



Microsoft UK for Precision Medicine and Genomics

Alessandro Riccombeni MBA, Ph.D.
National Genomics Officer, Microsoft UK

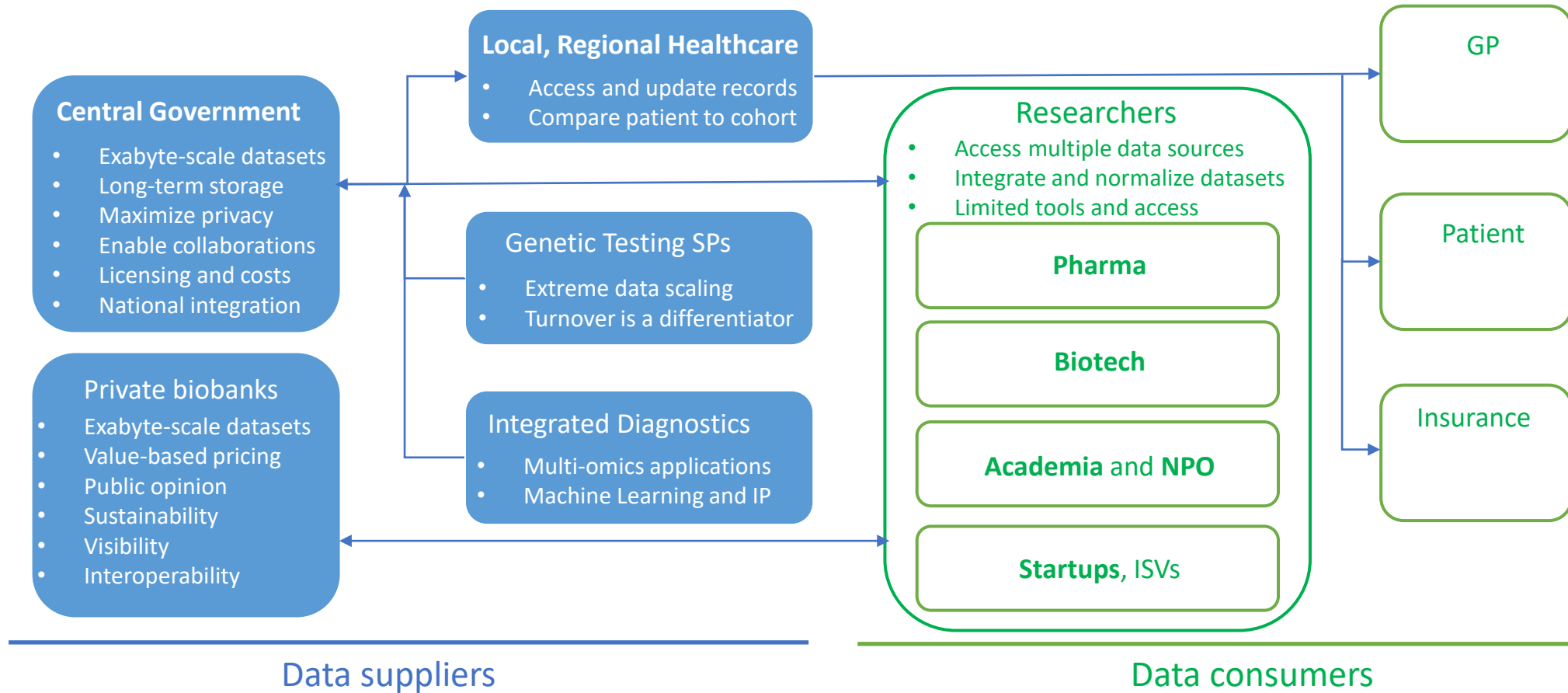
05.09.2022



Genomic Precision Medicine Initiatives



The Precision Medicine Landscape



The Cost of Genome Acquisition

<https://enlightenbio.com/news-and-features/2021/10/25/enabling-precision-medicine-life-sciences-in-2030/>

Whole Human Genomes	Overall basic data production cost at \$1000 per genome (\$M)	Basic storage footprint (PB)* *150GB per genome	Basic storage cost (\$10 per genome/year) (\$M)	Basic storage cost for the first decade (\$M)	Footprint with 2x data multiplier from researchers (PB)	Basic storage cost for the first decade, at 3x (\$M)	Footprint with 5x data multiplier from researchers (PB)	Basic storage cost for the first decade, at 5x (\$M)
5000	5	0.72	0.05	0.50	2.15	1.50	3.58	2.50
50000	50	7.15	0.50	5.00	21.46	15.00	35.76	25.00
100,000	100	14.31	1.00	10.00	42.92	30.00	71.53	50.00
500,000	500	71.53	5.00	50.00	214.58	150.00	357.63	250.00
1,000,000	1000	143.05	10.00	100.00	429.15	300.00	715.26	500.00
100,000,000	100000	14,305.11	1,000.00	10,000.00	42,915.34	30,000.00	71,525.57	50,000.00
334,000,000	334000	47,779.08	3,340.00	33,400.00	143,337.25	100,200.00	238,895.42	167,000.00
448,000,000	448000	64,086.91	4,480.00	44,800.00	192,260.74	134,400.00	320,434.57	224,000.00

Table 1: Examples of production and storage costs of large-scale, population genomics initiatives.



TRE Solution Accelerator

In development...

Today | TRE OSS capability

Self-service for administrators

Self-service for research teams

Agnostic
(tooling / data platform)

Extensible architecture

Desktops:
Windows, Linux

AzureML
inc Jupyter, R
Studio

InnerEye, ML
Flow, Gitea

Commercial
flexibility

Integrated
security (AAD /
CA)

Data exfiltration
control

Package and
repository
mirroring

Airlock

Cost reporting

What customers want to see...

PII scanning for
Airlock

Teams:
workspace
integration

Teams:
notifications /
approvals

Cataloguing
(Mauro, Purview)

Secure data
linking

Data sharing
across TREs

FHIR,
OMOP/OHDSI
Atlas

Cogstack,
OpenSafely

Cohort Selection
-> Data Access
request -> TRE

Databricks, Hue,
CKAN,
CycleCloud

Eyes Off/Remote
Task Execution

MLOps