



RWE04:

What we have observed

A summary of observations and some additional opportunities

Peter Grolimund

Nov-2018

Assumptions

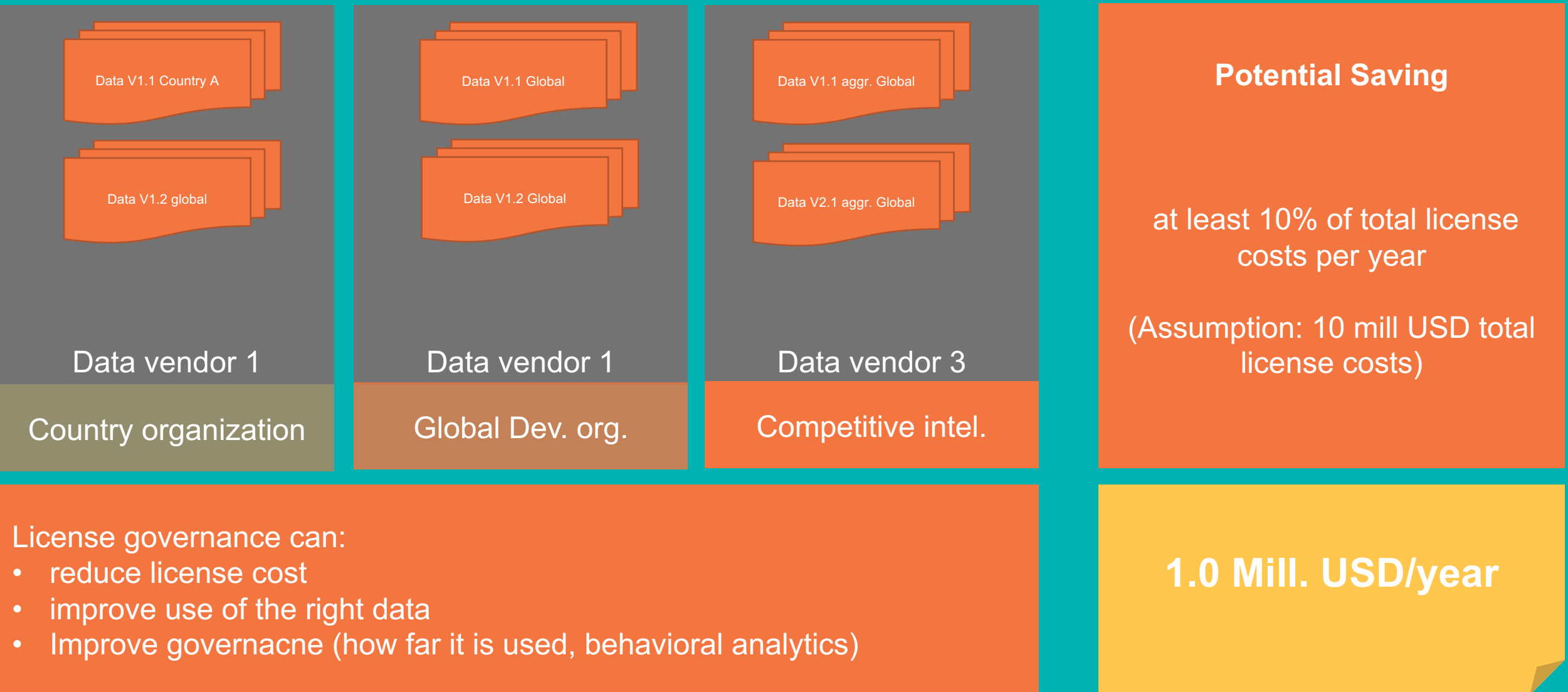
- **There is a need for an analytical system where ever it will be deployed and it needs to be very flexible and scalable**
- **There is a tremendous value using commercial RWE**
- **There is additional important value using public RWE**
- **The better integration of internal and external data, the easier and faster the opportunities can be leveraged**
- **There is value in collaboration with externals**
- **There is value to go with a patient centric approaches (legal aspects)**
- **There is an increasing need to use the data in different ways (analytical ecosystem, operational AI/ML)**

We only look at that stage for the savings...

The big basket..

The opportunities is a different story

License management



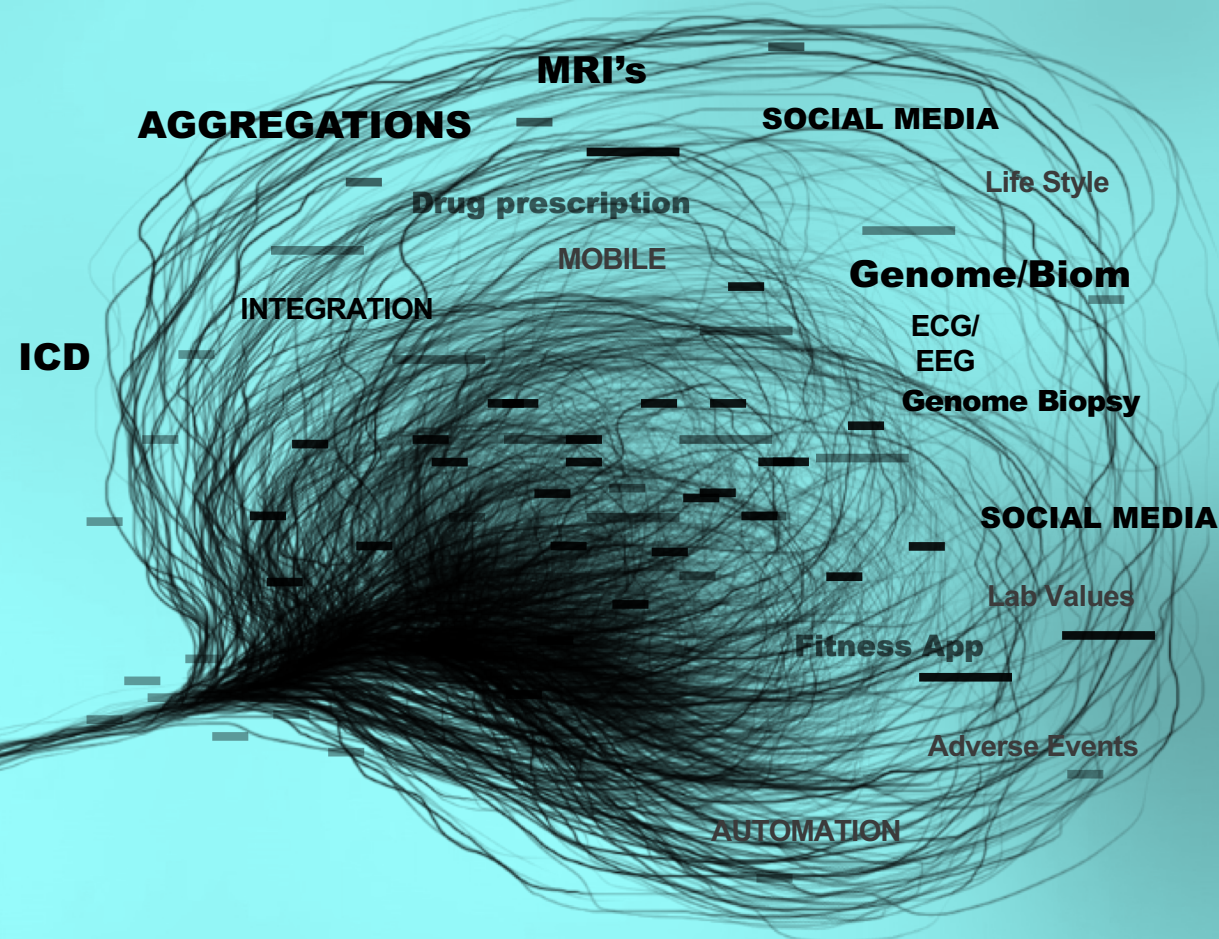
(RWE) Data Science & Analytics

90%

TIME WASTED SIFTING
THROUGH DATA

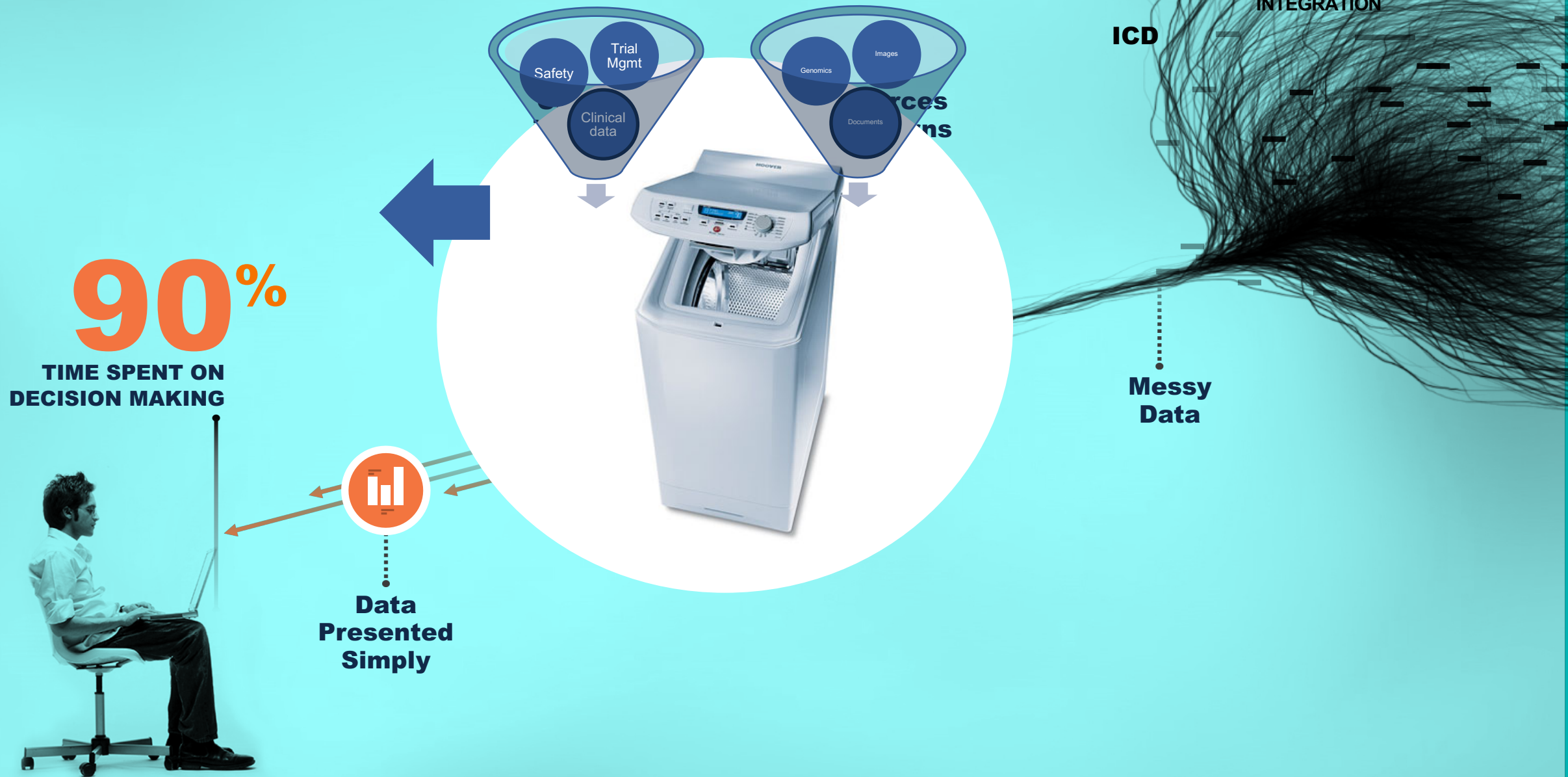


Messy
Data



What's the best approach for facilitating data-
access, usage, .. etc.?

Data Driven Operational Analytics



Open 'public' Data integration

FDA FEARS

Country organization

CT.GOV
FDA FEARS

Global Dev. org.

Demography data
EUROSTAT

Global B aggr. Global

Competitive intel.

Potential Saving

10 public data sources are used
in at least 5 units

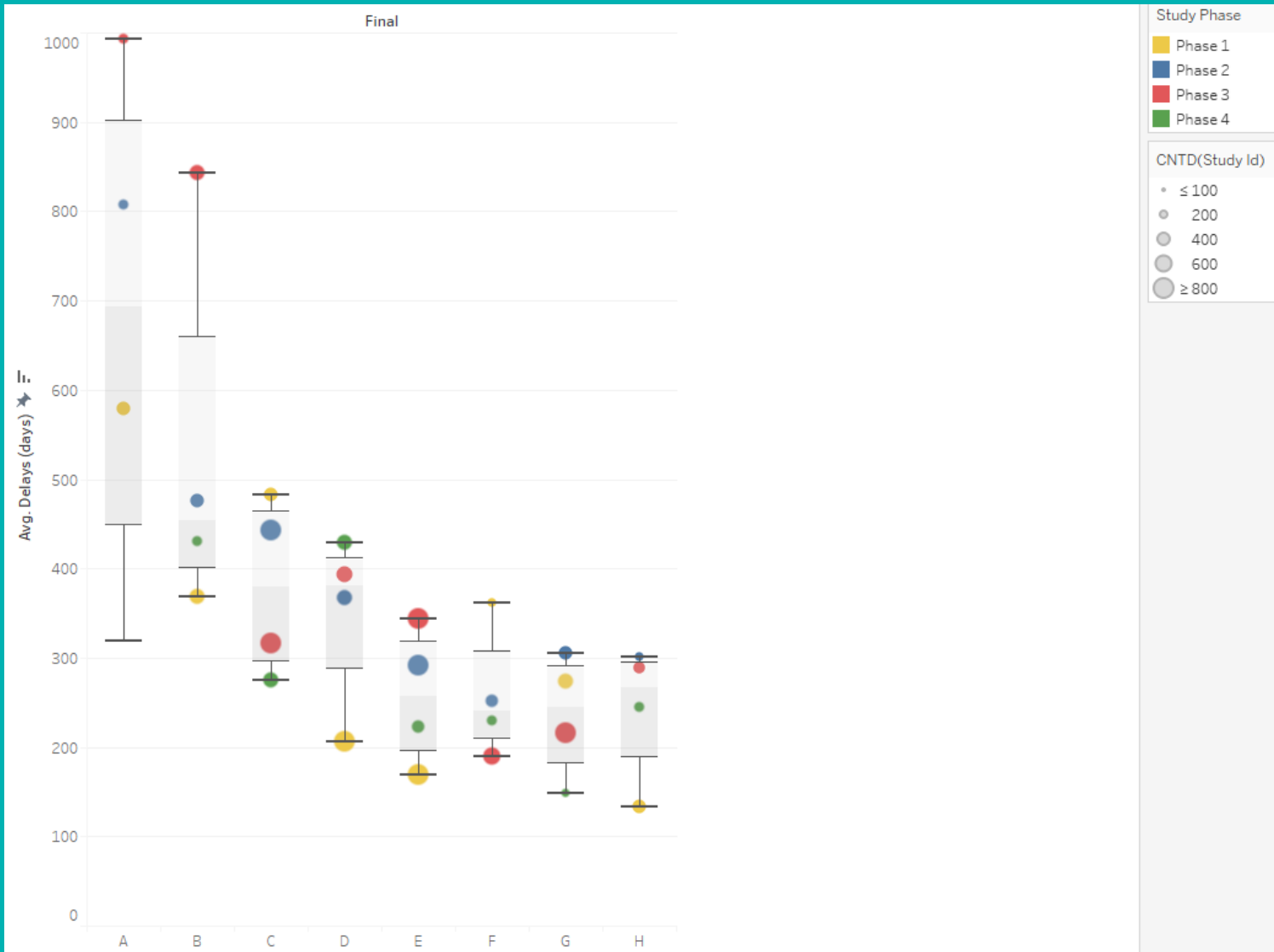
(Assumption: 1 FTE for 5
sources)
(savings FTE, storage etc.)

Governance can

- reduce integration, cleansing effort
- Guarantees actualization/operation
- Enables 'time-series' and hence new insights

1.5 Mill. USD/year

Open source and changes made in the planning (temporal)



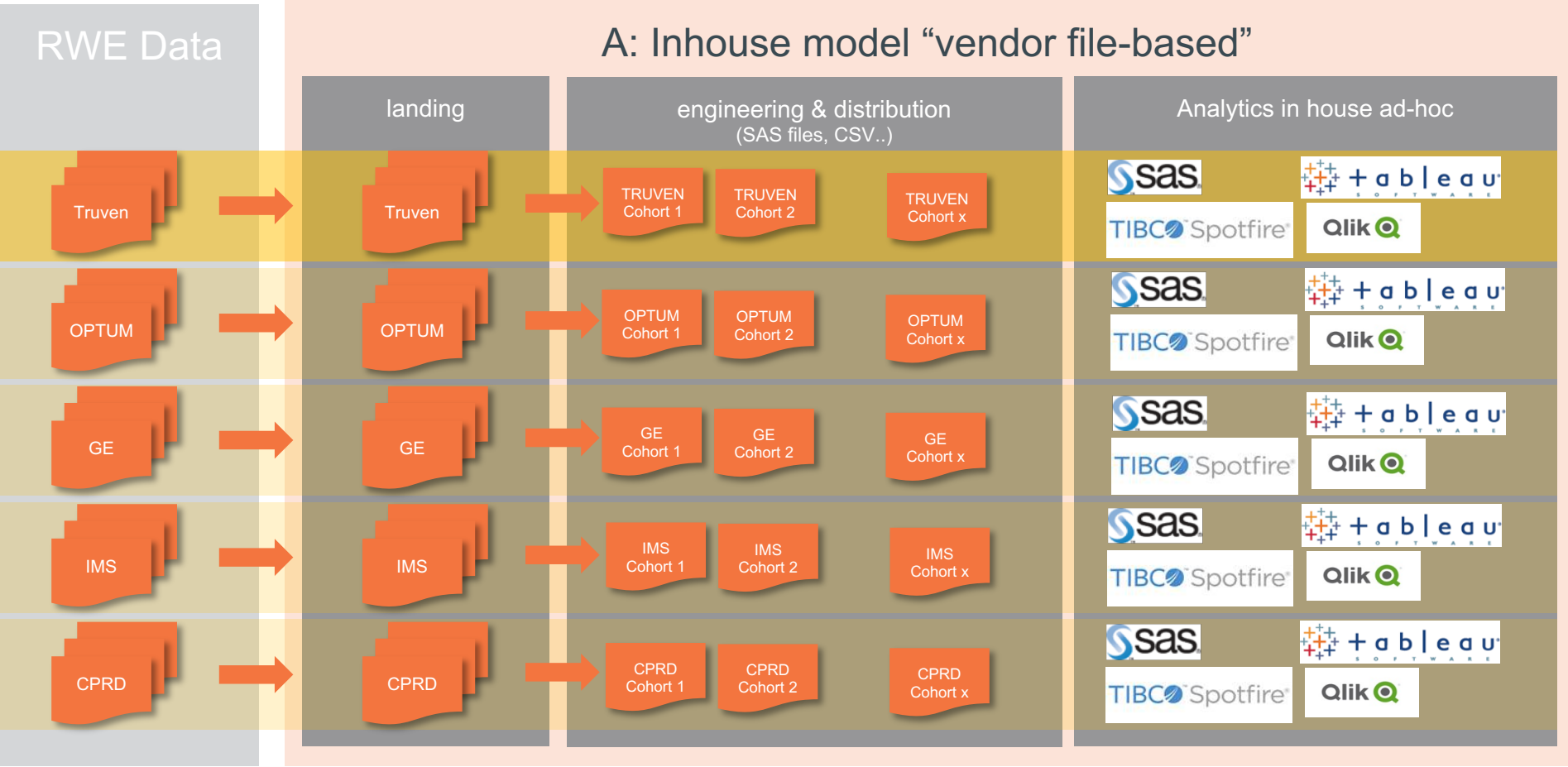
Explanation:

Changes in the study duration based on changed start and ending-dates over the time

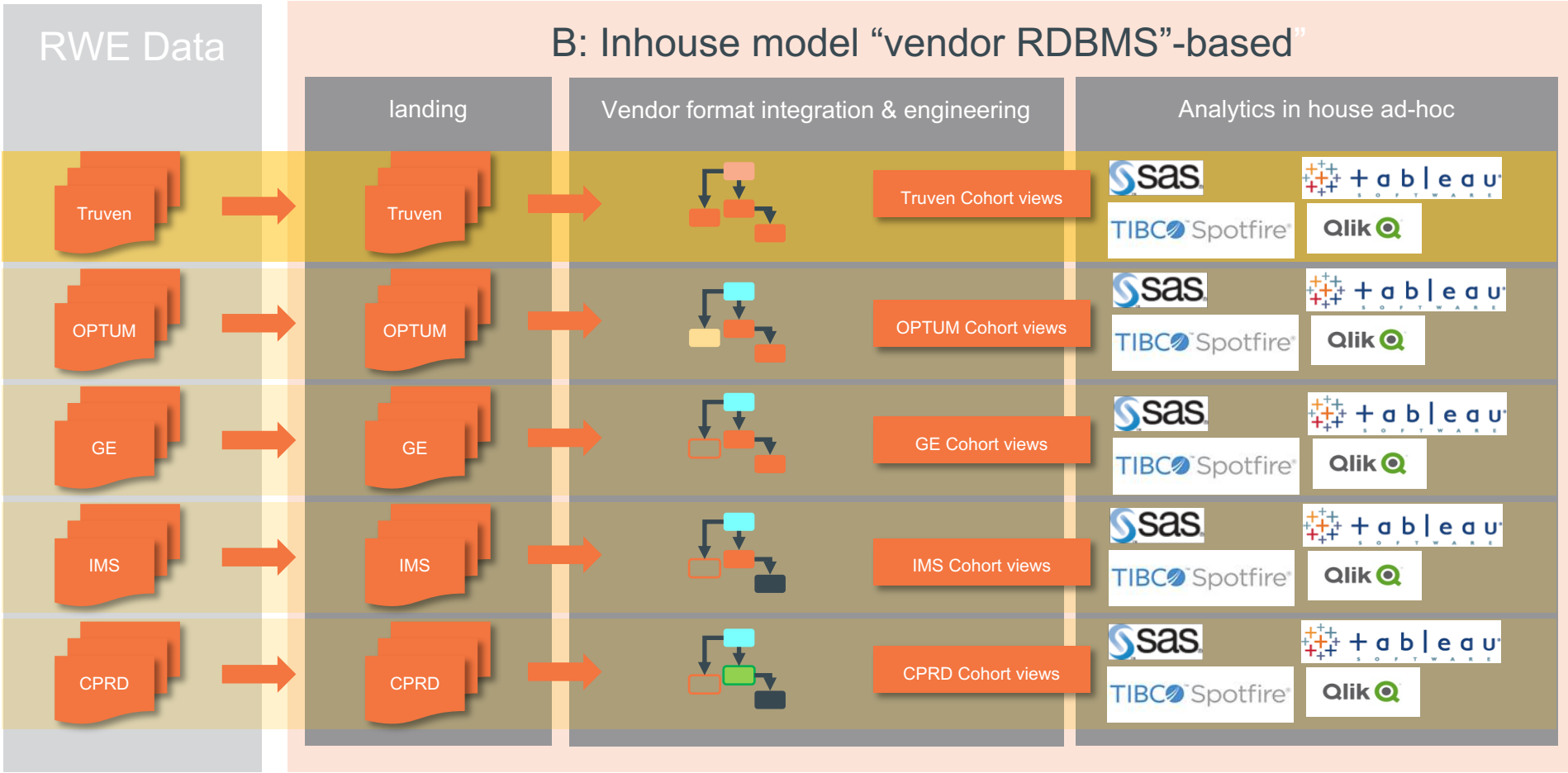
- By Phase (1-4) (colors)
- By Company (A-H)

P.S. size of the colored dots do respond to the number of studies (counts)

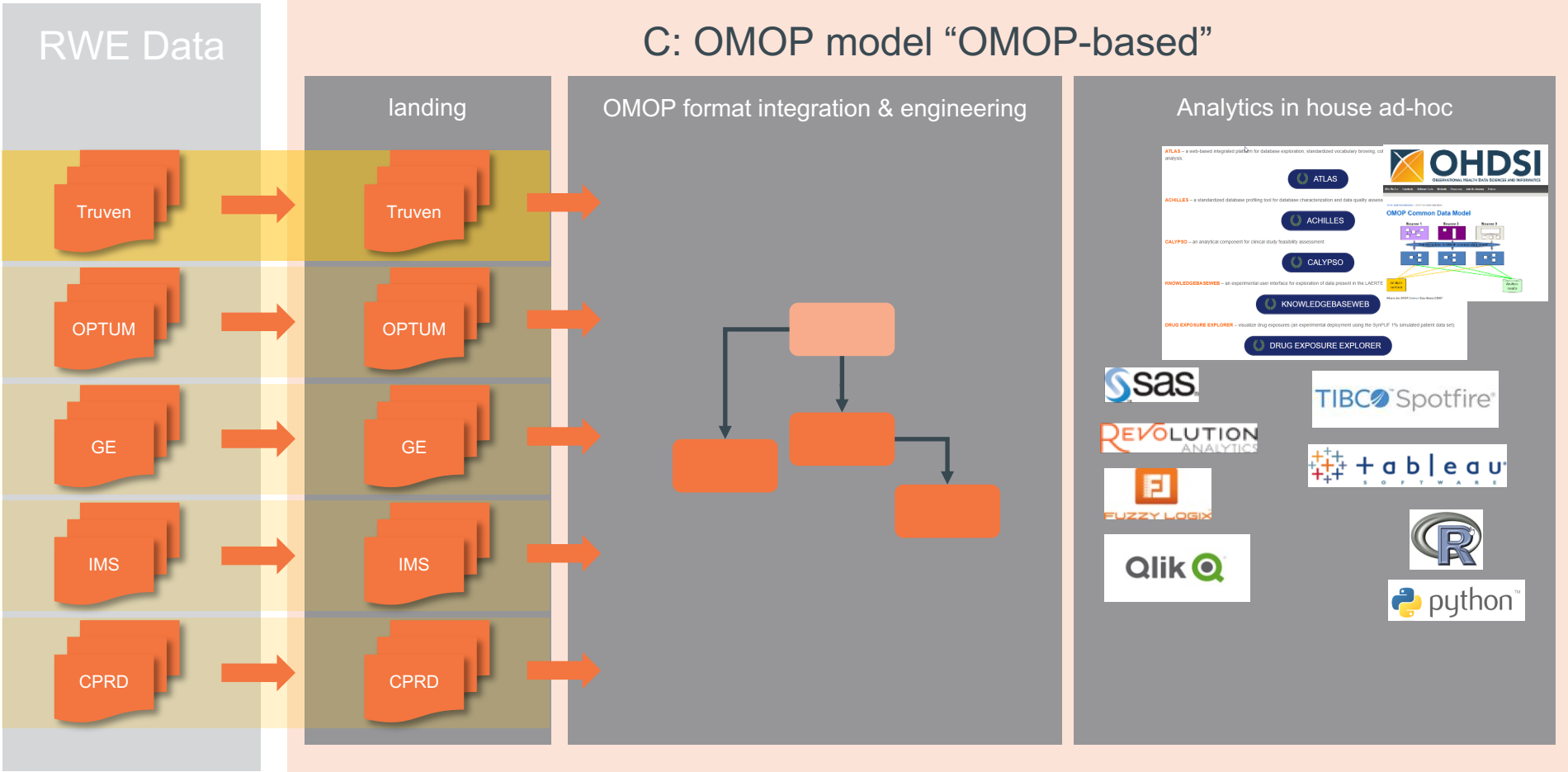
A: Traditional setting



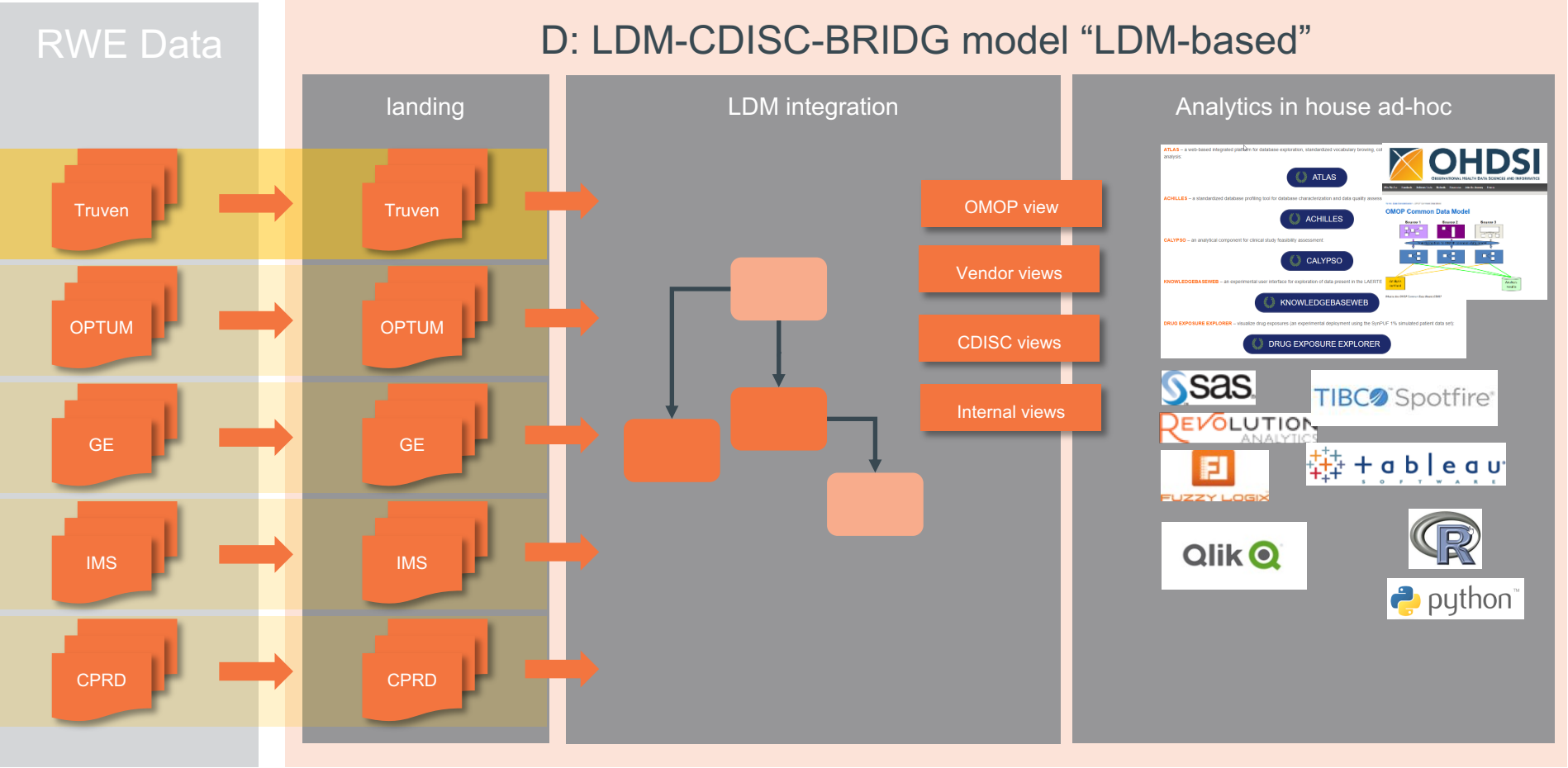
B: Relational and vendor specific



C: Relational OMPO



D: Relational LDM /CDISC/BRIDG



Integration commercial RWE

Easy / supported	
Medium / possible	
Hard / almost impossible	

Approach	Data format	Data Volume ¹⁾	Analytical capability								Analytics performance & scalability								Change management ²⁾
			SAS	R	Python	Tableau	Spotfire	Qlik	Aster	Audit trail analytics	future AI/ML	full volume	Complex models	Execution speed	Efficiency	Large user community	Clinical data integration	Re-usability	
A Traditional	SAS/CSV files	10x																	
B Relational and vendor specific	RDBMS	6x																	
C Relational OMOP	RDBMS	2x																	
D Relational LDM /CDISC/BRIDG	RDBMS	2x																	
E Vendor based	RDBMS	n/a																	
F Strategic	UDA	2x																	

Potential Saving

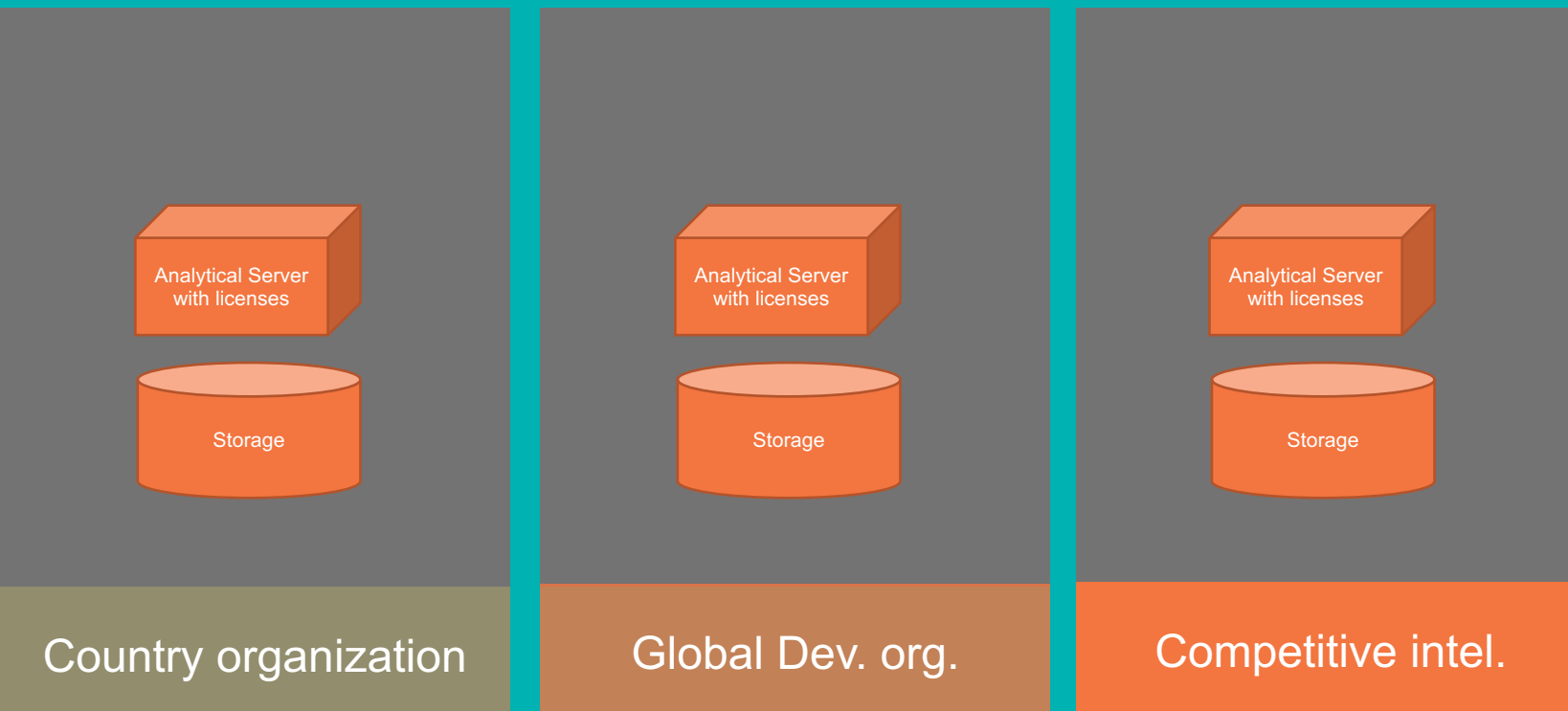
Manual integration 50% less
(2 FTE less)
Re-usability of code 20% less
(1 FTE less)
Server and License costs for analytics
(33% less)
No vendor provisioning

Governance can

- reduce integration, cleansing effort
- Guarantees actualization/operation
- Enables 'time-series' and hence new insights

1.0 Mill. USD/year

Infrastructure for analytics of RWE



Potential Saving

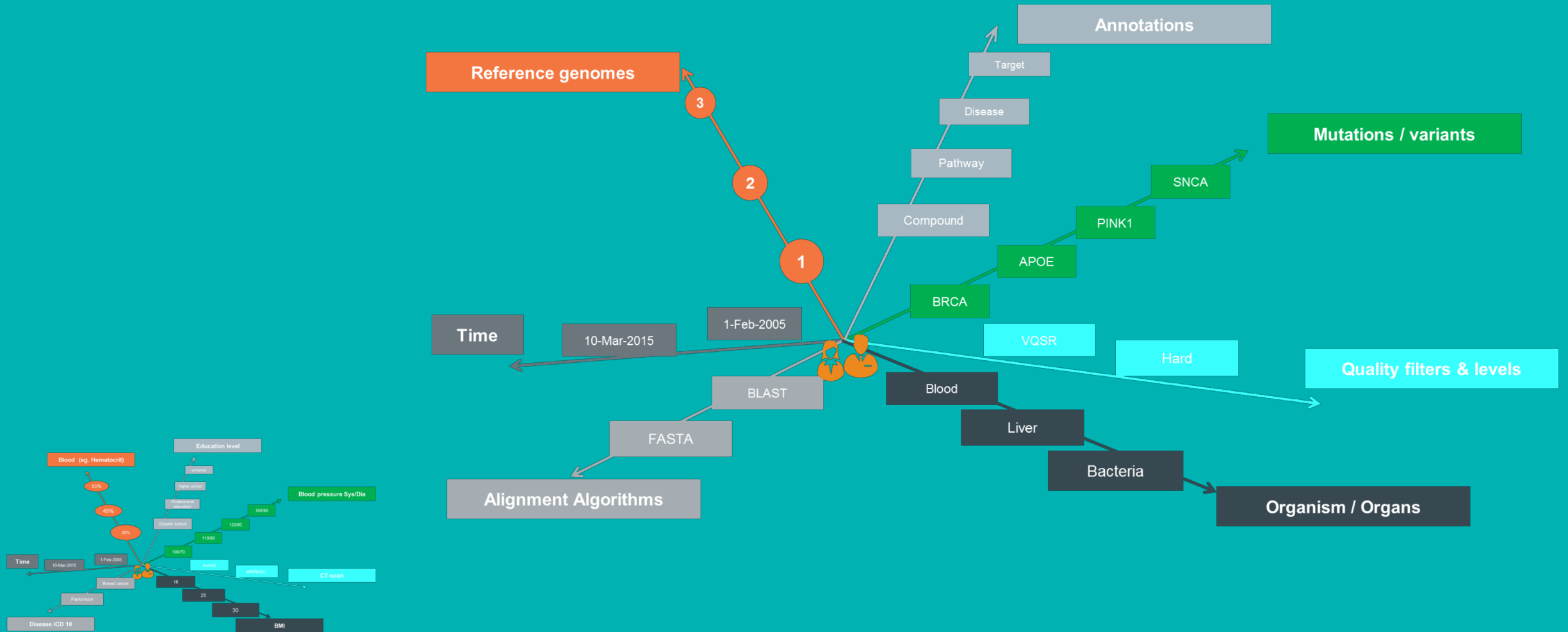
at least 2 environments less
(Assumption: 1 Mill USD total per environment)

Environments:

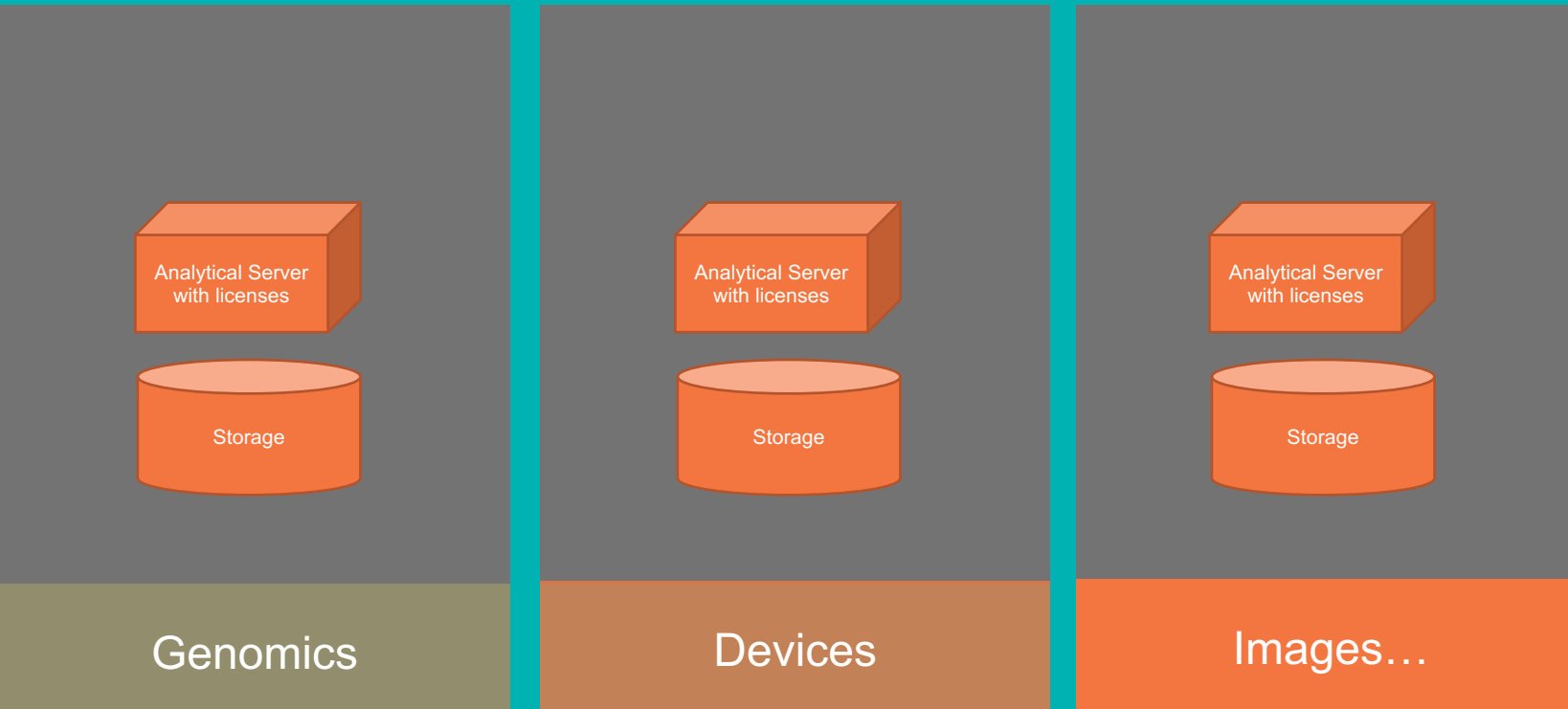
- Server and maintenance
- License costs for some of the analytical tools
- Effort to update and keep the environment going
- Services (e.g. user etc. and data mangament)

2.0 Mill. USD/year

GENOMICS / RWE and the dimensions in the data



New data types on a scalable system



Potential Saving

At least 1 new environment per year

(Assumption: 1 Mill USD total per environment)

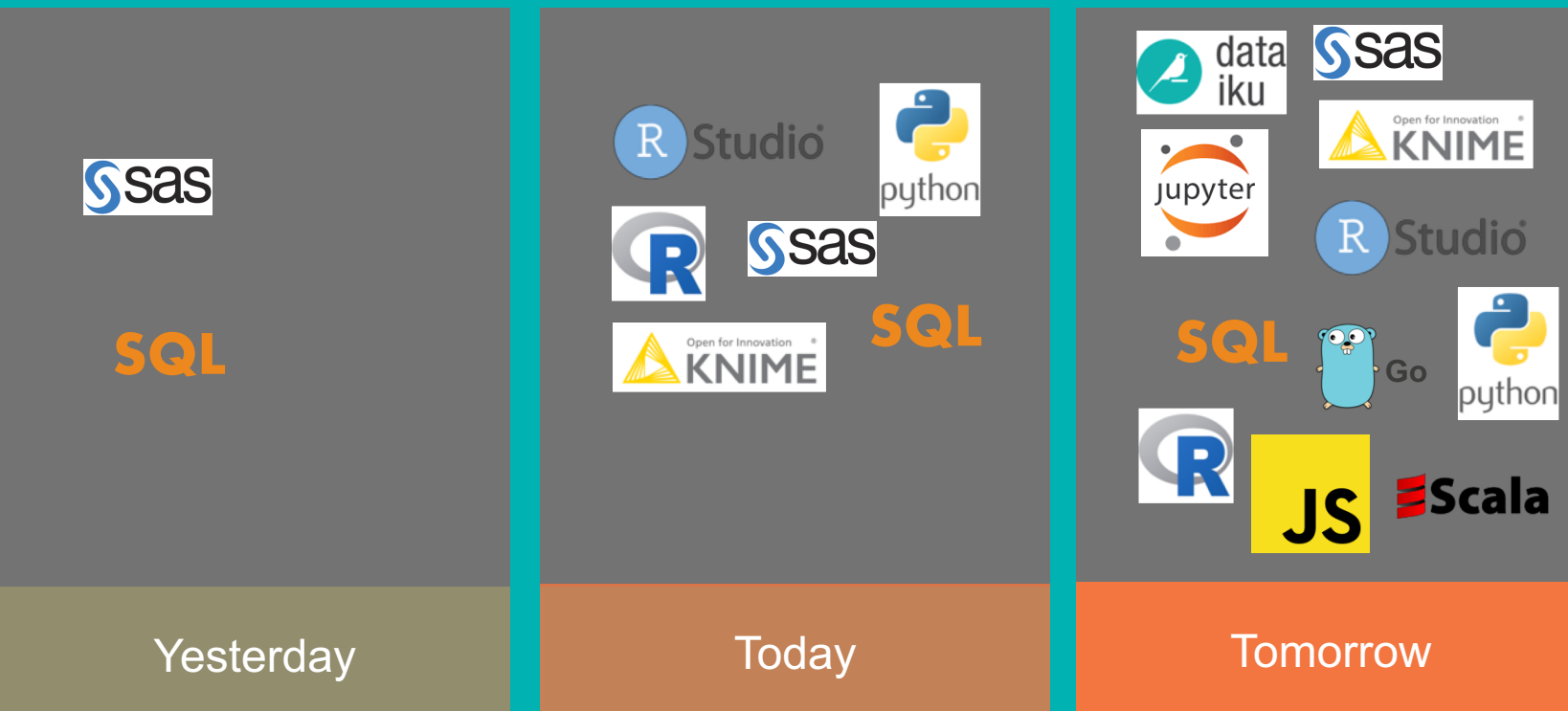
One instead of 3

Environments:

- No additional systems by data source
- Scalability given hence no migration need
- Re-use of existing infrastructure
- Fast deployment of

2.0 Mill. USD/year

New analytical engines, tools..



Potential Saving

At least 1 new analytics approach and language per year and at least 3 environments so limitation to one

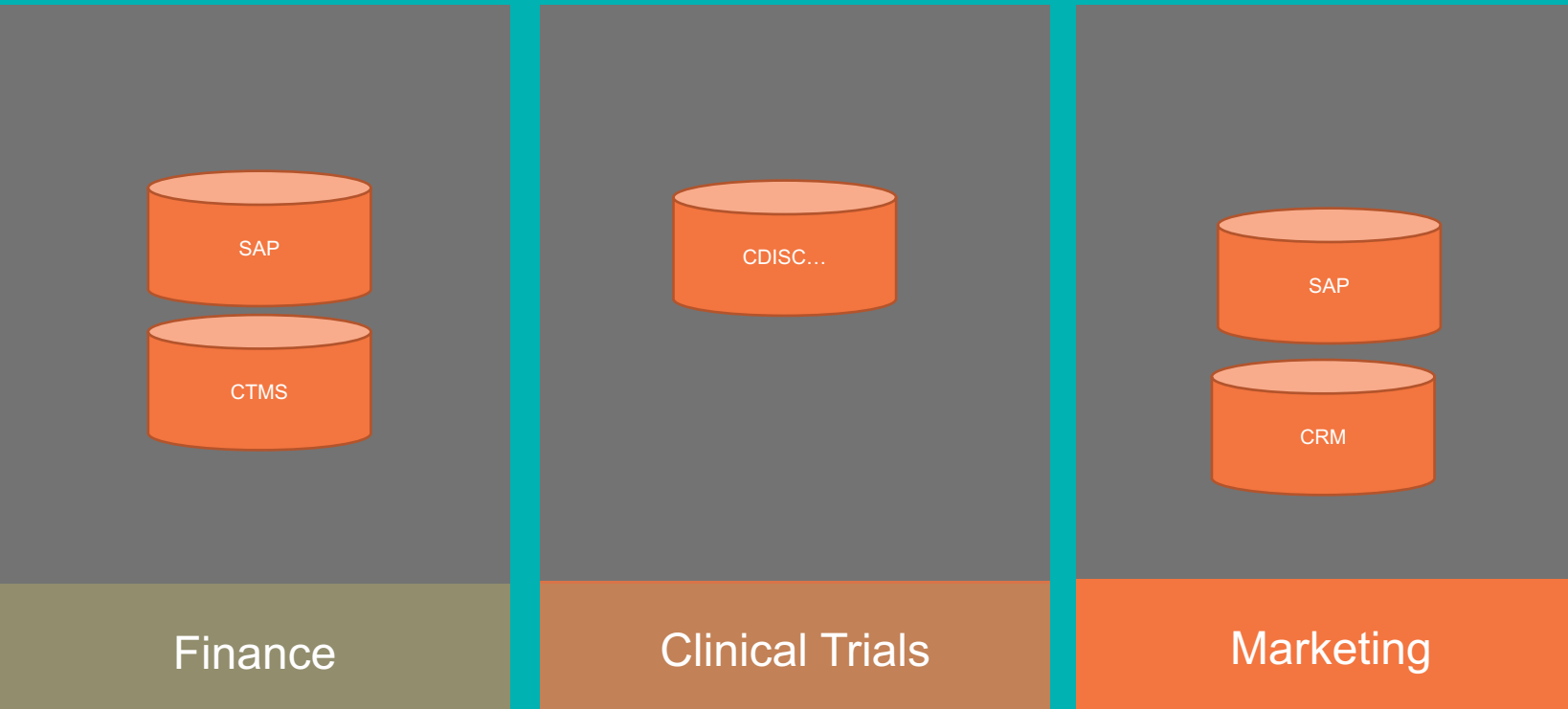
(Assumption: 1 Mill USD total per environment)

Environments:

- Limit the number of environments to cope with the changes
- Decoupling the data-store from the analytical engines (integration is independent)

2.0 Mill. USD/year

Existing data internally: operational integration



Potential Saving

Assumption is that there are at least around 5 to 10 integration activities per year
One data-pool per year because of a request (each 0.5 Mill)
(Assumption: 0.5 Mill USD total per integration and non sustainable)

One data pool per year à 1 mill can be avoided for clinical data integration)

Improvement

- Systematic integration (operational)
- Linking of the data (product details, profitability calculation for pricing)
- Trial cost details simulation and comparison with RWE

3.0 Mill. USD/year

Legal aspects

Public sector information, sometimes also referred to as government data, refers to all the information that public bodies produce, collect or pay for. Examples are geographical information, statistics, weather data, data from publicly funded research projects and digitised books from libraries. The European Commission's policies focus on generating value for the economy and society through the re-use of this type of data.

European Commission > Strategy > Digital Single Market > Policies >

Digital Single Market

POLICY

Proposal for a revision of the Public Sector Information (PSI) Directive



European Commission > Strategy > Digital Single Market > Policies >

Digital Single Market

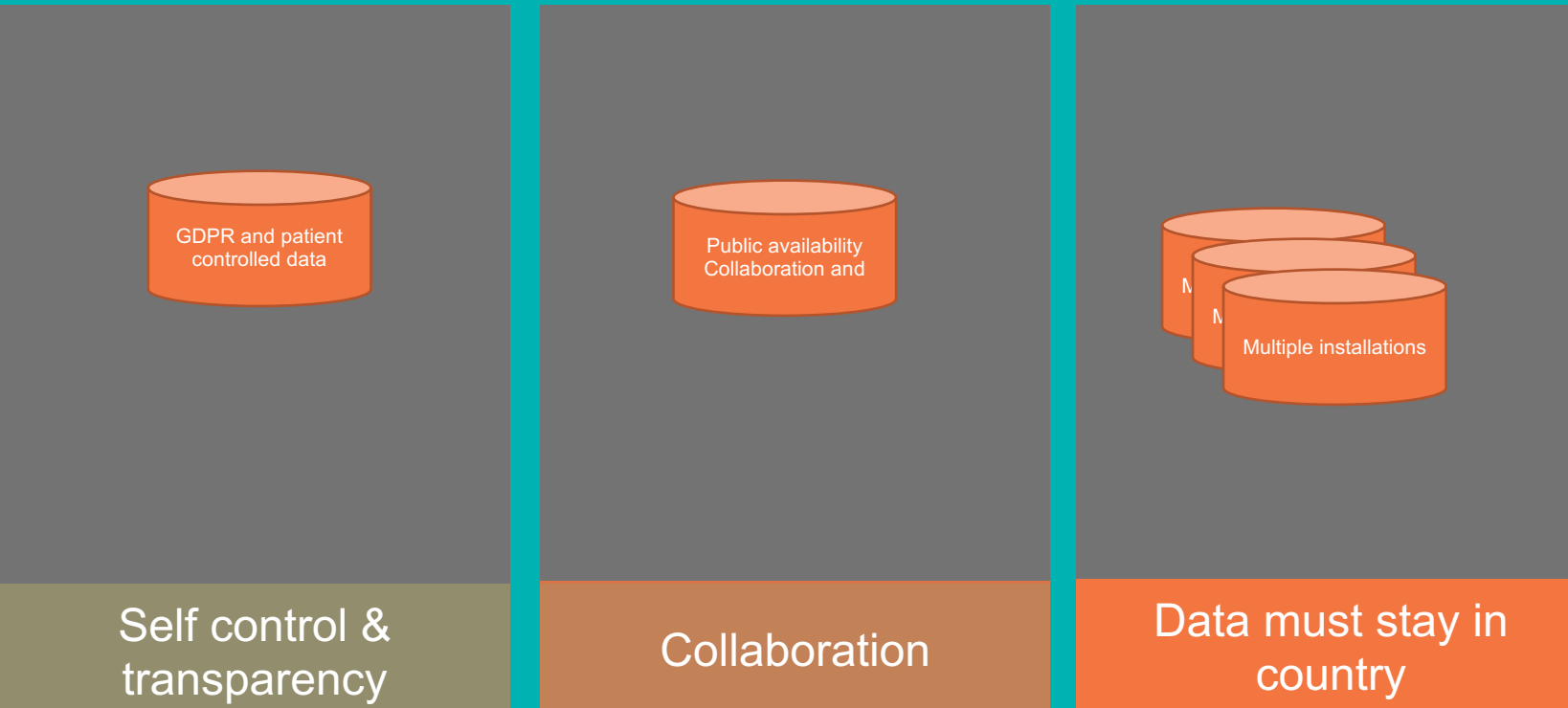
POLICY

Open data

Allowing public sector data to be re-used for other purposes, including commercial ones, can:

- stimulate economic growth and spur innovation: public data has significant potential for re-use in new products and services;
- help address societal challenges with the development of innovative solutions such as in healthcare or in transport;
- enhance evidence-based policymaking and increase efficiency in public administrations;
- become a critical asset for the development of new technologies, such as [artificial intelligence](#) (AI), which require the processing of vast amounts of high-quality data;
- foster the participation of citizens in political and social life and increase the transparency of government.

Legacy and collaboration related efforts



Potential Saving

As usually there are at least 3 different systems and in regard of country specific additional 2. IF a single system can handle it we save at least 3-4 systems

One system will cost at least 1 Mill / year

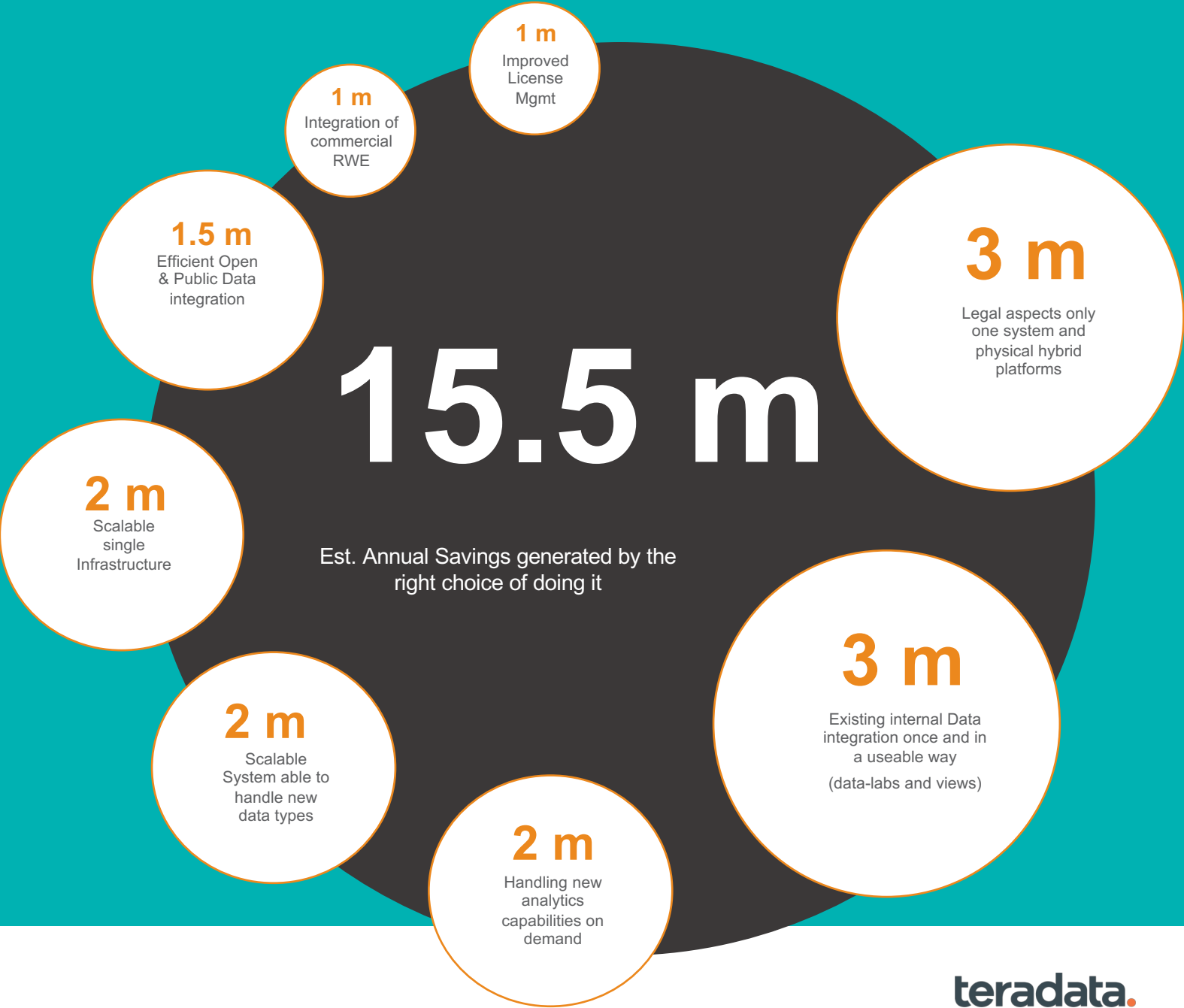
Environments

- Could be realized with the right design and capability on a single system
- Patients can control the data and see for what it used
- Behavioral analytics by patients, and institutions
- Governance and collaboration

3.0 Mill. USD/year

Savings per year:

Overview



Thank you.

teradata.

©2018 Teradata