



# Health data exploitation ecosystem

FinnGen company co-operation seminar  
16.3.2018, Team Finland –talo, Helsinki

Jaakko Lähteenmäki  
Principal Scientist

VTT Technical Research Centre of Finland

# Background



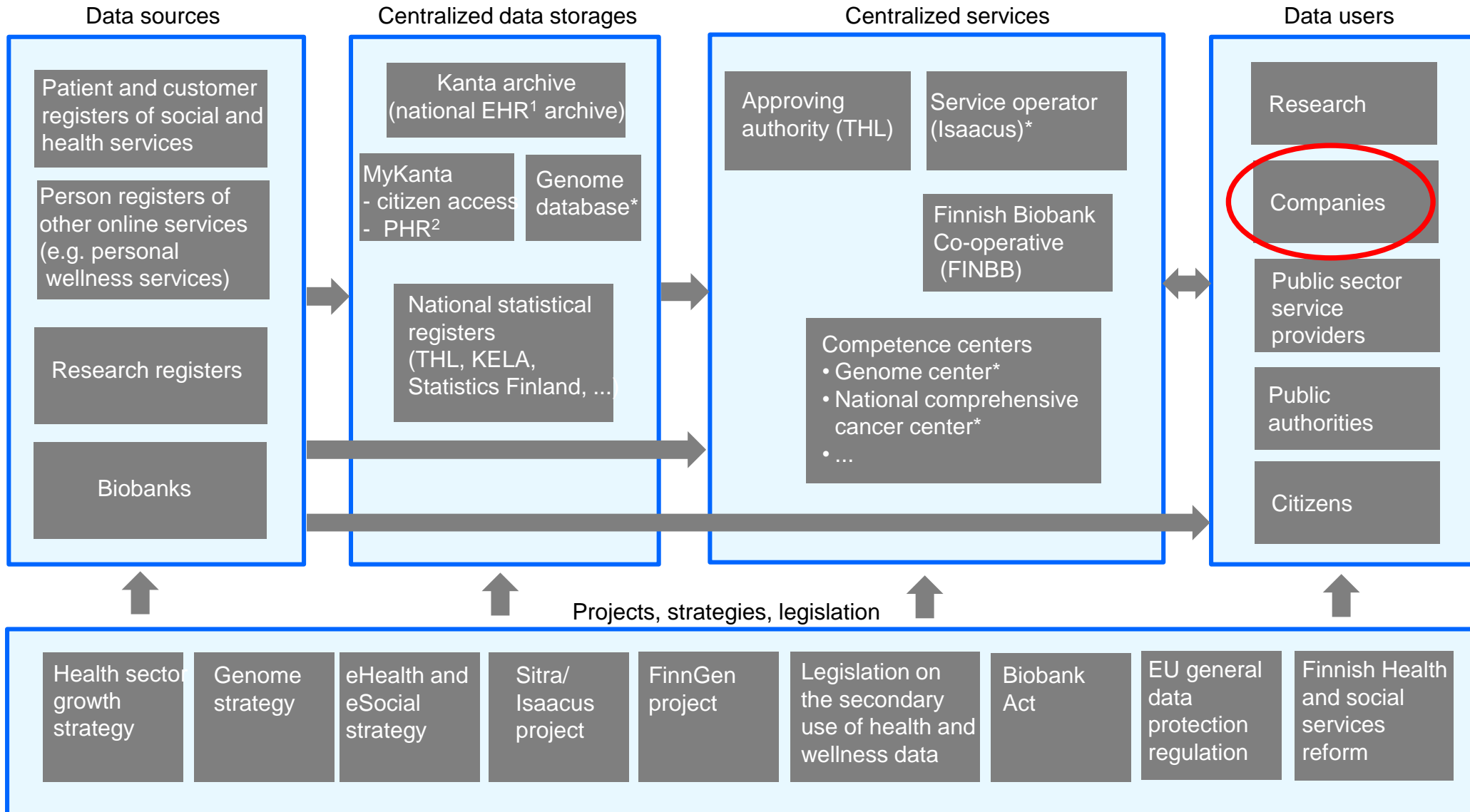
**Data will be increasingly available for secondary use**

→ How companies can benefit ?

## **VTT's PreMed project objectives:**

1. Identify and increase **business opportunities for Finnish companies** based on exploitation of health related data
2. Provide **recommendations for public bodies** concerning financing and other measures needed in boosting ecosystem development

# Finnish health data exploitation infrastructure



\* under development or planning

<sup>1</sup> Electronic Health Record

<sup>2</sup> Personal Health Record

# Three identified roles\* of private companies to exploit individual-level genotype/phenotype data

## Data owners

Organisations providing **access to data:**



- Isaacus operator
- Biobank operator
- Individual biobanks in Finland and abroad
- Other data controllers

biobank data/sample application, publication, results feedback



Isaacus service operator data access process (to be defined)



## 2. Research provider

Companies and organisations with the capability to **conduct research** and take responsibility of managing individual-level data/samples:



- Research and expert services for pharma and healthcare
- Universities and research institutes

request/answer



## 1. Research customer

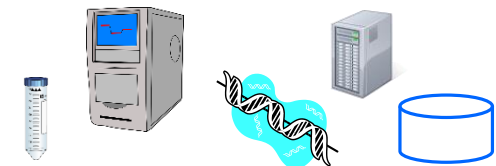
Companies with **research questions and information needs:**



- Pharma and diagnostics companies
- B2B genetic testing services
- B2C genetic testing services
- Healthcare services
- Insurance services
- Healthcare information system and application providers
- Medical device manufacturers
- Food industry
- ...

## 3. Service or technology provider

Companies **supporting data exploitation** by providing omics, data management, data analytics, software services and related technology



# Organisation classification and interviews carried out in PreMed phase 1\*



Organisation classification	Interview	Role
Biobank	Auria biobank, Finnish biobank Cooperative, Blood Service Biobank	Data operator
Research and expert services for pharma and healthcare	Medaffcon, Medengine, Crown CRO <sup>1</sup>	Research provider
University, research institute	VSSH, UEF, THL	Research provider
Bioinformatics software and services	BC Platforms, Euformatics, Genevia, Medisapiens	Service or technology provider
Sequencing, genotyping and diagnostic services	Blueprint Genetics <sup>1</sup> , Zora Biosciences <sup>1</sup>	Service or technology provider
Medical devices	Modulight	Service or technology provider
Genetic testing, interpretation and counselling services (B2B/B2C)	Abomics, DiagFactor	Research customer
Pharma and/or diagnostics company	Orion, Abbvie, Pfizer	Research customer
Healthcare information systems and applications	Duodecim, BCB Medical, Kaiku Health, Mediconsult <sup>1</sup>	Research customer
Healthcare services	Mehiläinen, Terveystalo, Docrates	Research customer
Insurance services	Lähtapiola	Research customer
Foresight	Sitra	Ecosystem

**\*Note:** The classification and company allocation is initial.  
It is subject to changes after further discussion with the listed companies

<sup>1</sup> Interviewed before PreMed project in 2016

# Opportunity for companies

- **High volumes of health data** is available in Finland and globally for secondary use
- **Access to data** is becoming easier
  - published instructions, criteria and processes
  - centralized, one-stop-shop services
- Dimensionality of data is widening
  - **personal lifestyle and environmental** data becoming more important
  - **population-based biobanks** growing fastest\*
- A **wide range of companies** may benefit from health data
- Accessing and processing individual-level genotype/phenotype data requires specific capabilities
  - specialised companies are needed in "**research provider**" role














# TECHNOLOGY «FOR» BUSINESS











# Background slides

# Examples of data/sample sources (1)

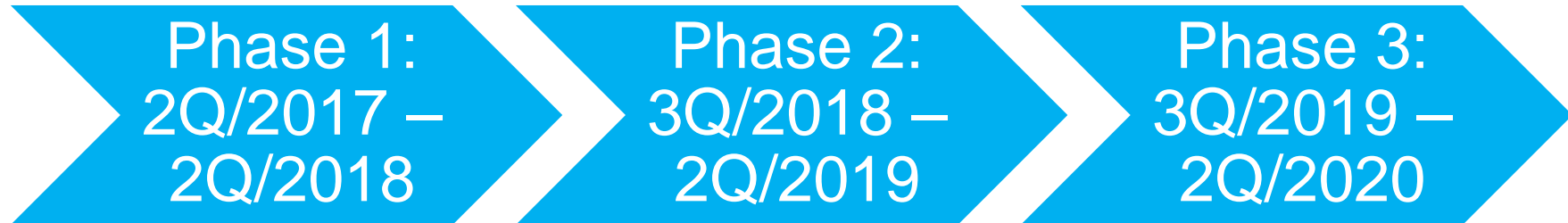


Organisation		Content	Data access process	Link
UK Biobank		Collection of questionnaire data, blood, urine and saliva samples and healthcare follow-up data of 500 000 individuals.	Instructions and costs for access published.	<a href="http://www.ukbiobank.ac.uk/">http://www.ukbiobank.ac.uk/</a>
The 100 000 Genomes Project		The project will sequence 100,000 genomes from around 70,000 people. Participants are NHS patients with a rare disease, plus their families, and patients with cancer.	Data access to researchers based on application. Instructions published. Companies are excluded from full access, but may participate the "Discovery Forum".	<a href="https://www.genomicsengland.co.uk/the-100000-genomes-project/">https://www.genomicsengland.co.uk/the-100000-genomes-project/</a>
Norway HUNT		Questionnaire data and biological specimens from the general population.	Instructions and costs for access published. Norwegian research partner required.	<a href="https://www.ntnu.edu/hunt/data">https://www.ntnu.edu/hunt/data</a>
Norwegian Institute of Public Health		Joint service for accessing data from health registries and health studies as well as samples from Norwegian biobanks.	Instructions and costs for access published.	<a href="https://www.fhi.no/en/">https://www.fhi.no/en/</a>
Cancer registry of Norway / Janus Serum Bank		Population based biobank for cancer research.	Instructions for access published. Norwegian research partner required.	<a href="https://www.kreftregisteret.no/en/Research/Janus-Serum-Bank/">https://www.kreftregisteret.no/en/Research/Janus-Serum-Bank/</a>
Danish National Biobank		National biobank with large sample collection and information about samples elsewhere in the Danish health system.	Instructions for access published. Danish research partner needed.	<a href="http://www.biobankdenmark.dk/">http://www.biobankdenmark.dk/</a>
Biobank Sweden		Centralized service point for Swedish biobanks.	Instructions and forms available for applying data (application shall be directed to the specific biobank owning the sample/data).	<a href="http://biobanksverige.se/research/basic-information-in-english/">http://biobanksverige.se/research/basic-information-in-english/</a>
LifeGene Sweden		Prospective cohort study including periodically collected questionnaire data, physical measurements and samples.	Instructions and costs for access published. Association with Swedish university or research institute required.	<a href="https://www.lifegene.se/">https://www.lifegene.se/</a>
The Estonian Biobank (University of Tartu)		Population based biobank including a cohort of > 50 000 people.	Instructions for access published.	<a href="https://www.geenivaramu.ee/en">https://www.geenivaramu.ee/en</a>
Hatwig Medical foundation (Netherlands)		A database of biopsy samples and clinical parameters of patients recruited for studies of the Center for Personalized Cancer Treatment (CPCT).	Instructions for access published.	<a href="https://www.hartwigmedicalfoundation.nl/en/">https://www.hartwigmedicalfoundation.nl/en/</a>
Biobank Graz		Comprises both population-based and disease-focused collections of human biological material.	Instructions and forms available for applying data.	<a href="http://biobank.medunigraz.at/en/">http://biobank.medunigraz.at/en/</a>

# Examples of data/sample sources (2)

Organisation	Content	Data access process	Link
Auria Biobank 	Provides biological samples collected in connection with normal health-care and medical examinations for use in medical research.	Instructions and forms available for applying data.	<a href="https://www.auriabiopankki.fi/auriapalvelut/auriabiopankki/en/index.php?lang=en">https://www.auriabiopankki.fi/auriapalvelut/auriabiopankki/en/index.php?lang=en</a>
Turku Clinical Research Centre 	Clinical informatics service responsible for maintaining the hospital district's patient information and making it available for research purposes.	Instructions for access published.	<a href="http://www.turkucrc.fi/en">http://www.turkucrc.fi/en</a>
Helsinki Biobank 	Helsinki Biobank provides samples with associated clinical information, for medical research and R&D purposes.	Instructions available for applying data.	<a href="https://www.terveyskyla.fi/helsinginbiopankki/en">https://www.terveyskyla.fi/helsinginbiopankki/en</a>
THL Biobank 	THL Biobank hosts a remarkable collection of population and disease-specific samples for research purposes.	Instructions and prices available for applying data.	<a href="https://thl.fi/en/web/thl-biobank/for-researchers">https://thl.fi/en/web/thl-biobank/for-researchers</a>
European Bio Informatics institute (EMBL-EBI) 	Shares data (e.g. nucleotide sequences) and tools originating from life science experiments.	Free open access (non-identifiable data). Access API provided.	<a href="https://www.ebi.ac.uk/">https://www.ebi.ac.uk/</a>
European Genome-phenome Archive (EGA) 	An EMBL-EBI service for permanent archiving and distribution of personally identifiable genetic and phenotypic data resulting from biomedical research projects.	Instructions for access published. Criteria for acceptance not listed.	<a href="https://ega-archive.org/access/data-access">https://ega-archive.org/access/data-access</a>
Elixir 	Brings together life science resources (including data, tools, cloud storage and supercomputing) from across Europe.	Free open access (non-identifiable data).	<a href="https://www.elixir-europe.org/">https://www.elixir-europe.org/</a>
China Kadoorie Biobank 	Population based biobank with questionnaire data and physical measurements, blood samples and health outcome data.	Instructions and criteria for access published.	<a href="http://www.ckbiobank.org/site/">http://www.ckbiobank.org/site/</a>

# VTT's PreMed project phases



- Establish understanding of existing companies (interviews)
  - Create initial ecosystem model
  - Identify initial business models and opportunities opened by data access
  - **Collect a group of interested companies to participate phases 2 & 3**
- Follow-up and keep project partners informed about on-going activities related to health data exploitation (FinnGen, Isaacus, data lakes, related international developments, ...)
  - Formulate data exploitation use cases based on the interests of participating companies
  - Execute a selected use case including: (1) research question formulation, (2) actions needed to access data (application, research plan, contract) and (3) results presentation and publication
  - Refine the initial ecosystem model and high-level simulation model resulting from phase 1
  - Implement a simulator to test various public strategies to support ecosystem creation
  - Provide recommendations for public bodies concerning financing and other measures

## PreMed tasks & leaders

- Task 1 / Stakeholder screening:  
Richard Fagerström, Principal Scientist
- Task 2 / Business enablers:  
Jaakko Lähteenmäki, Principal Scientist (project manager)
- Task 3 / System analysis:  
Peter Ylén, Principal Scientist
- Task 4 / Business strategy:  
Ville Valovirta, Senior Scientist
- Task 5 / AI opportunity:  
Mark van Gils, Principal Scientist, Principal Investigator