



Genoomika taristu arengusuundadest Euroopas ja Eestis

Hedi Peterson
Rakvere 19.10.2023



TARTU ÜLIKOOL
arvutiteaduse instituut



ERKI NOOL

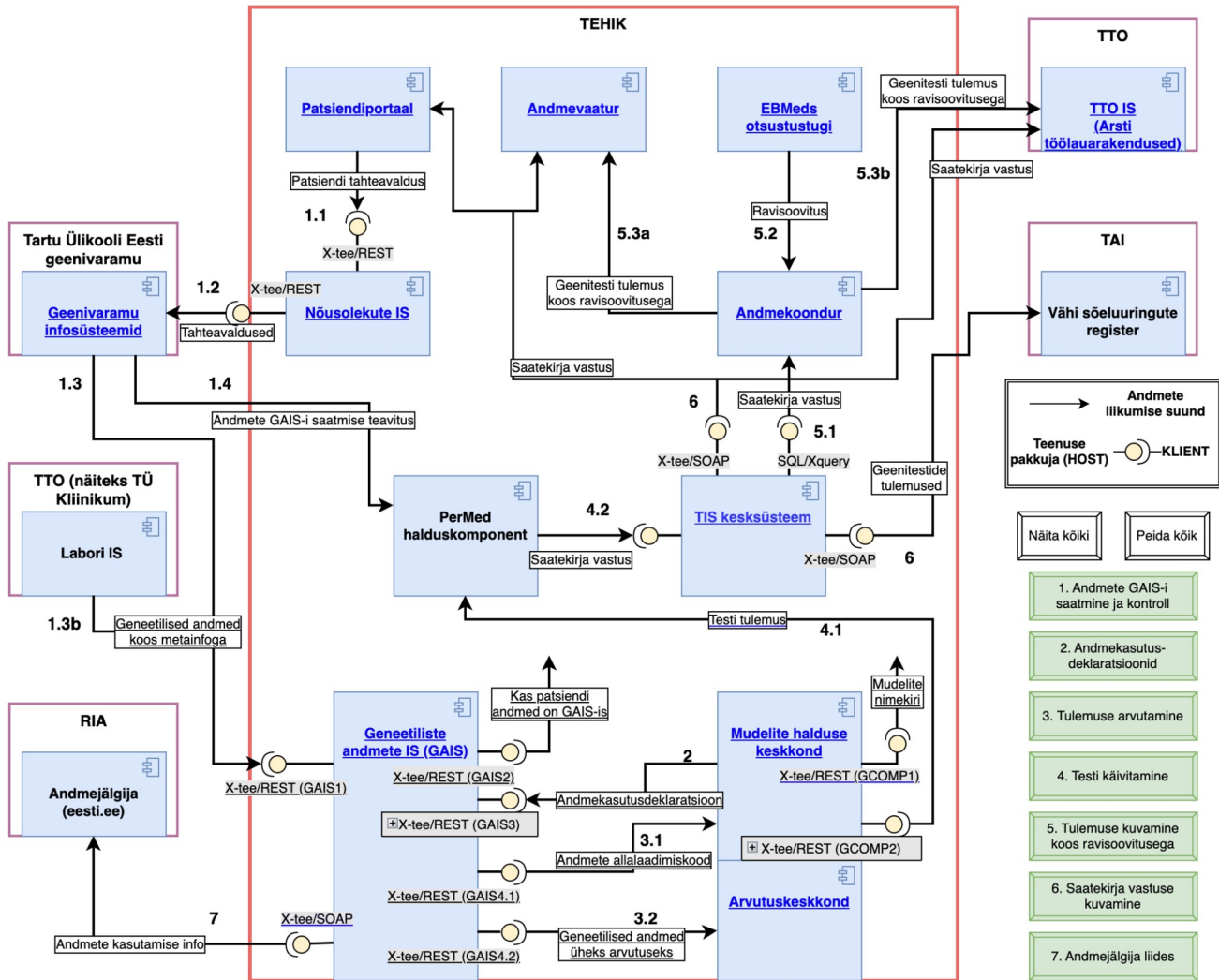
Sydney



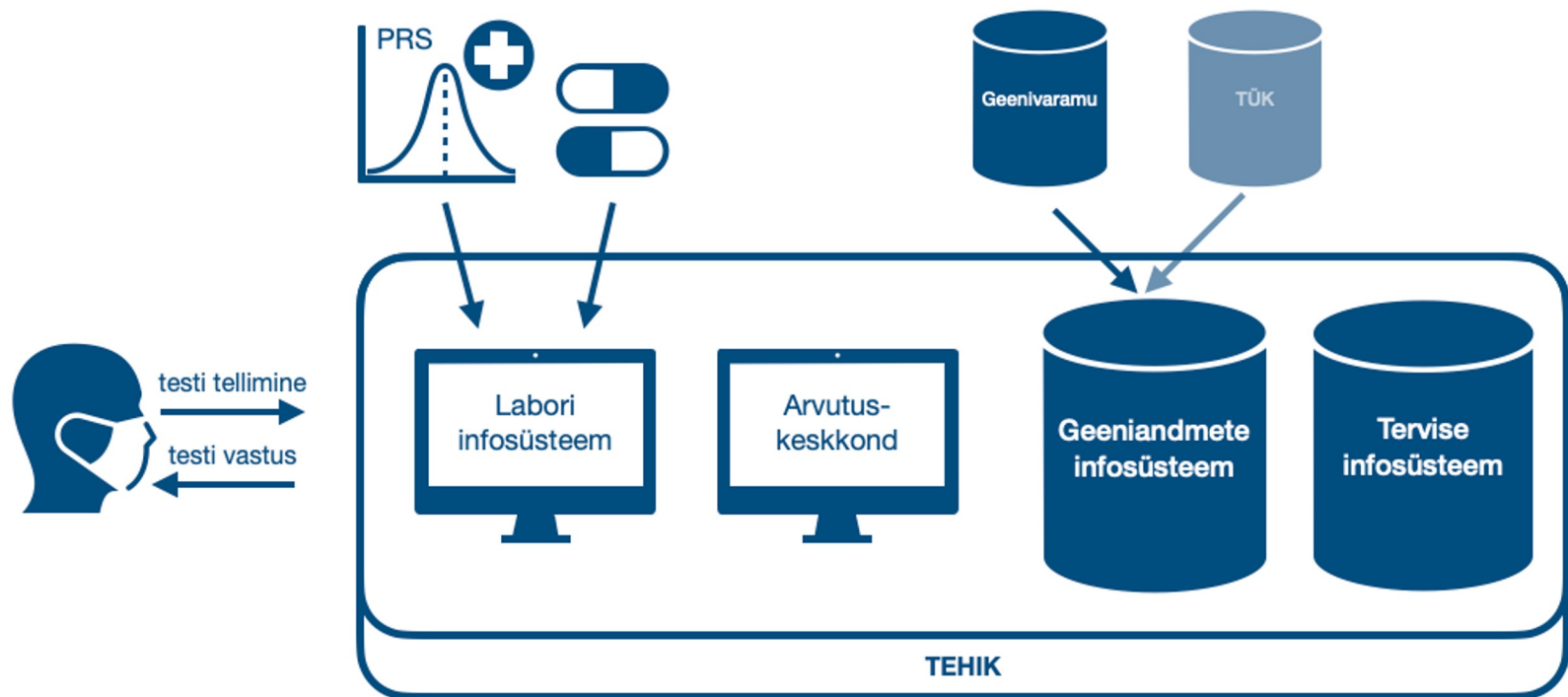
440
EESTI

SYDNEY 2000 DECATHLON GOLD MEDALLIST ERKI NOOL ESTONIA

PHOTO: J. K. / AP/WIDEWORLD

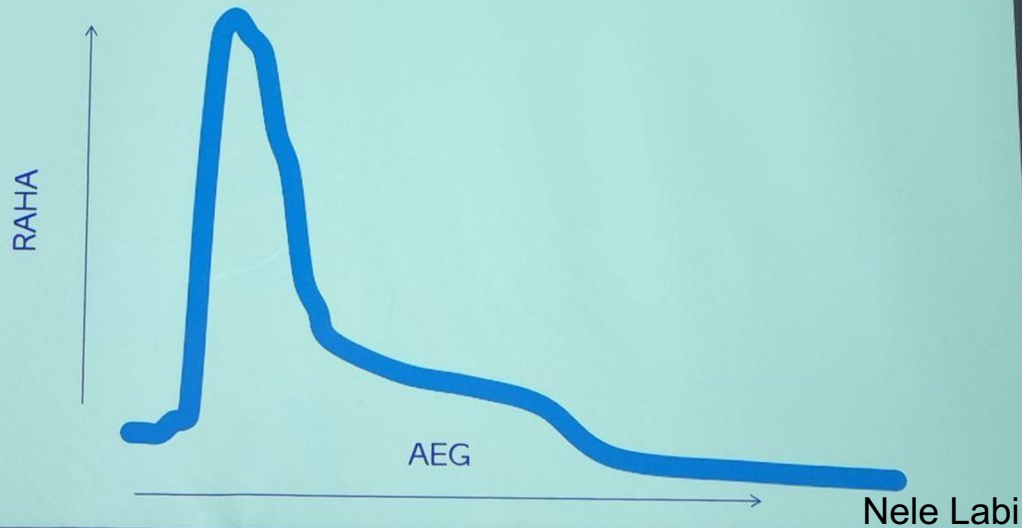


“Personaalmeditsiini rakendamine Eestis”



Suurimad väljakutsed

- 1) Andmete toomine teaduskasutusest kliinilisse kasutusse
- 2) Seadmete tootmine
- 3) Rahastamine (topelt väljakutse)



Parimad palad

Eesti saab esimeseks riigiks, kus on olemas süsteemselt teadusest kliinilisse andmete toomise praktika

Personaalmehitsiini kompetentsikeskuse loomine

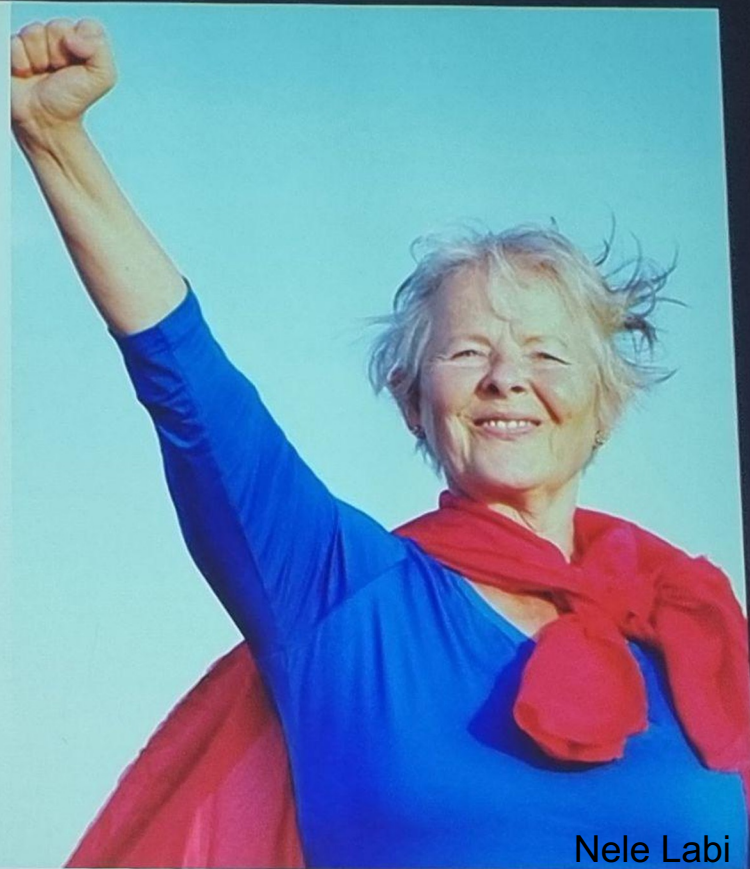
Rahvusvahelised projektid:

- 1+ Miljon Genoomi

- Genome of Europe

- Genomic Data Infrastructure

- European Partnership for Personalised Medicine



Nele Labi

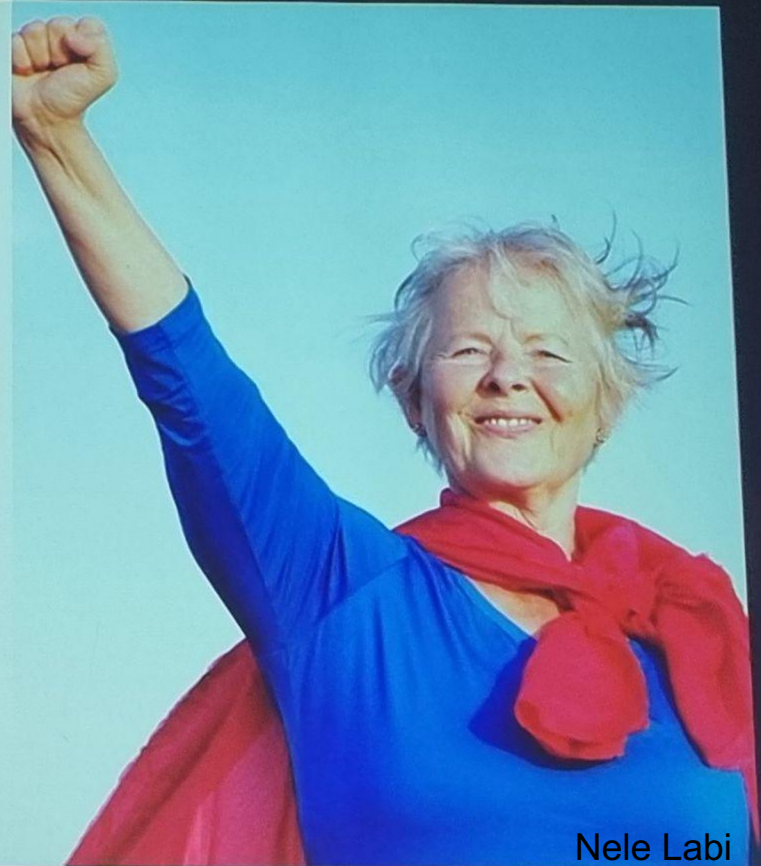
Parimad palad

Eesti saab esimeseks riigiks, kus on olemas süsteemselt teadusest kliinilisse andmete toomise praktika

Personaalmehitsiini kompetentsikeskuse loomine

Rahvusvahelised projektid:

- 1+ Miljon Genoomi
- Genome of Europe
- Genomic Data Infrastructure
- European Partnership for Personalised Medicine



Nele Labi

Suurbritannia

- Genomics England 100 000 genoomi
 - Patsiendid ja pereliikmed
 - Leiud tagasisidestatakse
- UK Biobank 500 000 osalejat
 - 40-69 aasta vanused, 2006-2010
 - 200 000 eksoomijärjestused
 - Pildiuuringus skaneeritakse aju, südant, luid ja kõhtu 100 000 osalejal
 - Isikuga seotud aga ind. tulemusi ei tagasisidestata
- Our Future Health
 - Fenotüüp ja elustiili info, lisaks SNP põhine genotüübiandmed
 - Varajane detekteerimine ja diagnostika, krooniliste haiguste ennetamine ja sekkumine

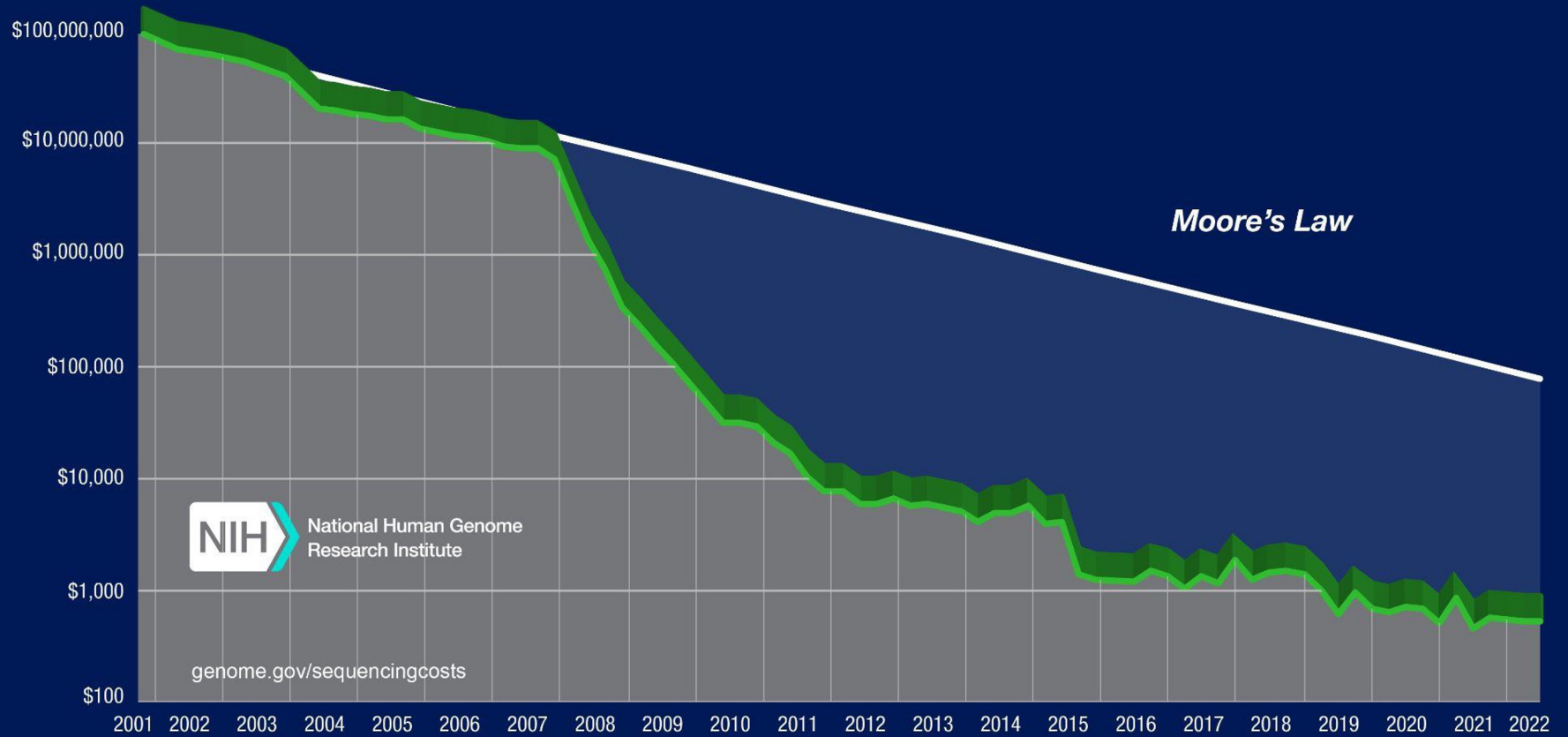
Järgmise viie aasta plaan:

- Uute kliiniliste näidustuste süstemaatiline juurutamine genoomse testimise jaoks
- Genoomitestide integreerimine kliinilisse ravisse
- Täppisravi kasutamise edendamine ja ravimite kasutamise optimeerimine genoomika kaudu
- Andmetaristu arendamine andmete jagamiseks ja kasutamiseks
- Käsikäes tipptasemele teaduse ja innovatsiooniga.

Genoomika hetkeseis Eestis

- TÜ EGVs on **200 000** geenidoonori **genotüübi andmed**
- **3000** geenidoonori **täisgenoomide** põhjal on loodud Eesti referentsgenoom
- Selle baasil on **genotüpiseeritud** proovid imputeeritud, mis on kuluefektiivne viis sagedasti esinevate polümorfismide analüüsiks sagedaste fenotüüpide puhul
- Täisgenoomide ja eksoomide sekveneerimine on lisaks oluline **harvikaiguste diagnoosimisel**, mida Eestis veab TÜ Kliinikum
- Kasvajate genoomide sekveneerimine tõusuteel
- Genoomika jõuab Tervisekassa kaudu inimeseni 2024

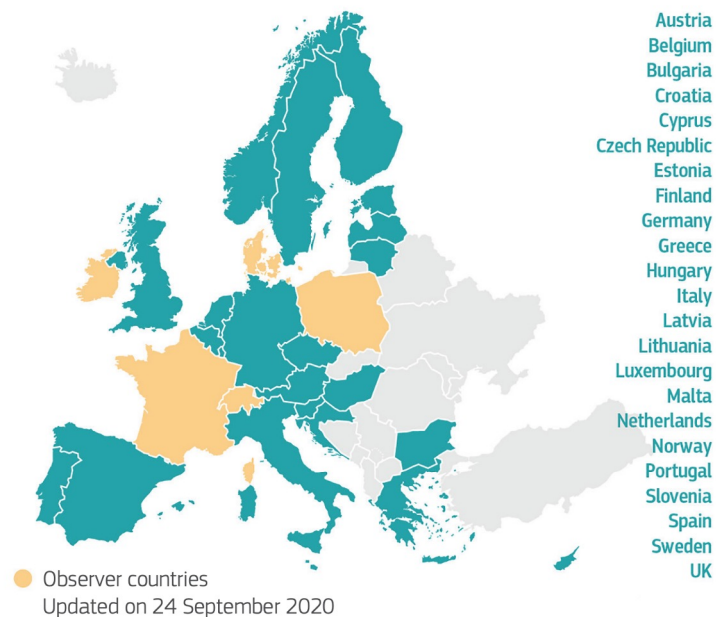
Cost per Human Genome



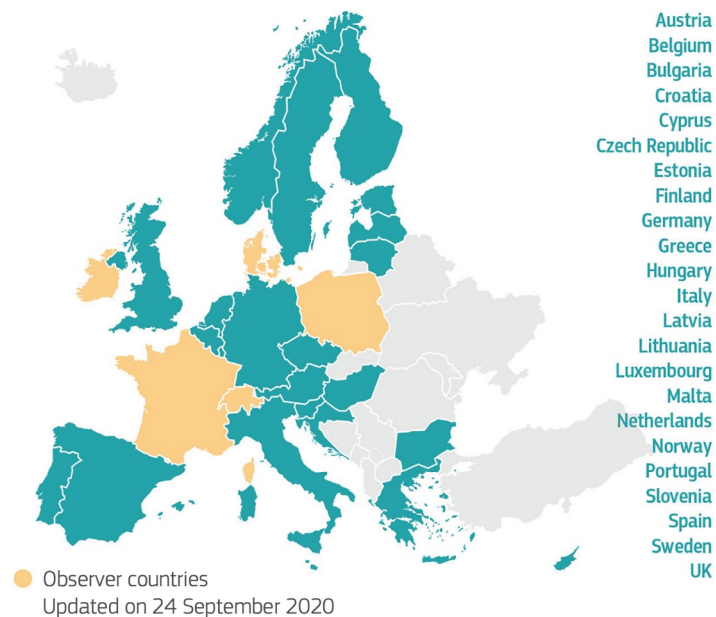
NIH National Human Genome Research Institute

genome.gov/sequencingcosts

- 1+MG deklaratsioon - aasta 2022 lõpuks miljon sekveneeritud genoomi
- Eesmärk on teha võimalikuks turvaline juurdepääs genoomiandmetele ja seotud kliinilistele fenotüübiandmetele kogu Euroopas üle riigipiiride.
- Seeläbi panustada paremate teadusuuringutesse, perosnaalsesse tervishoidu ja tervishoiupoliitika kujundamisse.



- 1+MG deklaratatsioon - aasta 2022 lõpuks miljon sekveneeritud genoomi
- B1MG: koordineerimine ja toetus (disain, soovitus, testimine) 2020-2023
- Genomic data infrastructure: arendus, juurutamine, toimimine, ülalpidamine. Nov 2022-2026

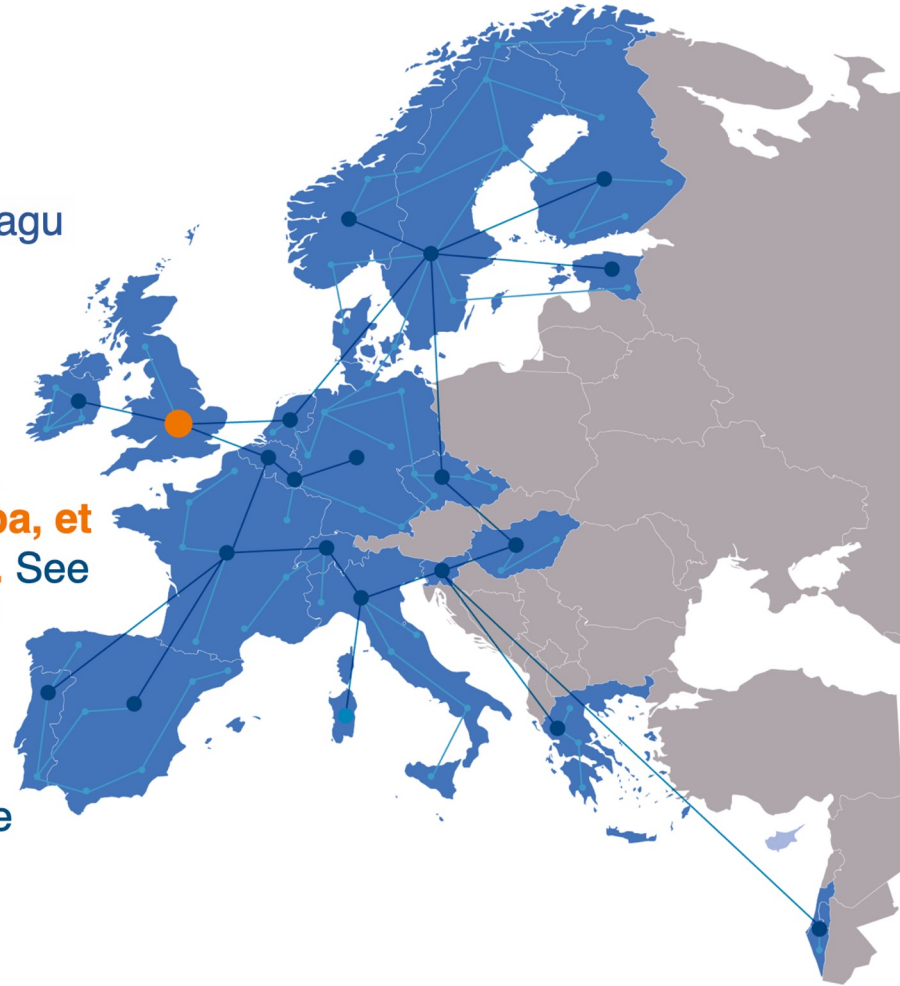


ELIXIR

ELIXIR on valitsustevaheline organisatsioon, mis koondab kogu Euroopast eluteaduste ressursse, nagu andmebaasid, tarkvara, koolitusmaterjalid, standardid ja arvutusressursid.

ELIXIRi eesmärk on **koordineerida eluteaduste ressursse üle Euroopa, et need moodustaksid ühtse taristu.** See muudab teadlaste jaoks lihtsamaks:

- Andmete leidmise ja jagamise
- Teadmishetuse
- Teadusuuringute parimate tavade kokkuleppimise





GDI numbrites

€40M (50% kaasrahastus)

45 projekti põhipartnerit

6 sidusüksust

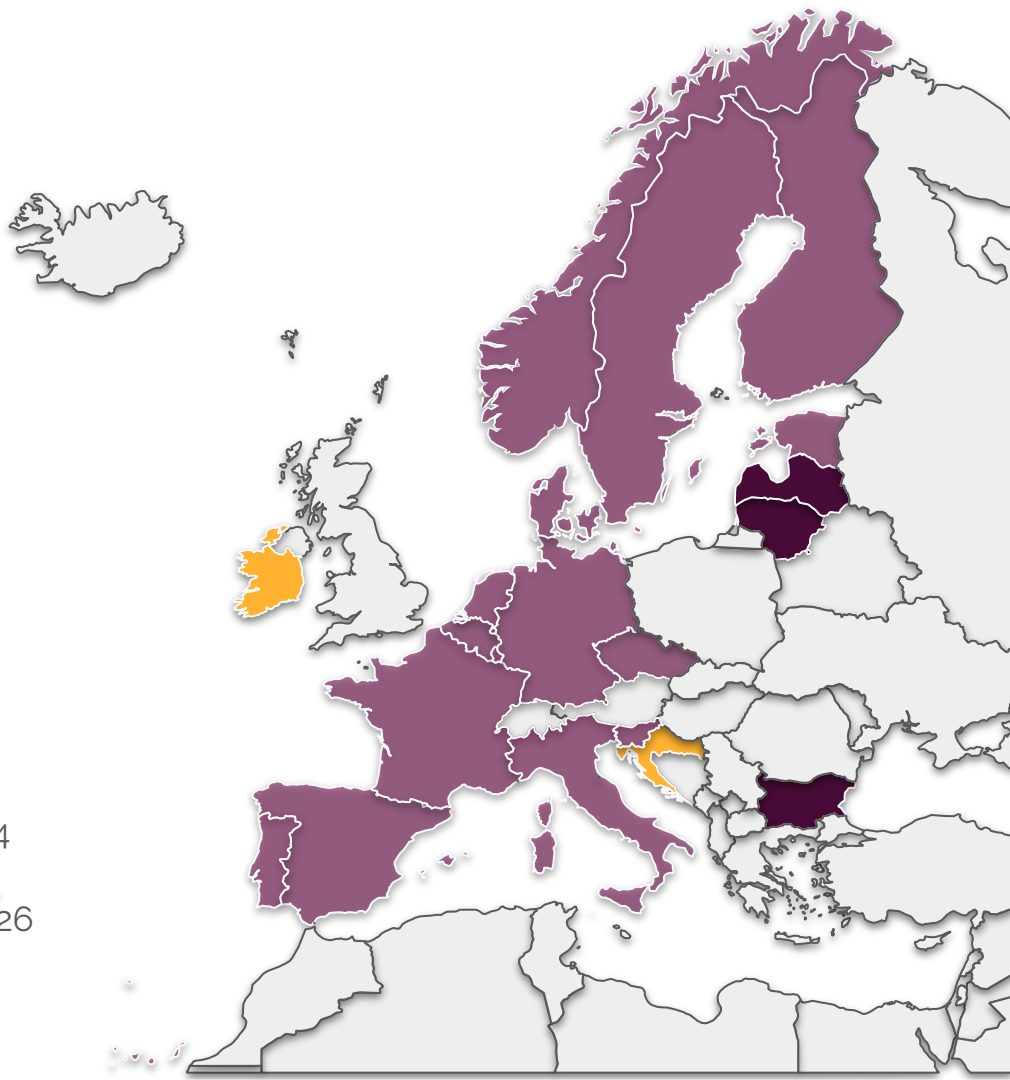
7 seotud partnerit

Kokku: **58** osapoolt

20+X Euroopa riiki

Vähemalt **6** riiki tehniliselt toimivaks aastal 2024

Vähemalt **15** riiki tehniliselt toimivaks aastal 2026

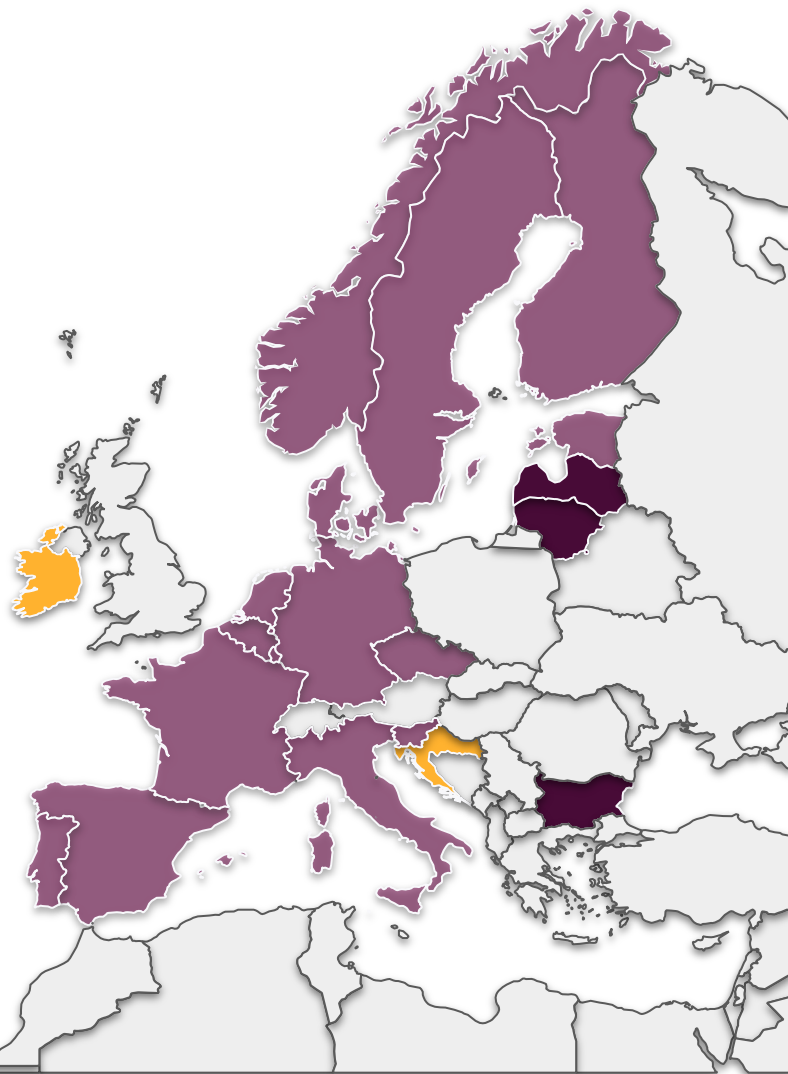




Riikide võetud kohustus GDI raames aastaks 2026



- Täielikult tegutsevad ja integreeritud ühtsesse 1+MG infrastruktuuri: **Belgia, Tšehhi Vabariik, Taani, Eesti, Soome, Prantsusmaa, Saksamaa, Itaalia, Luksemburg, Portugal, Sloveenia, Hispaania, Rootsi, Holland, Norra**
- Täielikult tegutsev riiklik *Node*, kuid mitte veel 1+MG infrastruktuuri integreeritud: **Bulgaaria, Läti, Leedu**
- Kaasamisfaasis: **Horvaatia, Iirimaa, Küpros, Ungari, Malta, Rumeenia**

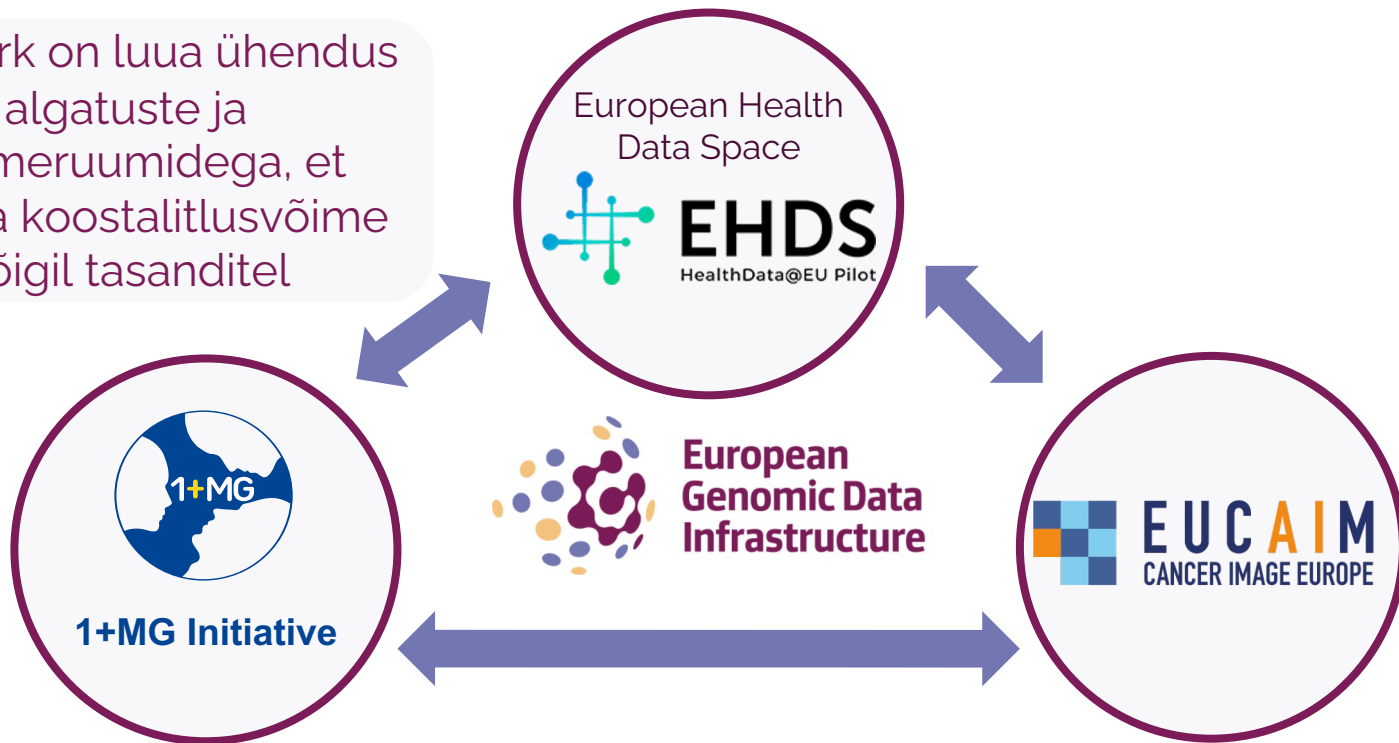


Infrastruktuur on olemas 2024. aasta lõpuks ja tehniline valmisolek on saavutatud vähemalt kuues riigis.



Erinevad üleeuroopalised initsiatiivid

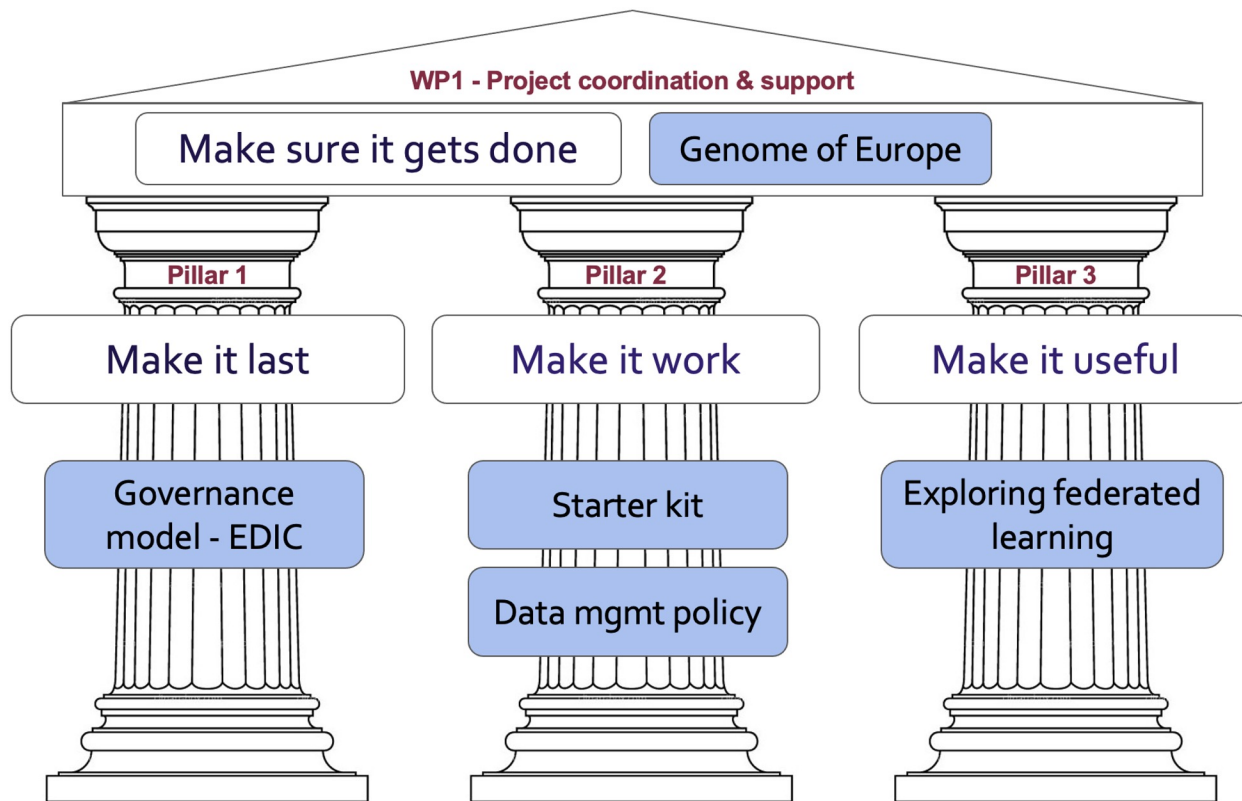
Eesmärk on luua ühendus algatuste ja andmeruumidega, et tagada koostalitlusvõime kõigil tasanditel



Euroopa partnerlussuhted: personaliseeritud meditsiin, haruldased haigused; Digitaalsed kaksikud, vähimissioon



GDI toetub kolmele sambale





Pillar I: Pikaajaline järjepidevus



Tagada infrastruktuuri toimimine pärast GDI projekti lõppu ja teha andmed kättesaadavaks kogu Euroopas

Seni tehtu:

Loomisel on European Digital Infrastructure Consortium (EDIC)



SOTSIAALMINISTEERIUM





Pillar II: Taristu arendus



Eesmärk on ellu viia 1+MG taristu: tagades, et Euroopa andmeressursid töötaksid osana osana 1+MG taristust

Seni tehtu:

Esmane stardikomplekt on olemas. Rakendused on juurutamiseks saadaval;

Esimesed GDI-Node-del on olemas taristu, mis kasutab 3000 realistlikku sünteetilist genoomi ja nendega seotud fenotüübiandmeid;

Riikide tehnoloogiline valmisolek on kaardistatud

Koostatud on ELSI parimate tavasid järgiv andmehalduspoliitika projekt





Pillar III: Kasutusjuhud



Pakkuda uuenduslikke lahendusi arstidele, teadlastele ja innovaatoritele

Seni tehtu:

Kasutusjuhtude esmased vajadused on kaardistatud

Esialgne analüüs hajutatud õppemudelite valideerimiseks on määratletud





GDI kasutusjuhud- Genome of Europe

Genome of Europe raames sekveneeritavate genoomide eesmärkideks on:

- i) üksikute geneetiliste variantide otsimine
- ii) polügeensete riskiskooride ümberkalibreerimine
- iii) põlvnemisest lähtuv imputatsioon





Genomic Data Infrastructure

20 1+MG allakirjutanud riiki GDIs on võtnud eesmärgiks, et:

- Iga riik seab üles Andmekeskuse või *Node*
- Iga riik haldab oma andmeid ise (sealhulgas kohalikke andmekeskuseid)
- Andmekeskused pakuvad piirideülest andmeanalüüsi kasutades ühtset andmeraamistikku (standardeid ja APIsid)
- Andmetaristu viis peamist funktsiooni on



Andmete
avastamine



Juurdepääsu
haldamise
tööriistad



Andmetöötlus

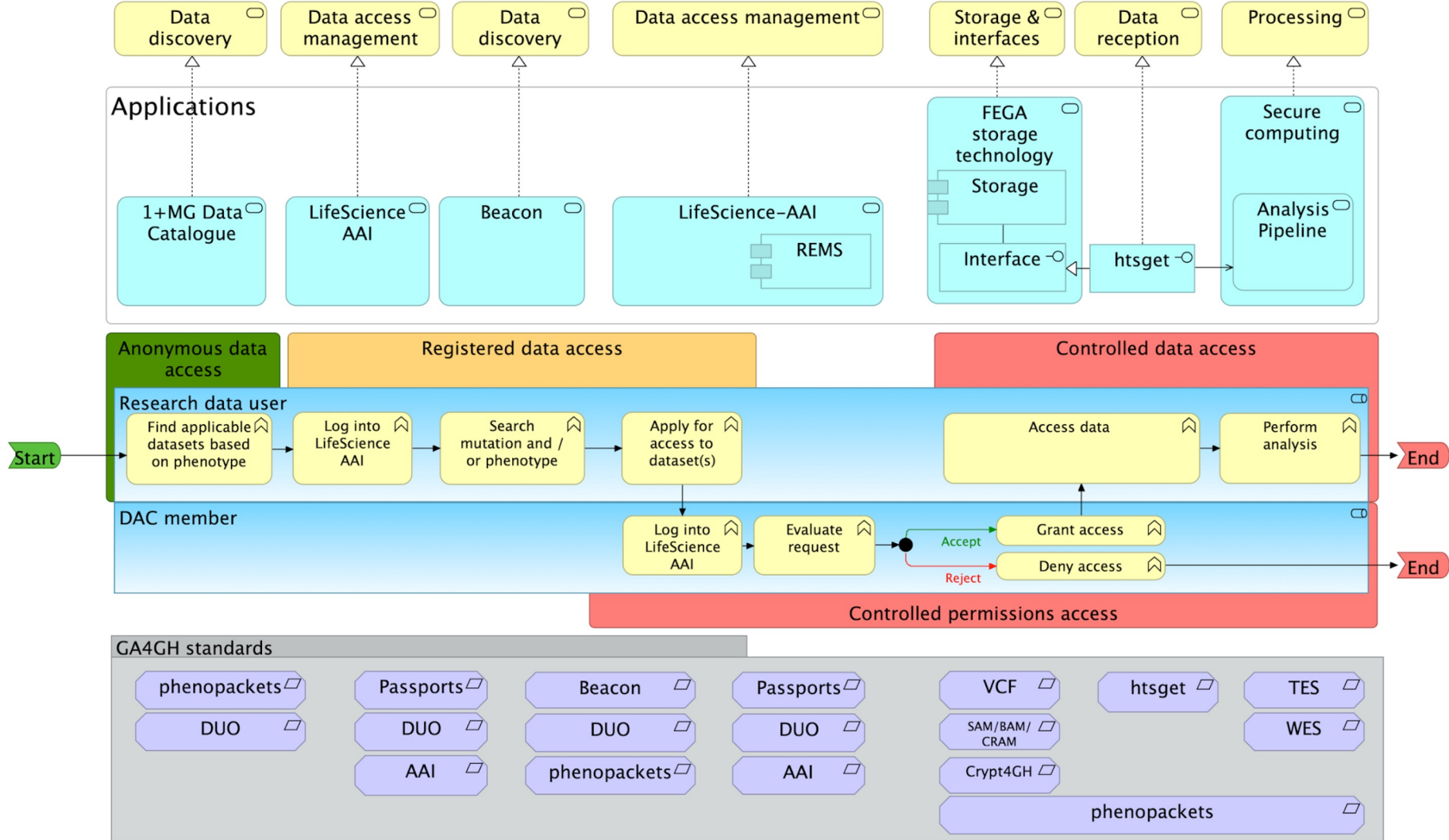


Andmete
vastuvõtt



Salvestus ja
liidesed







DisGeNET



THL Biobank

Kas konkreetne
alleel eksisteerib
kindlas
genoomses
asukohas?



TÜ Geenivaramu

Elixir Beacon - Estonia



Query available datasets

Select a dataset: EGCUT2535 (variants)

Assembly GRCh37

Chromosome 10

Position 60221

Reference Bases T

Alternate Bases A

Find

Beacon search history

Clear results

| Dataset | Asm | Chr | Pos | Ref | Alt | Exists |
|-----------|--------|-----|-------|-----|-----|--------|
| EGCUT2535 | GRCh37 | 10 | 60221 | T | A | true |

| Dataset ID | Short title | Access type |
|------------|--------------------------|-------------|
| EGCUT2535 | ELIXIR Estonia EGCUT2535 | PUBLIC |

Total Items: 1

Beacon v2



- Registreeritud kasutaja avastab variandi ELIXIR Beacon Networki abil andmekeskusest #1
- Taotleb juurdepääsu 1+MG andmetele REMS-i kaudu
- Vastutav töötleja / DAC annab andmetele juurdepääsu
- Logib LifeScience AAI abil sisse andmekeskuse 1 GPAP-i
- Kasutab Beacon V2, et teha päringuid teistest andmekeskustest tõeväärtuse vastuse tasemel, et leida teisi sarnase fenotüübi ja genotüübiga isikuid
- Leiab sarnase näidise andmekeskusest #2
- Kasutaja logib sisse andmekeskuse 2 GPAP-i ja visualiseerib htsgeti kaudu genoomijärjestused



GDI stardipakett

1+MG
5 FUNCTIONALITIES



Andmete
avastamine



Juurdepääsu
haldamise
tööriistad



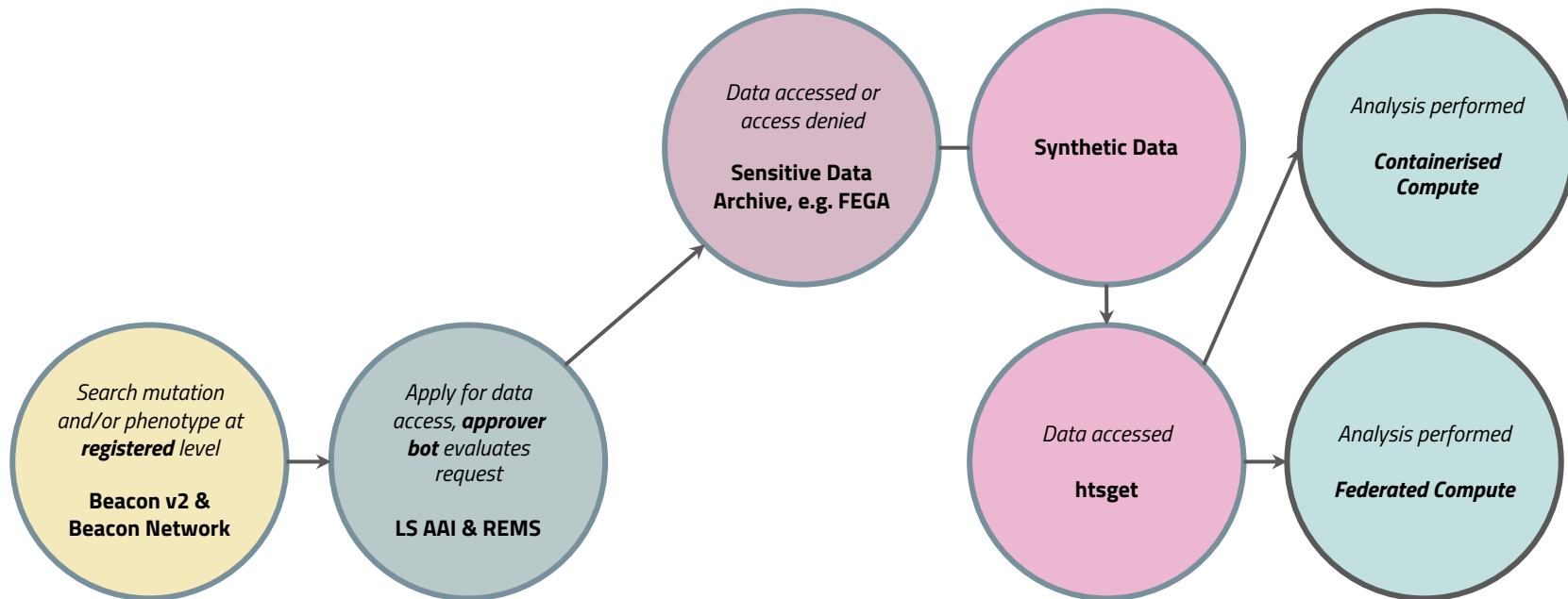
Salvestus ja
liidesed



Andmete
vastuvõtt



Andmetöötlus



STARTER KIT



GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.



GDI kasutajaloo tervikpilt

1+MG
5 FUNCTIONALITIES



Andmete
avastamine



Juurdepääsu
haldamise
tööriistad



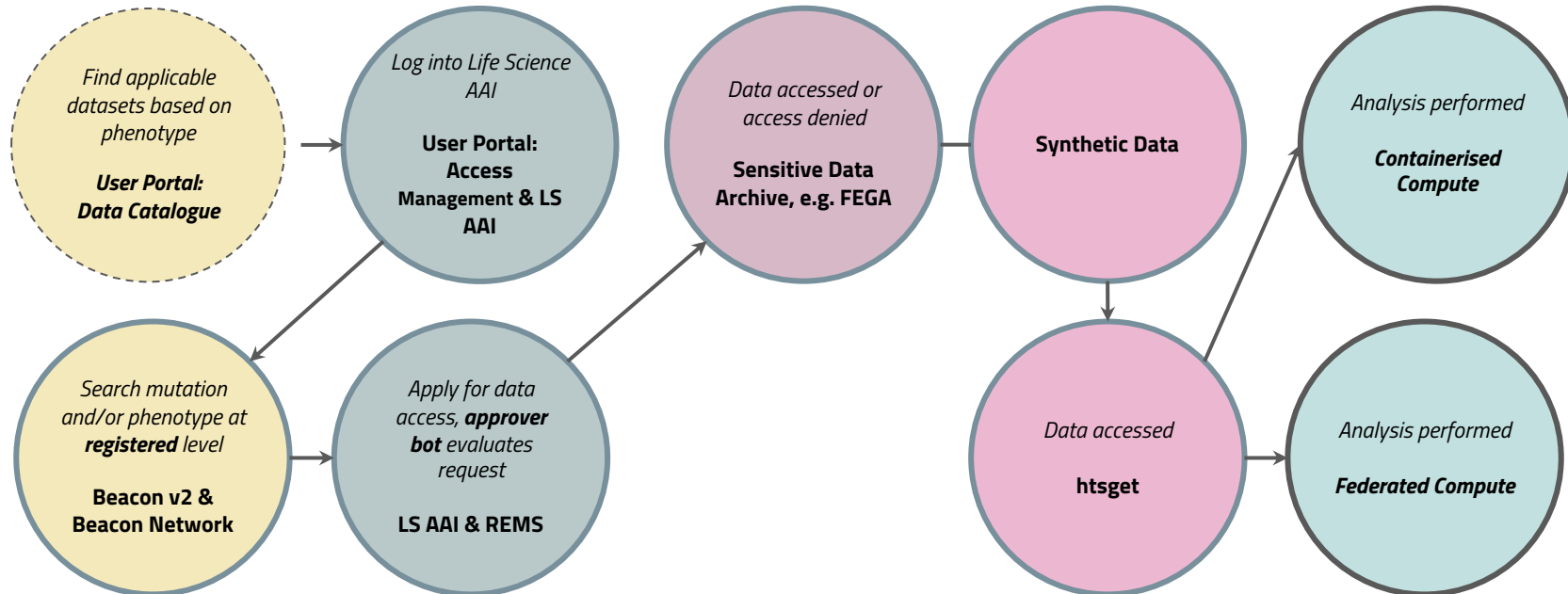
Salvestus ja
liidesed



Andmete
vastuvõtt



Andmetöötlus



STARTER KIT



GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.



GDI "Starter Kit" = Esimesed tarkvarapakettide demod 2023

| Product | Outline | Prod | PO | Function |
|----------------------------------|--|------|-------------|----------|
| Sensitive Data Archive | Securely stores data | ✓ | Sweden | |
| LifeScience AAI | Provides a federated Identity | ✓ | Czechia | |
| REMS | Allows data access applications and decisions | ✓ | Finland | |
| Beacon | Genetic and phenotypic data discovery | ✓ | Spain | |
| Beacon Network | Federated network of Beacons | ✓ | Finland | |
| Synthetic Data | Artificial anonymous data | | Finland | |
| htsget | Secure genetic data distribution | | Sweden | |
| Containerised Computation | Computation via virtualised portable software packages | ✓ | Czechia | |
| Federated Computation | Interoperable distributed workflows | | | |
| Packaging and Deployment | Packaging and deployment of the starter kit | | Spain | |
| User Portal – Data Catalogue | European level catalogue of data within nodes | | Netherlands | |
| User Portal – Access management | European level application / access management | | Luxembourg | |

European level



GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.



Global Alliance
for Genomics & Health

<https://github.com/GenomicDataInfrastructure>

Hetkeseis ja lähitulevik

- Genoomide järjestamine muutub igapäevaseks
- Andmemahud on suured
- Andmete nähtavaks tegemine on oluline
- Andmepuur turvalisteks arvutusteks on vajalik
- Geenandmed TIS osaks, et jõuaks arsti ja patsiendini
- Metaandmed on kriitilised - fenotüübid, kriitiline komponent teaduses ja ravis

Kõik panustavad!

- TÜ arvutiteaduse instituut <https://www.cs.ut.ee/>
 - Jaak Vilo, Sulev Reisberg, GenMed tiim (GAIS!)
- ETAIS TÜ HPC <https://etais.ee/> (SAPU!)
 - Ulvi Talas, Ott-Eric Oopkaup, Vladislav Tuzov, Sander Kuusemets, Tommy Tomson
- TÜ Eesti Geenivaramu <https://genomics.ut.ee/et/eesti-geenivaramu> (TÄISGENOOMID)
 - Andres Metspalu, Mait Metspalu, Reedik Mägi, Lili Milani, Kristjan Metsalu, Priit Kleemann
- TÜ Kliinikum Geneetika ja personaalmeditsiini kliinik; Harvikaiguste kompetentsikeskus <https://www.kliinikum.ee/geneetikakeskus> <https://www.kliinikum.ee/harvikaigused/> (TÄISGENOOMID)
 - Sander Pajusalu, Katrin Õunap
- SoM ja TEHIK (PIKK PLAAN JA TARISTU)
 - Raili Sillart, Artur Novek
- ELIXIR <https://elixir-europe.org/> (RAHVUSVAHELISED PARTNERID)





TARTU ÜLIKOOL
arvutiteaduse instituut



TARTU ÜLIKOOL
genoomika instituut



SOTSIAALMINISTEERIUM

