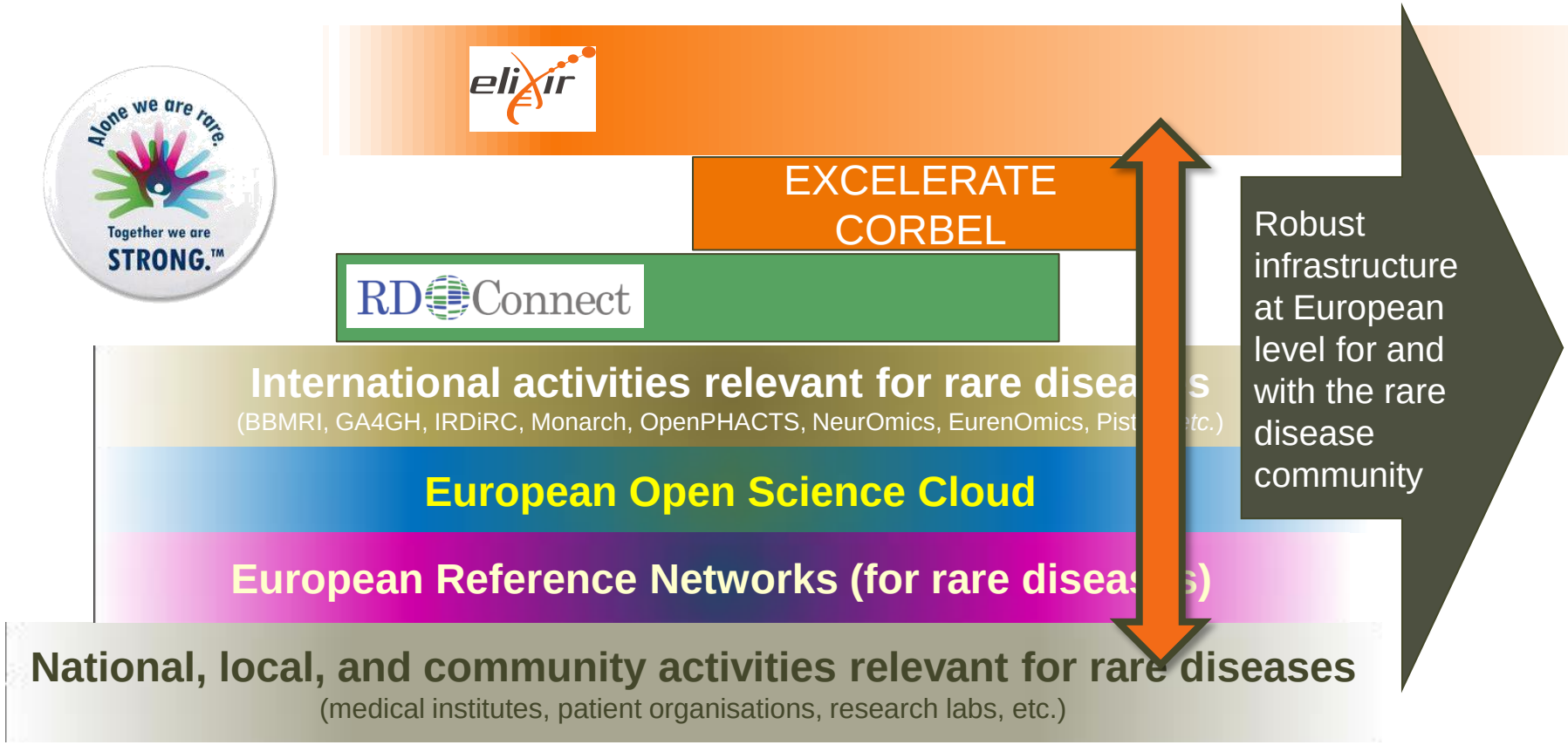
IRDiRC

SEVENTH FRAMEWORK PROGRAMME

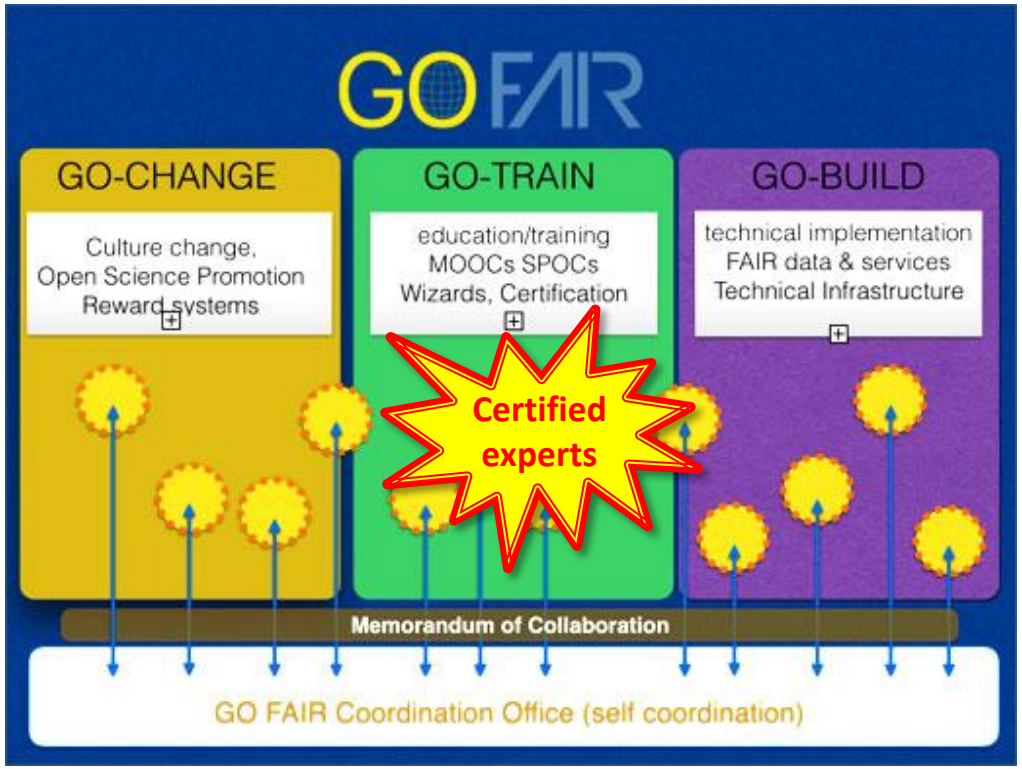
elixir

Marco Roos, Paris, October 13-14, 2017

Collaborating infrastructures

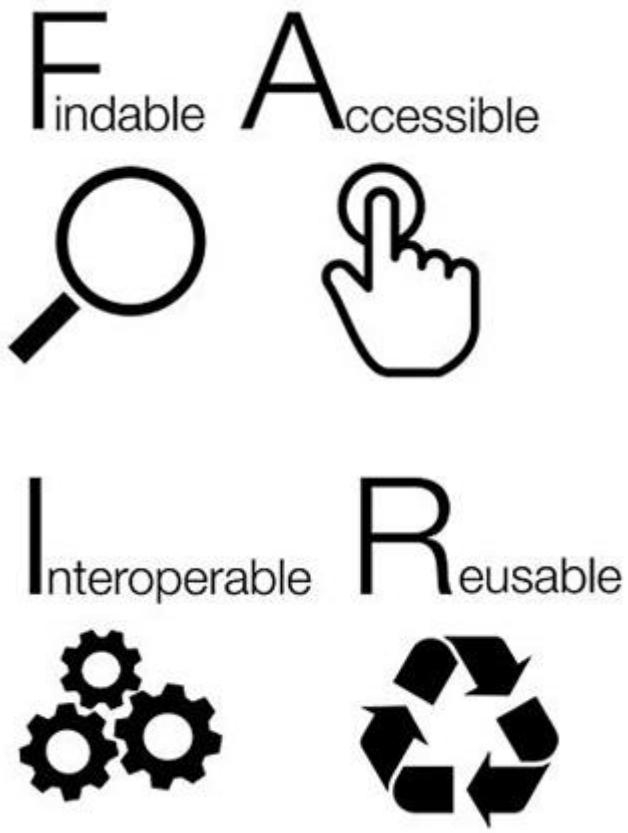


GO FAIR network for ERNs and rare diseases



- Self-organizing network
- Implements the European Open Science Cloud
- GO FAIR office supports coordination
- ELIXIR & others help building

FAIR vision and purpose

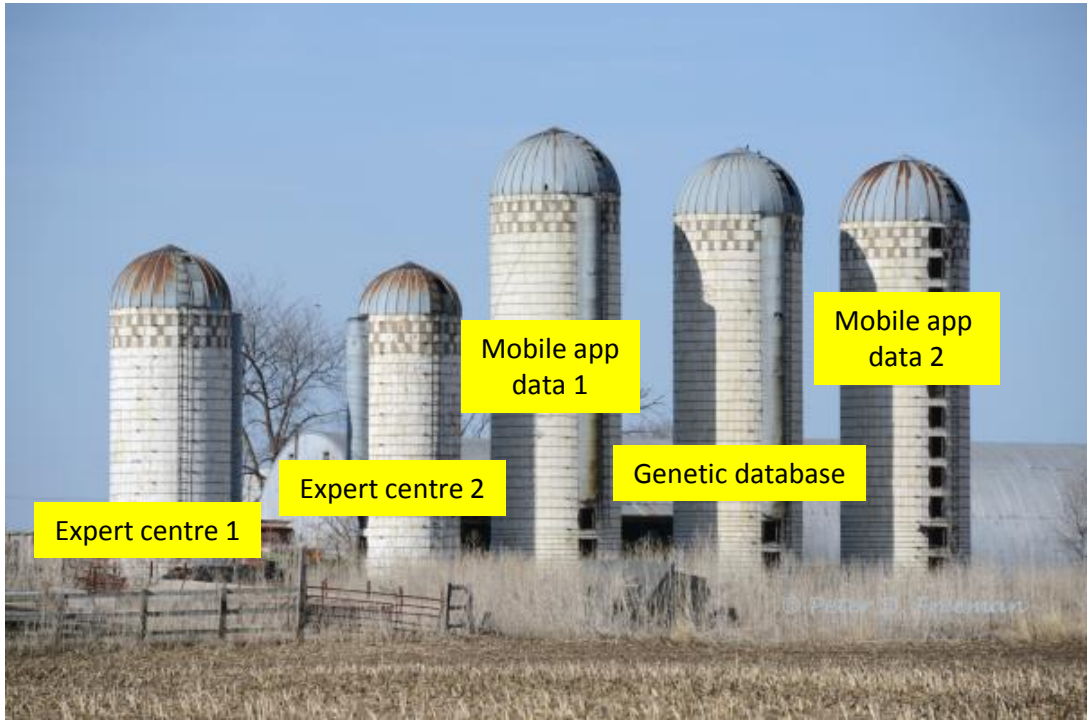


Vision: VASCERN stakeholders use

- appropriate data
- from *across* data sources in VASCERN
- and relevant data outside of VASCERN
- for basic queries and advanced analysis
- without bottlenecks & errors caused by data incompatibilities and opaque access

Who is the expert on your data?

The problem: information siloes



Data and samples lost for diagnosis, treatment, research

Inefficient (re)use of data, many errors

Data/samples unnecessarily recreated, less time for smart use of data

Personal Health Train concept

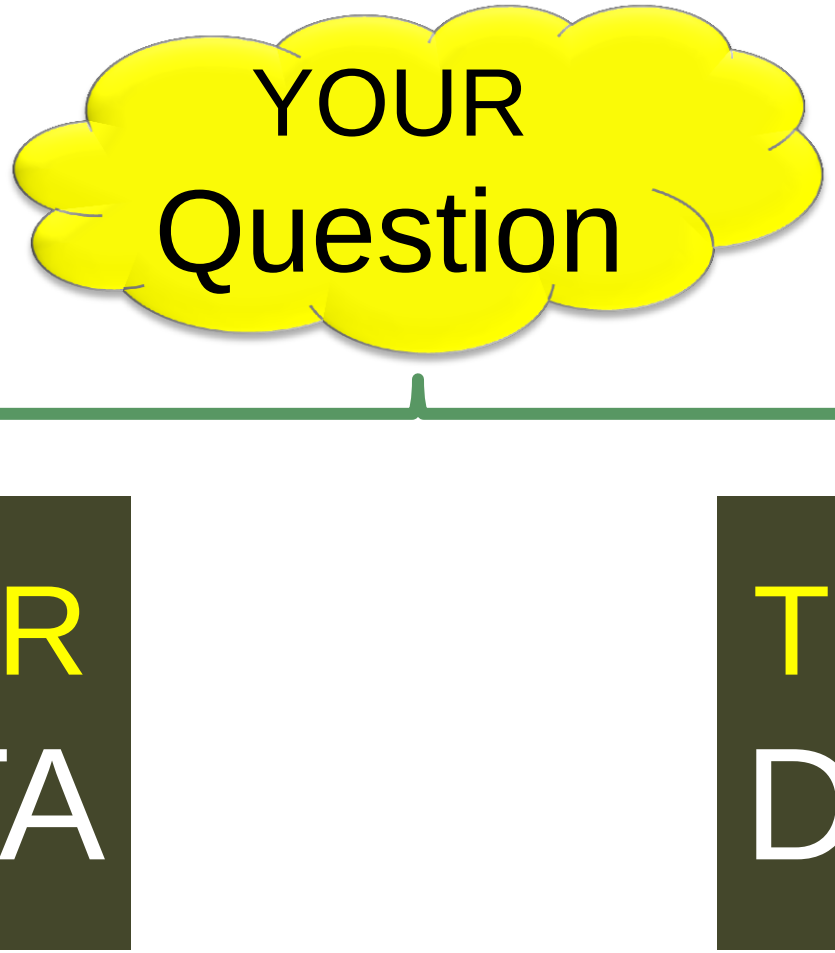
7

EVOLVED RAPIDLY INTO A GLOBAL MOVEMENT

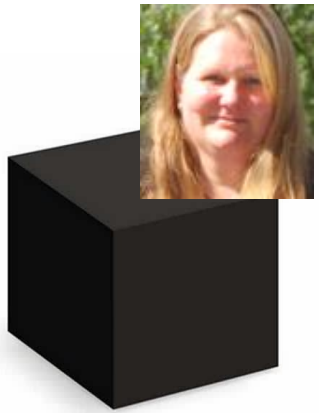
Rapid acceptance and endorsement process

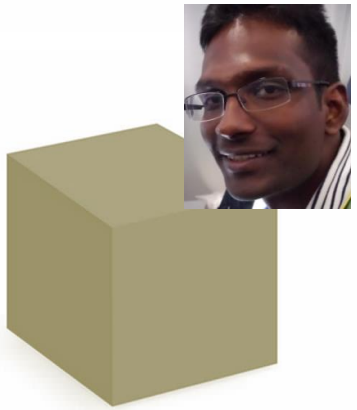
- ✓ The conference
- ✓ The Website
- ✓ Research Data Alliance endorsement
- ✓ DTL flagship project
- ✓ FORCE11 international partner
- ✓ Articles accepted in NATURE
- ✓ NIH accepts FAIR compliance in Life Sciences Commons
- ✓ DTL director Prof. Barend Mons Chair High Level Expert Group EC
- ✓ The Personal Health Train Initiative started
- ✓ EC announces European Open Science Cloud with FAIR as leading principle

Purpose of FAIR data principles



Scenario





Patients in four independent registries

Is there a candidate treatment for Monica?

Their data

Is there a candidate treatment for Monica?




Disclaimer: mock examples!!!!

Four registries

Monika

Krankheit Ringbildung Chromosom 14, Salaam-Anfälle, (Keine Behandlung)

Rajaram

பெர்ரி நோய்க்குறி, வலிப்பு தாக்குதல்கள், லாமோட்ரைஜின்

The ‘I’ of FAIR

Assuming data are findable and accessible: can we find the candidate treatment efficiently?

Annika

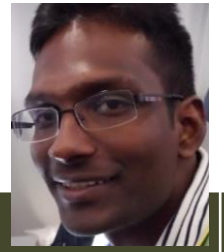
Ring-14-sjúkumynd, sankta Vitusar dansur, eingin viðgerð

Pietro

sindrome Perry, sbalzi d'umore estremi, ossalato

RD Connect elixir BBMRI Biobanking and Biomolecular Resources Research Infrastructure LUMC Leiden University Medical Center

Find the treatment experiment

Monika 	Annika 	Rajaram 	Pietro 	FAIR
Krankheit Ringbildung Chromosom 14, Salaam-Anfälle, (Keine Behandlung)	Ring-14-sjúkumynd, sankta Vitusar dansur, eingin viðgerð	பெர்ரி நோய்க்குறி, வலிப்பு தாக்குதல்கள், லாமோட்ரைஜின்	sindrome Perry, sbalzi d'umore estremi, ossalato	Local registry

Find the treatment experiment

Monika	Annika	Rajaram	Pietro	FAIR
Krankheit Ringbildung Chromosom 14, Salaam-Anfälle, (Keine Behandlung)	Ring-14-sjúkumynd, sankta Vitusar dansur, eingin viðgerð	பெர்ரி நோய்க்குறி, வலிப்பு தாக்குதல்கள், லாமோட்ரைஜின்	sindrome Perry, sbalzi d'umore estremi, ossalato	Local registry
Ring-14 disease, Salaam seizures, (no treatment)	Ring-14 syndrome, Chorea, (no treatment)	Perry syndrome, Epileptic attacks, lamotrigine	Perry syndrome, extreme mood swings, oxalate	English

Find the treatment experiment

Monika	Annika	Rajaram	Pietro	FAIR
Krankheit Ringbildung Chromosom 14, Salaam-Anfälle, (Keine Behandlung)	Ring-14-sjúkumynd, sankta Vitusar dansur, eingin viðgerð	பெர்ரி நோய்க்குறி, வலிப்பு தாக்குதல்கள், லாமோட்ரைஜின்	sindrome Perry, sbalzi d'umore estremi, ossalato	Local registry
Ring-14 disease, **Salaam seizures**, (no treatment)	**Ring-14 syndrome**, **Chorea**, (no treatment)	**Perry syndrome**, **Epileptic attacks**, **lamotrigine**	Perry syndrome, **extreme mood swings**, **oxalate**	English

Semantics:
Person, Disease, Phenotype, Treatment

Scenario

obo: http://purl.obolibrary.org/obo/
ordo: http://www.orpha.net/ORDO/

Monika	Annika	Rajaram	Pietro	FAIR
Ring-14 disease, Salaam seizures, (no treatment)	Ring-14 disease, Chorea, (no treatment)	Perry syndrome, Epileptic attacks, lamotrigine	Perry syndrome, extreme mood swings, oxalate	English
ORPHA1440, HP:0011097	ORPHA1440, HP:0011097	ORPHA178509, HP:0011097, CHEBI_6367	ORPHA178509, HP:0000720 CHEBI_132952	Coded

Identifiers for
Disease, Phenotype, Treatment

Putting the pieces together

ORPHA178509, HP:0000720, CHEBI_132952

ORPHA1440, HP:0011097

ORPHA178509, HP:0011097, CHEBI_6367

ORPHA1440, HP:0002072

Not a global format that computers can understand

No meaning for a computer

Scenario

Monika	Annika	Rajaram	Pietro	FAIR
ORPHA1440, HP:0011097	ORPHA72, HP:00027072	ORPHA178509, HP:0011097, CHEBI_6367	ORPHA178509, HP:0000720 CHEBI_132952	Coded
Monika ***has disease*** Ring-14 disease, and ***has phenotype*** Salaam seizures	Annika ***has disease*** Ring-14 disease, and ***has phenotype*** Chorea	Rajaram ***has disease*** Perry syndrome, and ***has phenotype*** Epileptic seizures. Epileptic seizures ***are treated by*** lamotrigine	Pietro ***has disease*** Ring-14 disease, and ***has phenotype*** Extreme mood swings. Extreme mood swings ***are treated by*** the drug Oxalate	full meaning

Semantics:
Person *has disease* Disease,
Person *has phenotype* Phenotype,
Phenotype *is treated by* Treatment

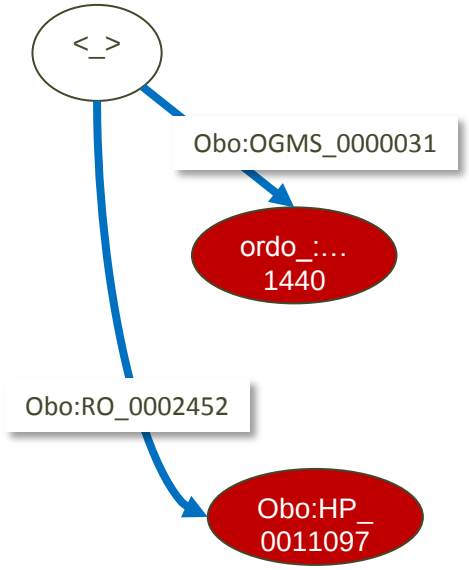
Scenario

obo: http://purl.obolibrary.org/obo/
ordo: http://www.orpha.net/ORDO/

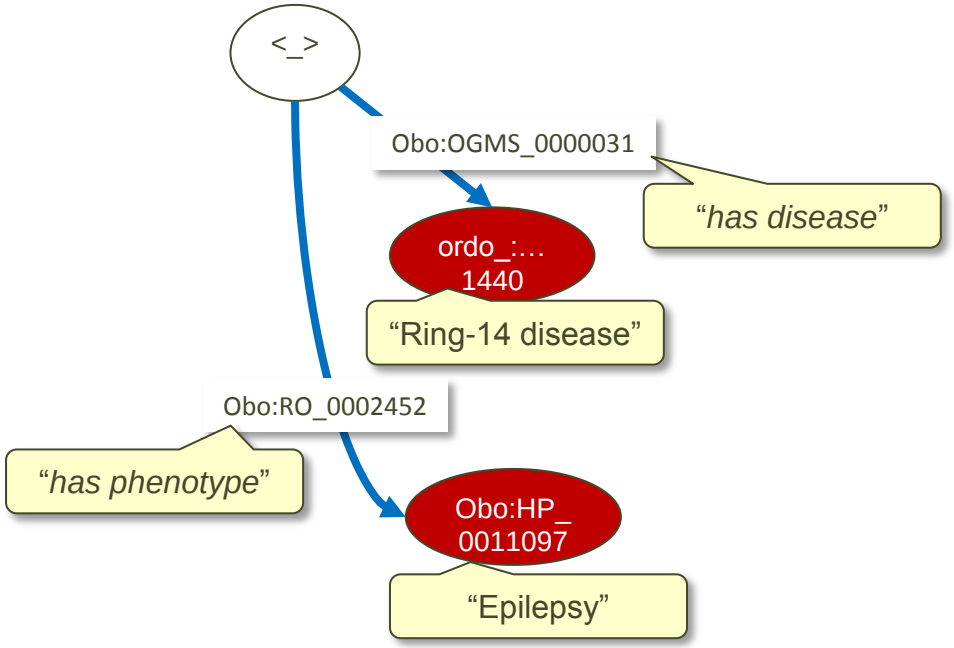
Monika	Annika	Rajaram	Pietro	FAIR
Monika ***has disease*** Ring-14 disease, and ***has phenotype*** Salaam seizures	Annika ***has disease*** Ring-14 disease, and ***has phenotype*** Chorea	Rajaram ***has disease*** Perry syndrome, and ***has phenotype*** Epileptic seizures. Epileptic seizures ***are treated by*** lamotrigine	Pietro ***has disease*** Ring-14 disease, and ***has phenotype*** Extreme mood swings. Extreme mood swings ***are treated by*** the drug Oxalate	Full meaning
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952	Interoperable & Machine readable

Machine readable semantics (RDF) for:
Person *has disease* Disease,
Person *has phenotype* Phenotype,
Phenotype *is treated by* Treatment

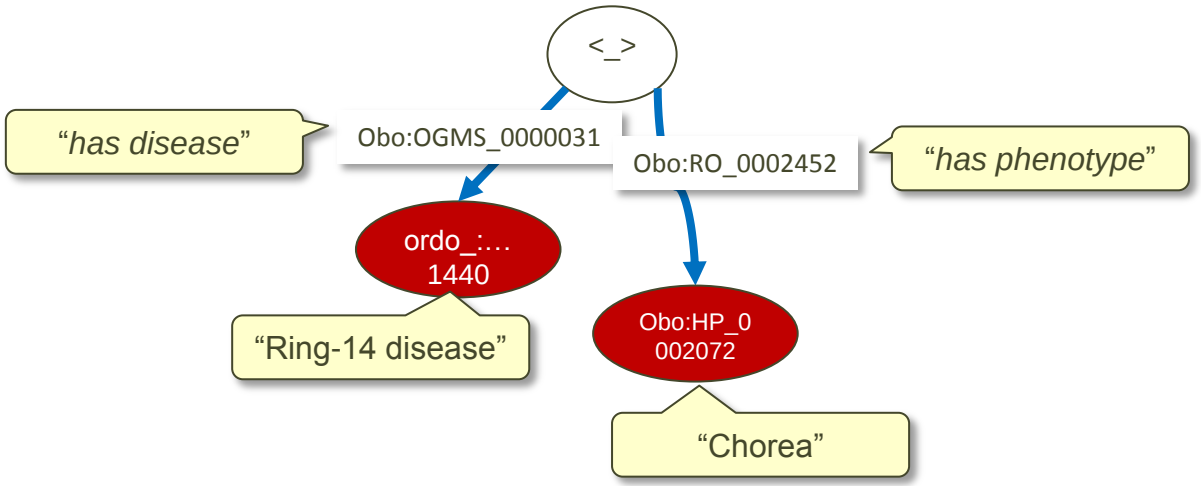
Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



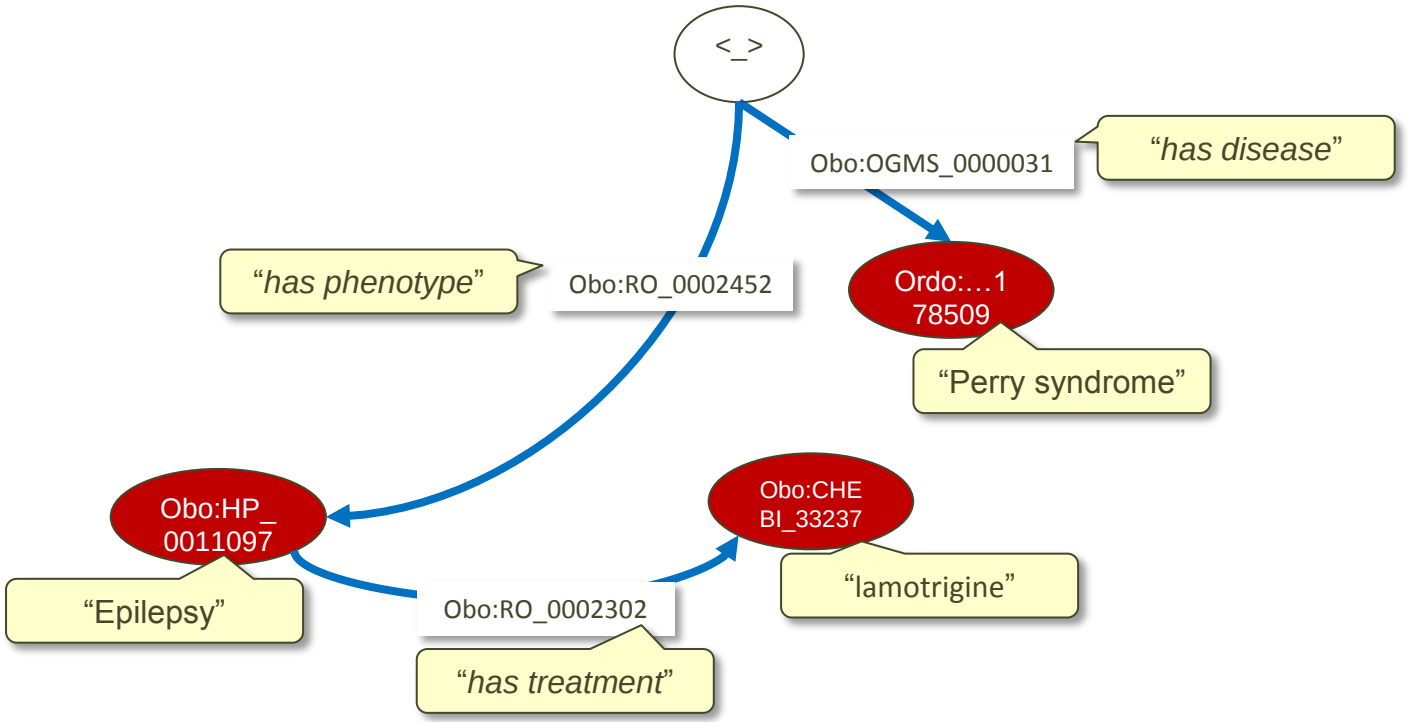
Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



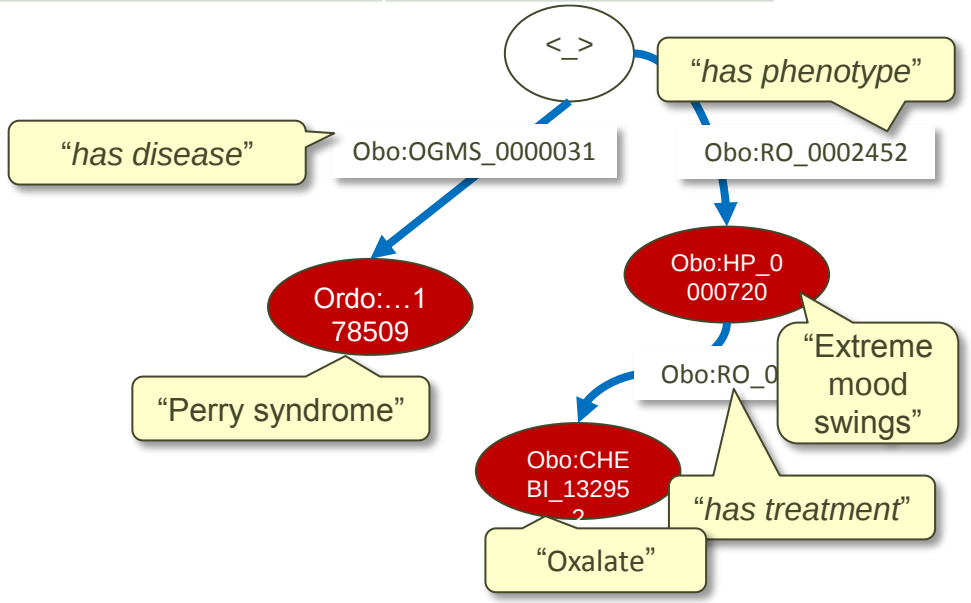
Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



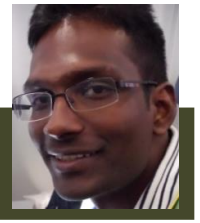
Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



Recap

Monika

Krankheit Ringbildung Chromosom 14, Salaam-Anfälle, (Keine Behandlung)



Rajaram

பெர்ரி நோய்க்குறி, வலிப்பு தாக்குதல்கள், லாமோட்ரைஜின்

Starting point:
4 independent,
incompatible data sources

Not findable, accessible, interoperable, reusable

Annika

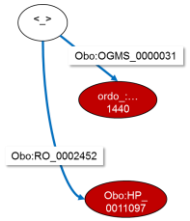
Ring-14-sjúkumynd, sankta Vitusar dansur, eingin viðgerð

Pietro

sindrome Perry, sbalzi d'umore estremi, ossalato

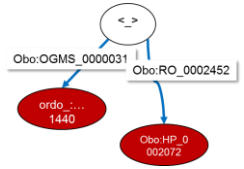
RD-Connect elixir BBMRI Biobanking and Biomolecular Resources Research Infrastructure LUMC Leiden University Medical Center

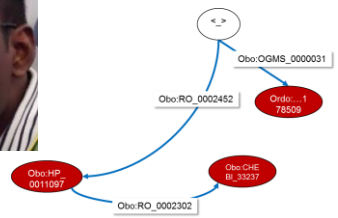
FAIR data landscape

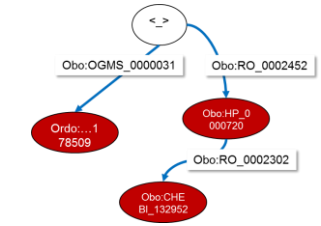


Now we have four independent **interoperable** data sources, *still under control of the local data manager*

(e.g. HCP, patient organisation, patient)

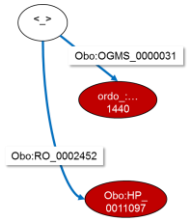




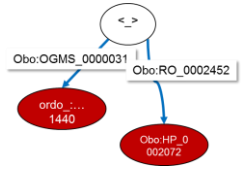


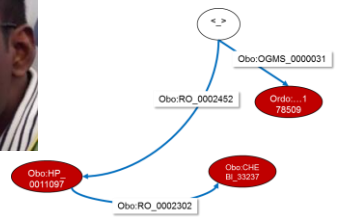
RD Connect elixir BBMRI Biobanking and Biomolecular Resources Research Infrastructure LUMC Leiden University Medical Center

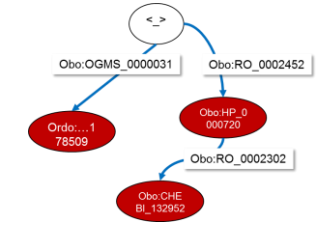
FAIR data landscape



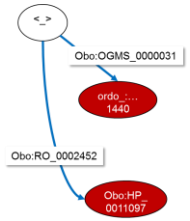
Data at each source is **self-explaining** through global standards that computers understand

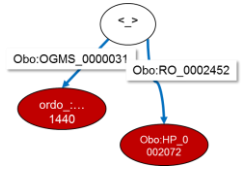


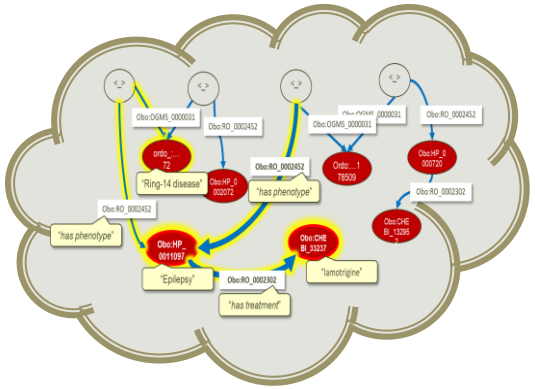


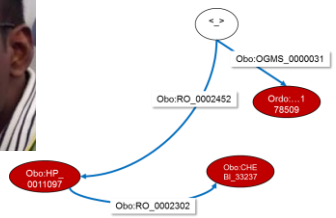


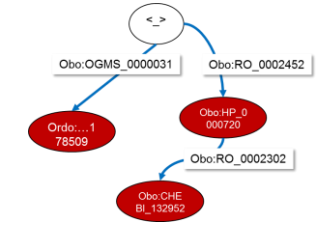
FAIR data landscape



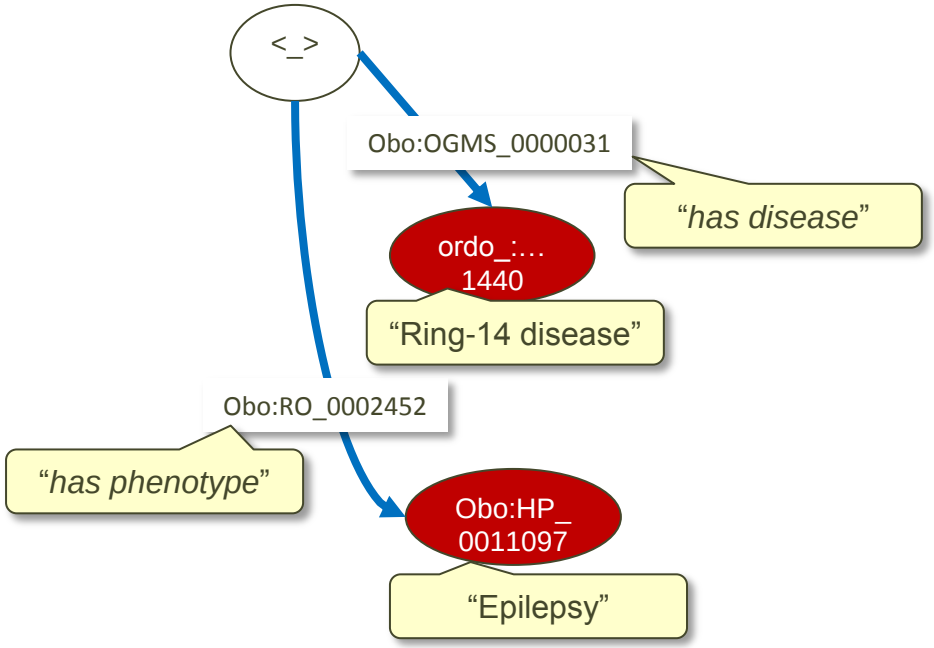




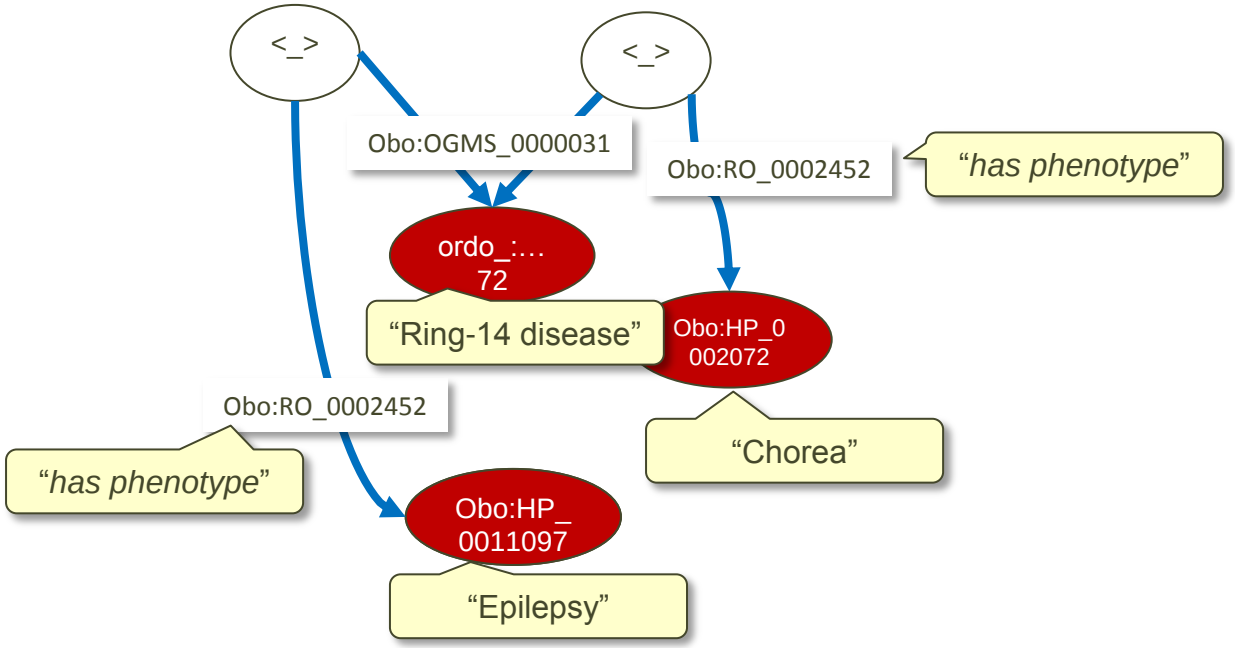




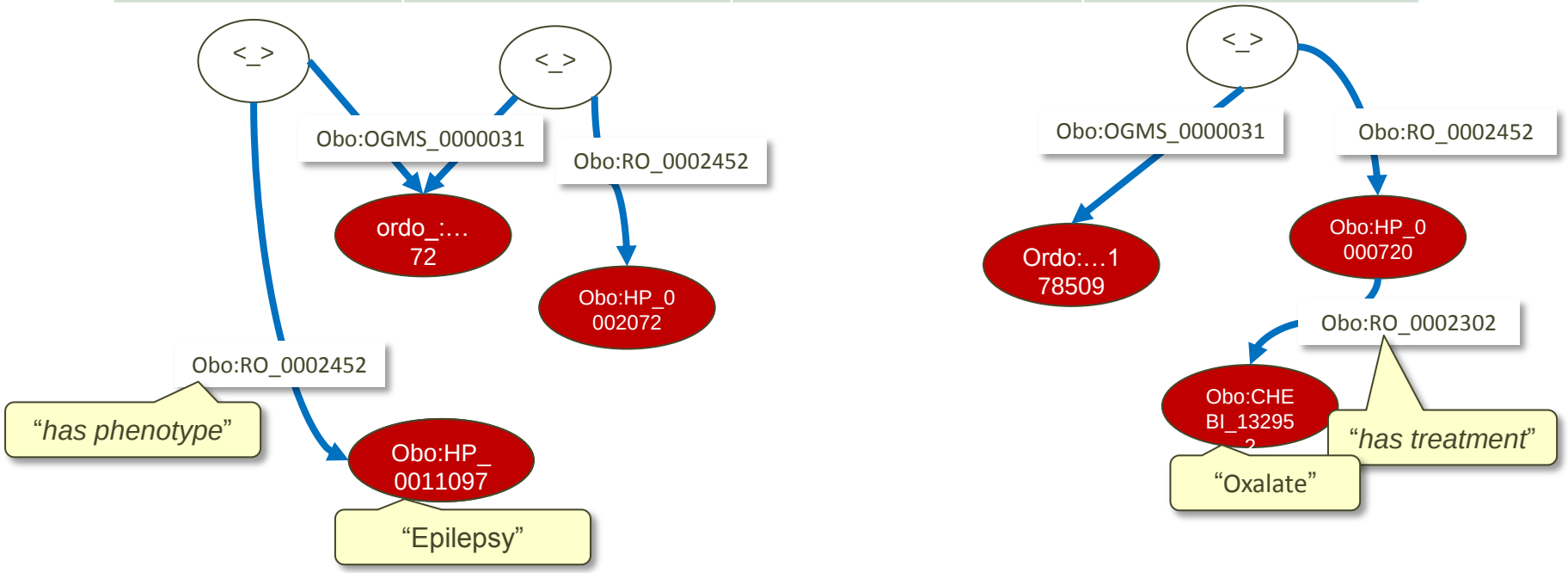
Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



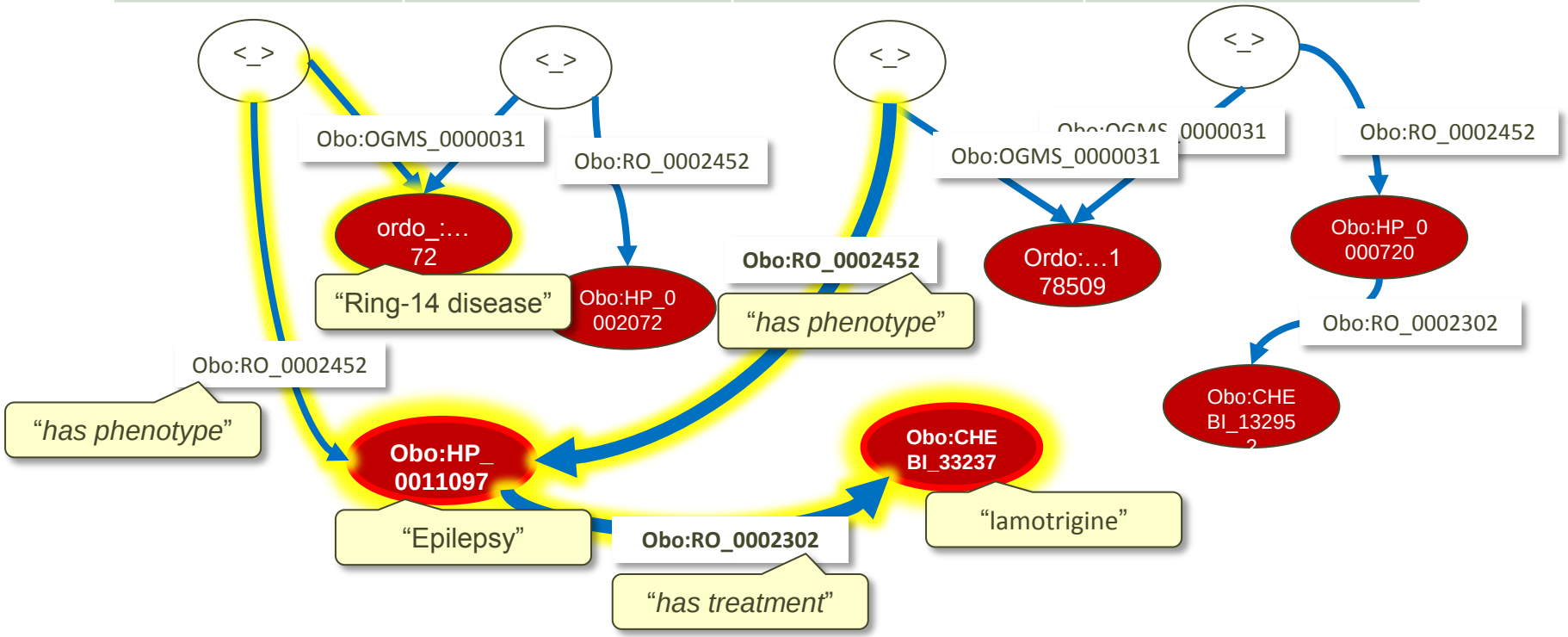
Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952

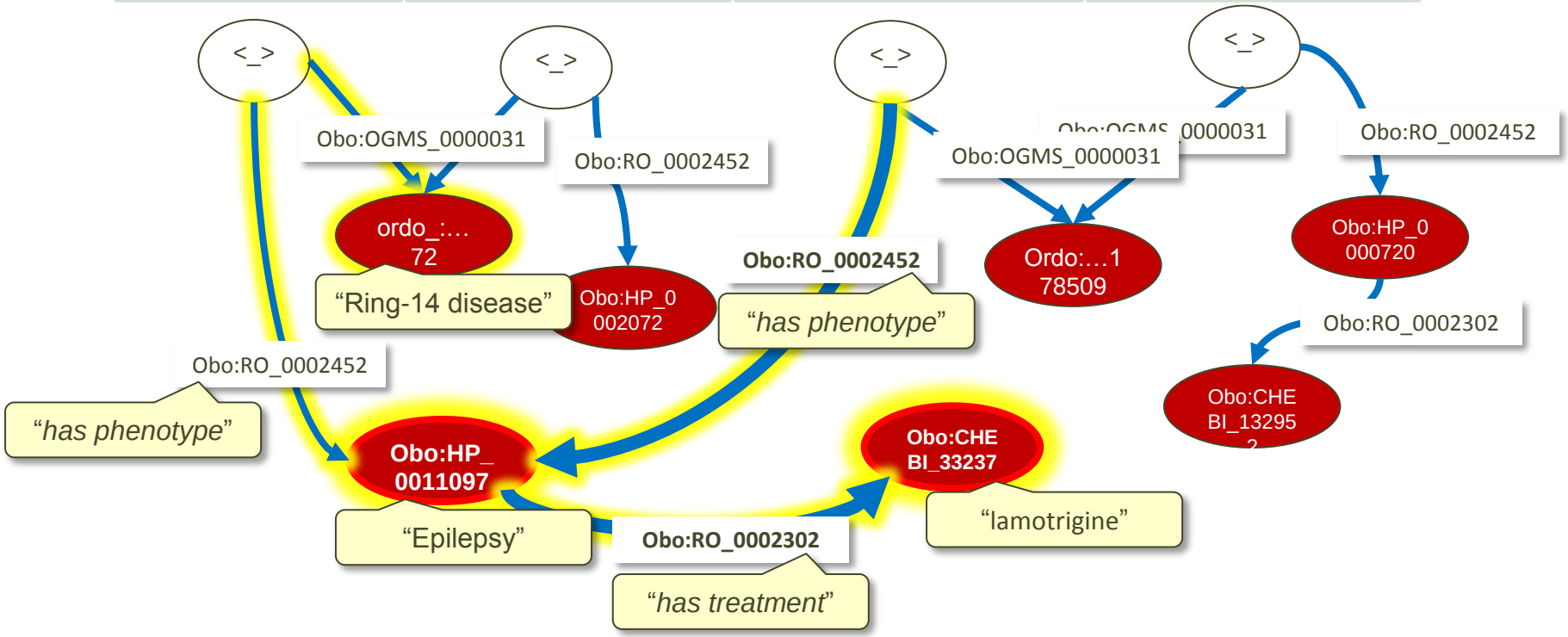


Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



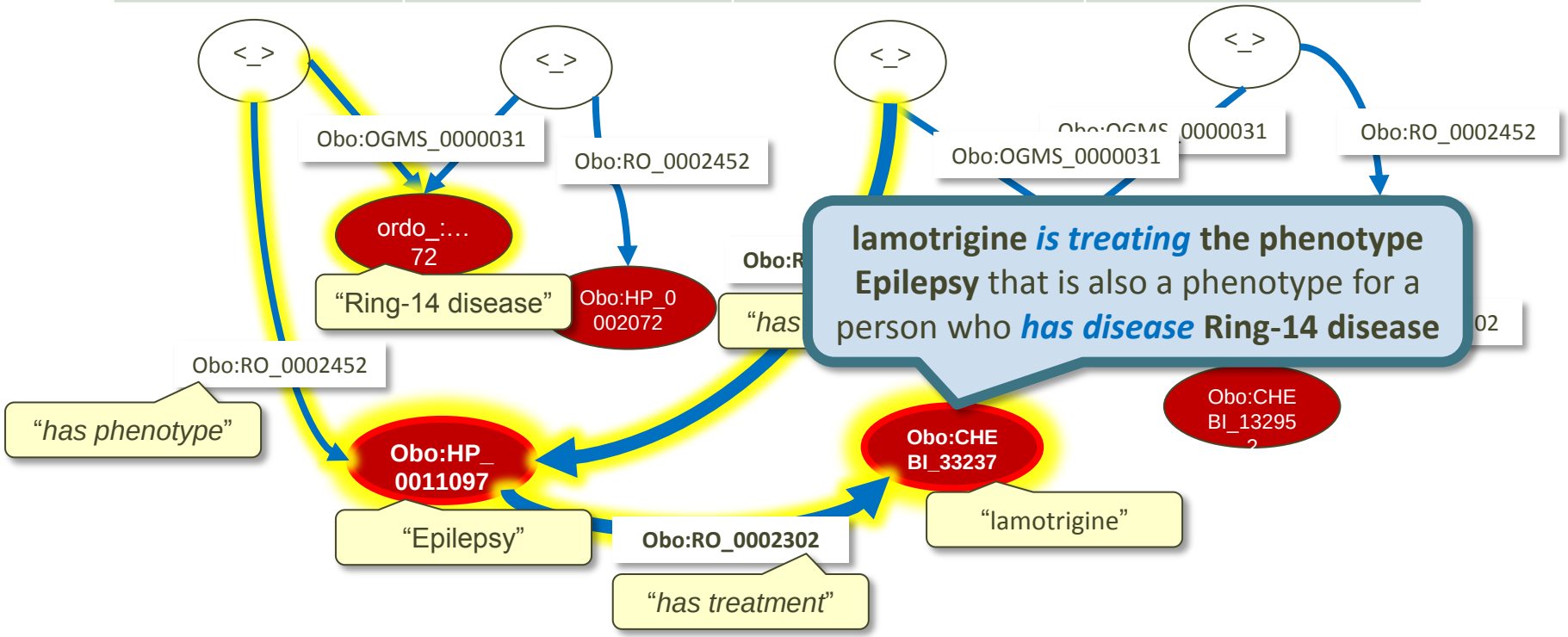
Which **treatment** ***is treating*** **the phenotype** that is also a phenotype for Monica who ***has disease*** **Ring-14 disease**

Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952



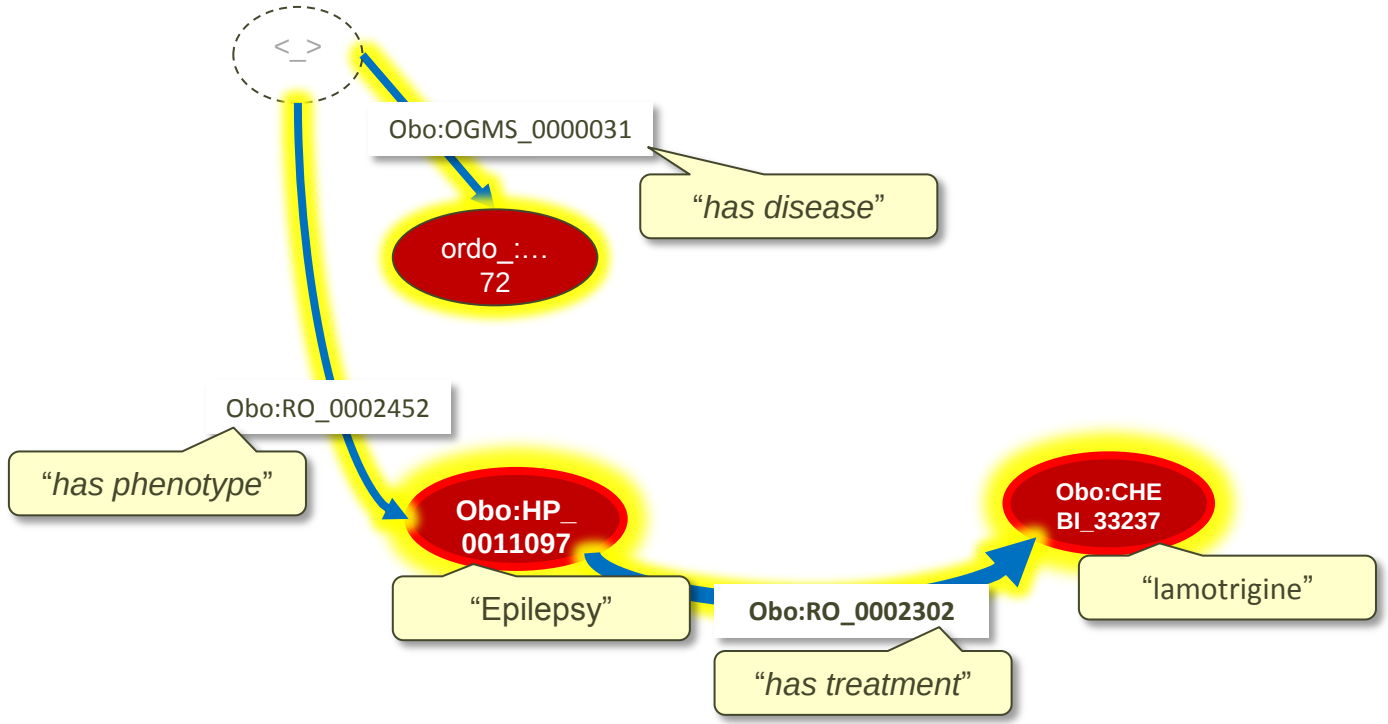
Which **treatment *is treating*** **the phenotype** that is also a phenotype for Monica who ***has disease*** **Ring-14 disease**

Monika	Annika	Rajaram	Pietro
<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440 ***obo:RO_0002452*** obo:HP_0011097.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_1440, ***obo:RO_0002452*** obo:HP_0002072.	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0011097 ***obo:RO_0002302*** obo:CHEBI_33237	<_> ***obo:OGMS_0000031*** ordo:Orphanet_178509, ***obo:RO_0002452*** obo:HP_0000720 ***obo:RO_0002302*** obo:CHEBI_132952

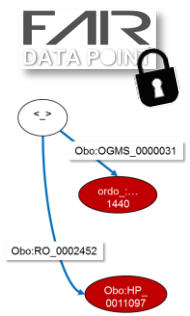


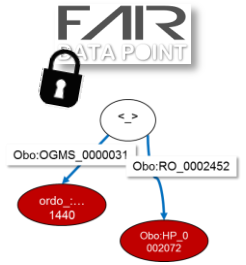
Result

Disease | global machine readable code (URI)		**Treatment** | URI		**Phenotype** | URI	
Ring-14 disease	ordo:Orphanet_1440	Lamotrigine	Obo:CHEBI_33237	Epilepsy	Obo:HP_0011097

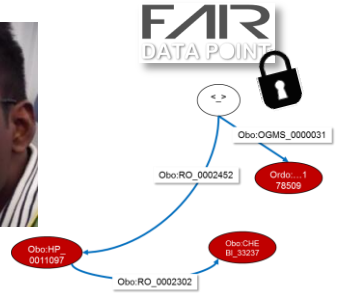


FAIR data landscape

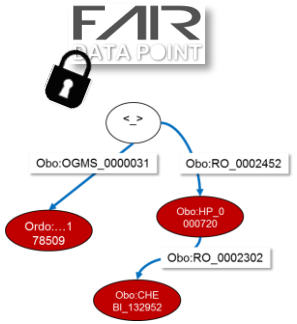




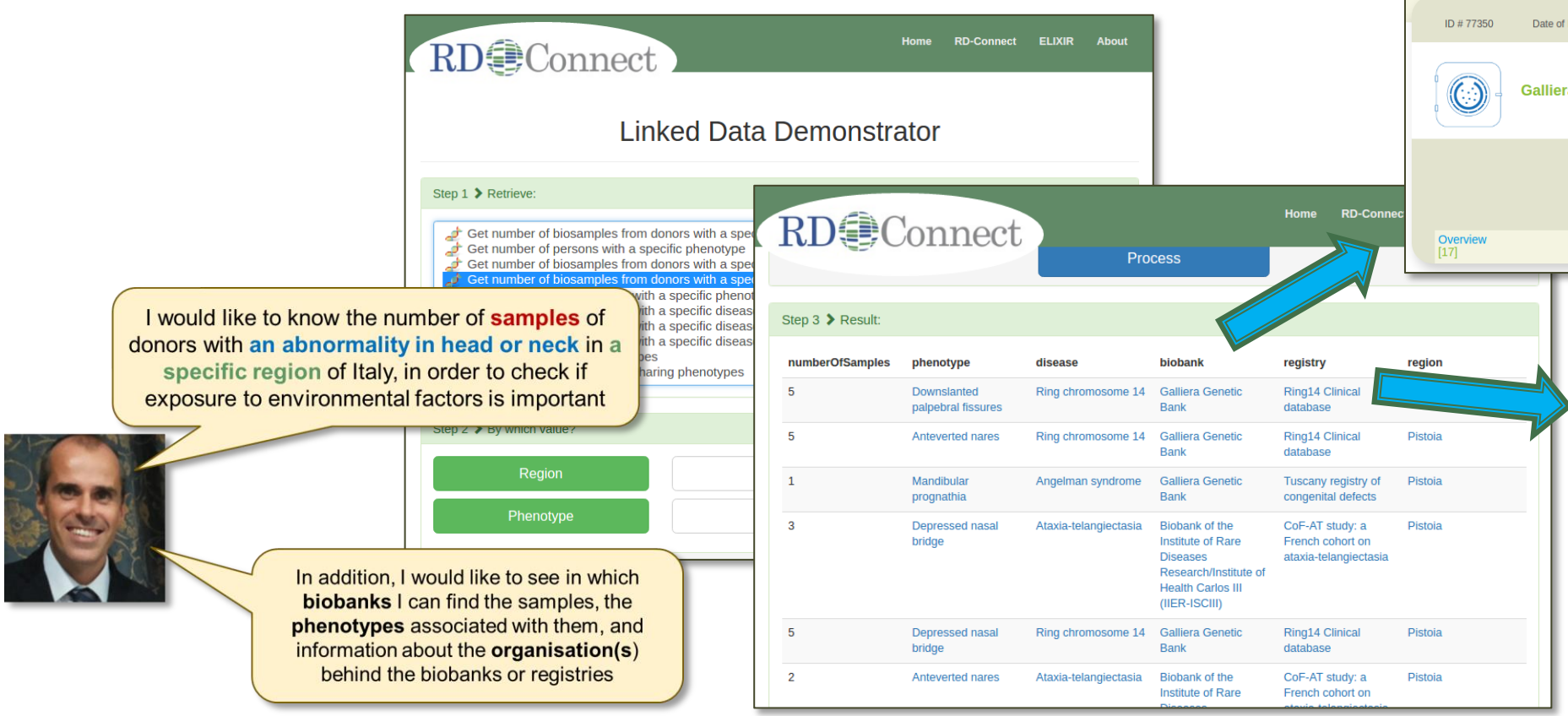
Data at each source is **self-explaining** through global standards that a computer program can understand



Data is communicated via a FAIR data point *under well-defined conditions*



RD-Connect: proof of concept questions across biobanks and registries
ELIXIR: Test interoperability components



Driving user questions

Demonstrator UI

ID Cards

RD Connect elixir BBMRI Biobanking and Biomolecular Resources Research Infrastructure LUMC Leiden University Medical Center

Proposal: define strategy

Who is the expert on your data?

Who (and what) is needed to make VASCERN data sources FAIR?

Who should make the data FAIR?

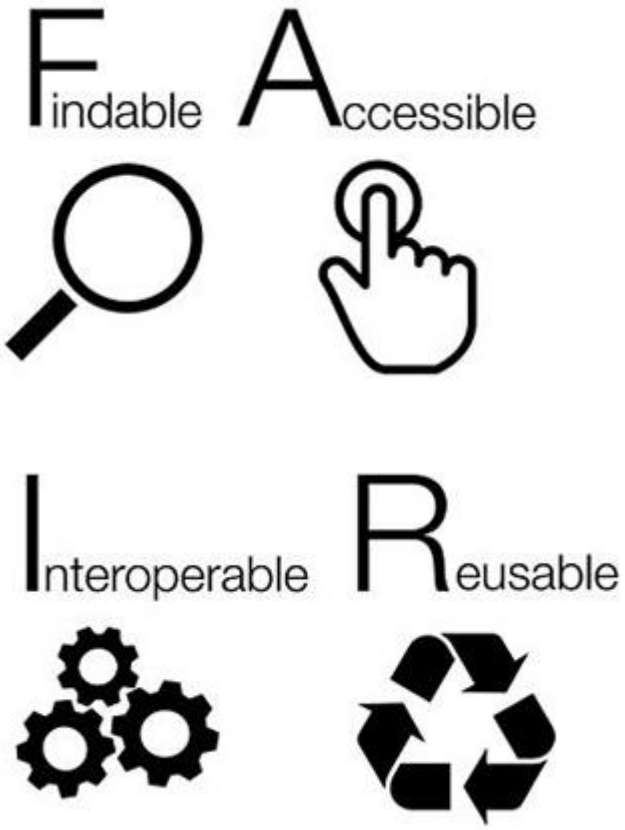
Guidelines/protocols

Tooling

Training

Data stewards

FAIR implementation strategy (proposal)



Strategy: strengthen data sources

- Organise FAIR data stewardship
  - local, national, international
- **Do it** while contributing to infrastructure development
  - define & apply coding that conveys the meaning of VASCERN data cf. global standards
    (i.e. ontologies and linked data, e.g. OMOP, FIHR, LOINC, HPO, ORDO, bio-ontologies)
  - use global data access mechanisms (FAIR data points)
    - Incorporate ELSI rules, PPRL, dynamic consent
  - Start with minimal list of data elements; add data elements for VASCERN research questions

LUMC FAIR data stewardship and Research ICT

FAIR data stewards
BioSemantics group
LUMC Library
LUMC ICT

IT experts contributing to FAIR data software development
(Ontologies, mapping services, FAIR software)

- Support researchers write a FAIR data management plan (with DM budget)
- Require FAIR data principles in software tenders
- Structural financing for FAIR data technical expert
- Librarians in training to become FAIR data stewards
- Contribution to international FAIR software engineering team

RD Connect elixir BBMRI Biobanking and Biomolecular Resources Research Infrastructure LUMC Leiden University Medical Center

Thank you

Acknowledgements

FAIR Data engineering team
Luiz Bonino, Rajaram Kaliyaperumal, Kees Burger, Nuno Nunes, Shamanou van Leeuwen, Mark Thompson

Skunk team / FAIR metrics
Mark Wilkinson, Michel Dumontier

LUMC Human genetics department BioSemantics group, ADM,

Elixir: Team Evelo (Maastricht), Team Goble (Manchester), Team Gut/Beltran (CNAG), Team Parkinson (EBI), Team Poch (Strassbourg);

GO-FAIR: Barend Mons; **DTL**: Mascha Jansen, Celia van Gelder, Erik Schultes, Albert Mons

ERNs, patient representatives and patient organisations

Collaborating RD developers / FAIR data experts
David van Enckevort (UMCG), Annika Jacobsen (LUMC), Andra Waagmeester (Micelio); Heimo Muller, Robert Reihs (Graz, Austria); Pedro Sernadella, Jose Oliveira (Aveiro, Portugal); Marc Hanauer, Ana Rath (Orphanet, Paris); Roxana Merino (Karolinska, Sweden); Matthias Brochausen (USA); Katy Wolstencroft (LIACS); Rob Cornelisse (LUMC); Developers of Castor, RDRF, OSSE, MolGenis

Liaisons/case owners: Rachel Thompson, Libby Wood, Claudio Carta (Rome, Italy), Domenica Taruscio (Rome, Italy), Marco Crimi, Estrella Gomes, Marina Mordenti, Freddie Ehrhart, Carina van Vleuten, Maria Jongma

RDs GO FAIR seed group & liaisons
Virginie Bros-Facer, Vicky Hedley, Ana Rath, Marc Hanauer, Rachel Thompson, Ronald Cornet, Mark Wilkinson, David van Enckevort, Leo Schultze Kool

RD Connect | LUMC Leiden University Medical Center | elixir