

Joint controllership: the case of HEAP

Martin Boeckhout, senior consultant, MLCF
April 13 2021, webinar on joint controllership

www.heap-exposome.eu

MEMBER OF THE



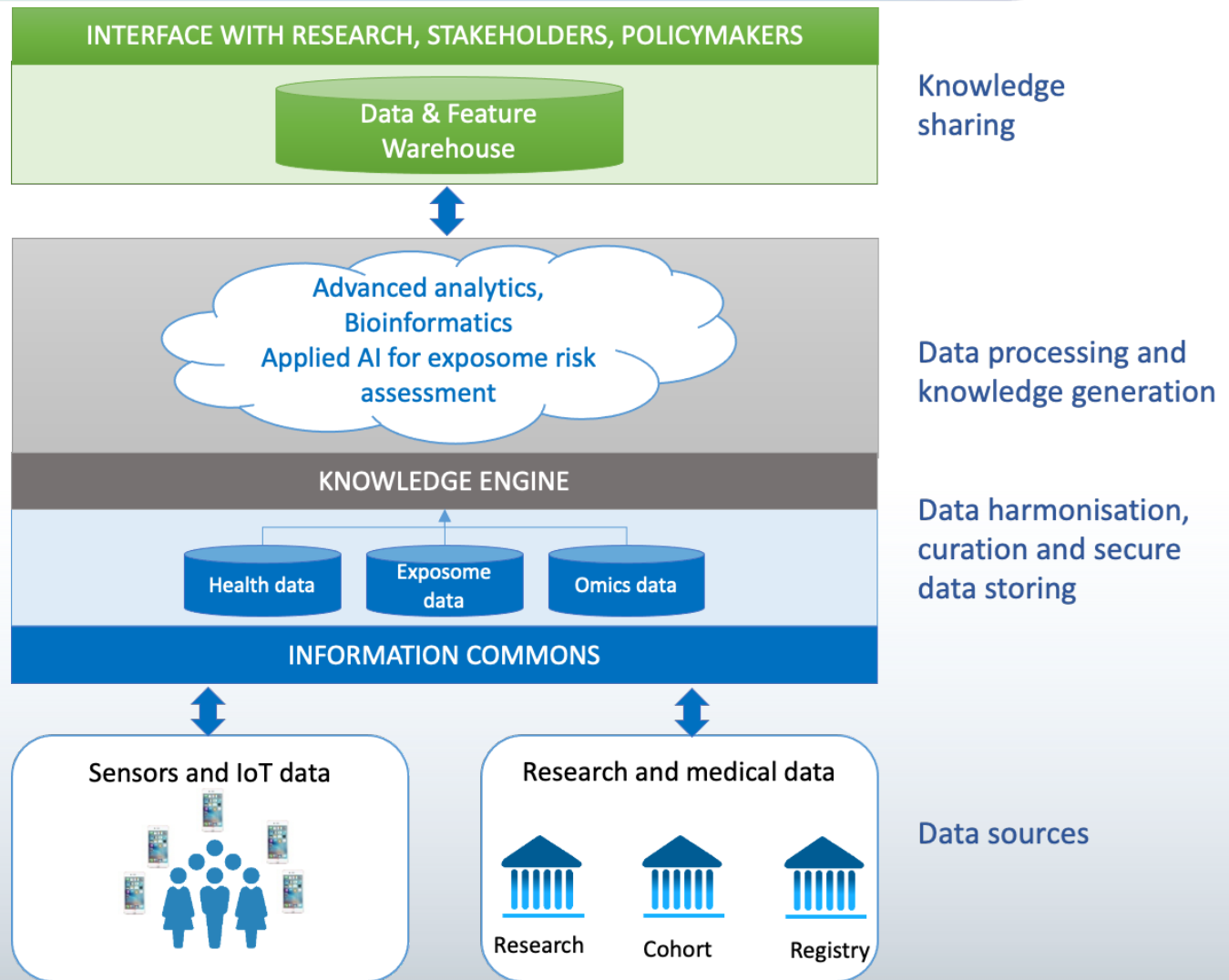
European Human Exposome **NETWORK**

HEAP

Human Exposome Assessment Platform

Data Management Plan

10.5281/zenodo.4043111



Data Summary: “use cases”

- **Cervical screening cohort (Sweden, KI):** ~800 000 samples from ~450 000 women, 140 000 new women enrolled yearly. Sequencing analysis from samples from healthy tissue, pre-neoplastic lesions and cancer to characterize risk factors.
- **Maternity cohort (Finland, UOULU):** Data from the Finnish National prenatal screening, FMC screening results (including HIV, syphilis, hepatitis B, rubella). Data collected over generations from ~2M women, part to be sequenced to understand the exposome, over time and over generations.
- **HPV vaccination cohort (Sweden, KI):** 60 000 participants, samples and data from intervention trial, microbiome data including HPV and chlamydia, for microbiomics, epidemiological and impact analysis.
- **Consumer cohort (Denmark, SSI):** consumer purchase data aiming to 10000 participants to continuously model and assess the household consumed exposome's impact on health.
- **Wearable data collection study (Maternity cohort, KI):** Pilot study to monitor exposures on 100 pregnant women in relation to healthy childbearing and produce continuous **personal exposome profiling** of the participants.

FAIR data

- **Making data findable, including provisions for metadata**
 - Discoverability of data (metadata provision)
 - Standards for metadata definition
 - Persistent and unique identifiers
 - Search keywords
 - Versioning
- **Making data openly accessible**
 - Tools for data availability (Information Commons)
- **Making data interoperable**
 - Hopsworks knowledge engine (standardised data management)
- **Increase data re-use**
 - Guarantee widest possible reuse
 - Data quality assurance processes
 - Security + GDPR
- **BYOD workshops for Data FAIRness**

Governing controllership



A spectrum of potential arrangements, each with pros & cons

Platform as a 'pure' service

- Data providers share controllership with users on a case-by-case basis
- Governance sets minimal terms of data provision & use
- Contracts of data controllers and – processor and data provider user on a case by case basis

Platform as a joint facility

- Data providers and users have an active role in furthering the platform's capabilities and appropriate use
- Governance sets terms for collaboration and conditions for use

The HEAP Consortium

More than just a Platform

- ✓ **A “common pool resource”:**
intermingling of roles of data providers,
researchers and platform developers
- ✓ **Data collection:** data for some use
cases collected as part of HEAP
- ✓ **Data generation:** Production &
archiving of data and ML features
- ✓ **Terms of (re)use:** FAIR data & re-use
– also by third parties

Governing controllership



+

- Recognisable proposition to third parties
- Clear roles from a DP perspective: platform operators are 'pure processors'

+

- Fits the model of a research (infrastructure) consortium
- FAIRness and re-use 'by organisational design'



Platform as a 'pure' service



Platform as a joint facility

-

- Current roles intermingled
- Platform development
- Furthering FAIRness, re-use
- What about consortium-funded data?

-

- Clarity of roles from a DP perspective:
- Consortium members merely facilitating data processing: support, development, maintenance

Possible steps towards resolution

Sorting out controllership

<i>Kinds of processing</i>	Data provision	Research access & use within consortium	Research access & use by third parties	Data archiving (longer-term data availability)
<i>Kinds of data</i>				
"Extant" data sources: use cases	Provider	Provider <i>and</i> consortium researchers?	* / Depends: does consortium decide on terms of use?	N/A
Consortium-funded use cases	Provider <i>and</i> consortium researchers?	Provider <i>and</i> consortium researchers?	* / Depends: does consortium decide on terms of use?	* / Post-consortium
Third-party data sources	Provider	N/A	* / Depends: does consortium decide on terms of use?	* / Post-consortium
Data generated on the platform	N/A	Consortium researchers?	* / Depends: does consortium decide on terms of use?	* / Post-consortium

* / Post-consortium : to be settled in a governance and data access framework
Conditional on H2020 consortium agreement; FAIR and re-use terms, ...

HEAP PROJECT

EU Funding



The HEAP project has received funding from the European Union's Horizon 2020 research programme under grant agreement No 874662.

HEAP PROJECT

Partners



Karolinska
Institutet



STATENS
SERUM
INSTITUT



C S C

LOGICAL CLOCKS



Med Uni
Graz

universität
innsbruck

STICHTING MLC FOUNDATION



UNIVERSITY
OF WARSAW



Tays
For Life



UNIVERSITY
OF OULU

International Agency for Research on Cancer



World Health
Organization