Roadmap to Mobile Phone Data

Building preparedness towards use of mobile data

Siim Esko

Positium

Estonia

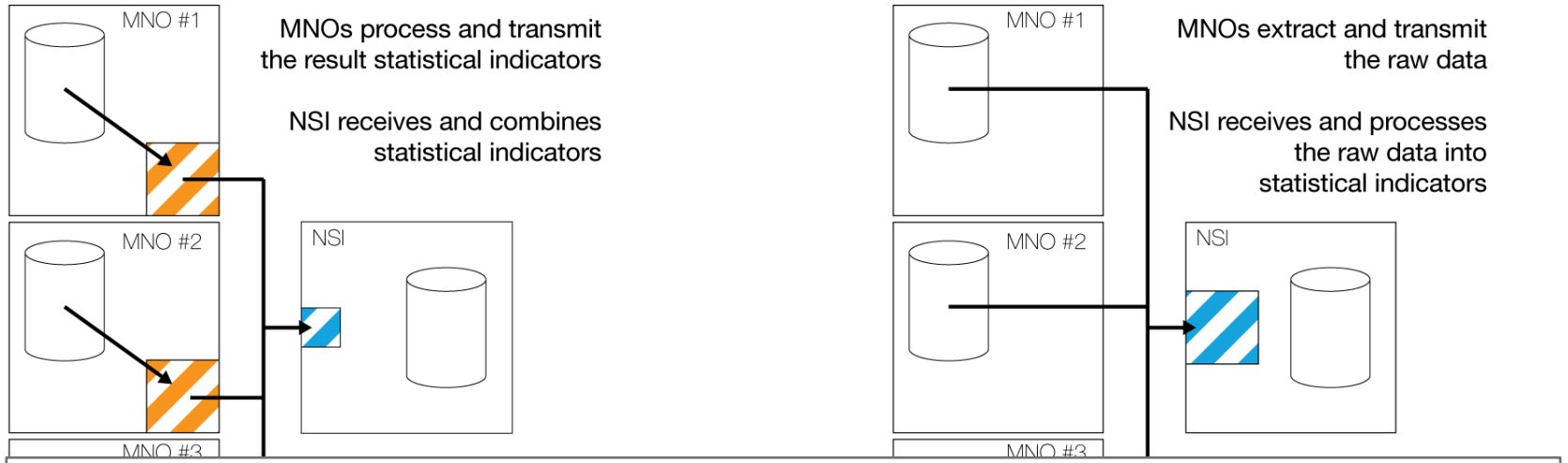
@positium

www.positium.com

The Puzzle Pieces – What do you need

Data Mobile operators	**Infrastructure** Data centre	**Skills** Staff, consultants, universities
Legal Privacy preservation Legal supervision	**Technology** Mature big data tech	**Funding** Low-hanging fruit International co-funding

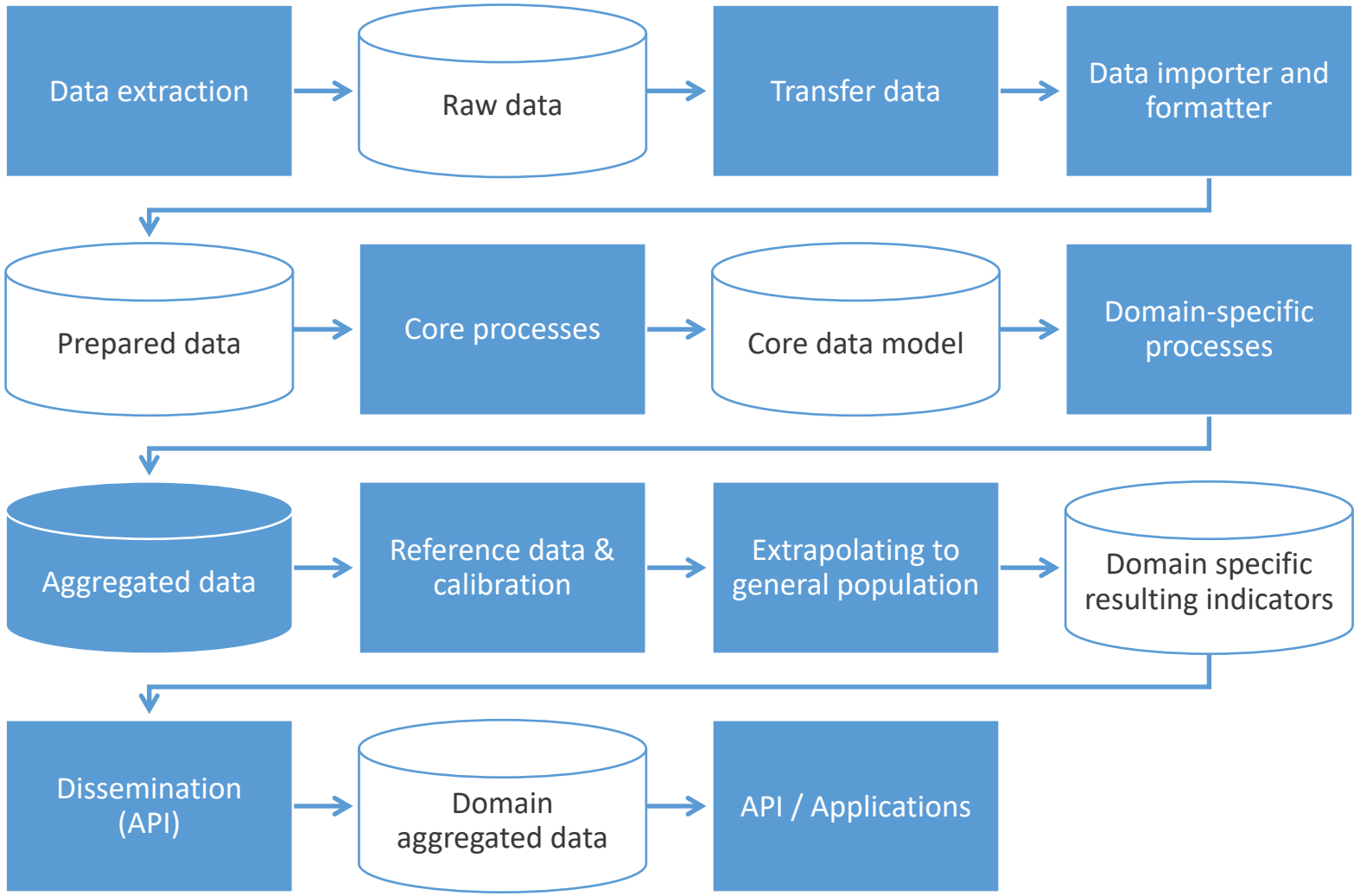
Processing Location



Depends on:

- Legal basis -> What is possible?
- Method -> Is centralized processing necessary for validity?
- Potential Future Scope -> What is efficient?

Components of the system – open source



hadoop

APACHE Spark

PostgreSQL

python

R

sdmx
Statistical Data and Metadata eXchange

CARTO

BigData
UN Global Working Group

positium

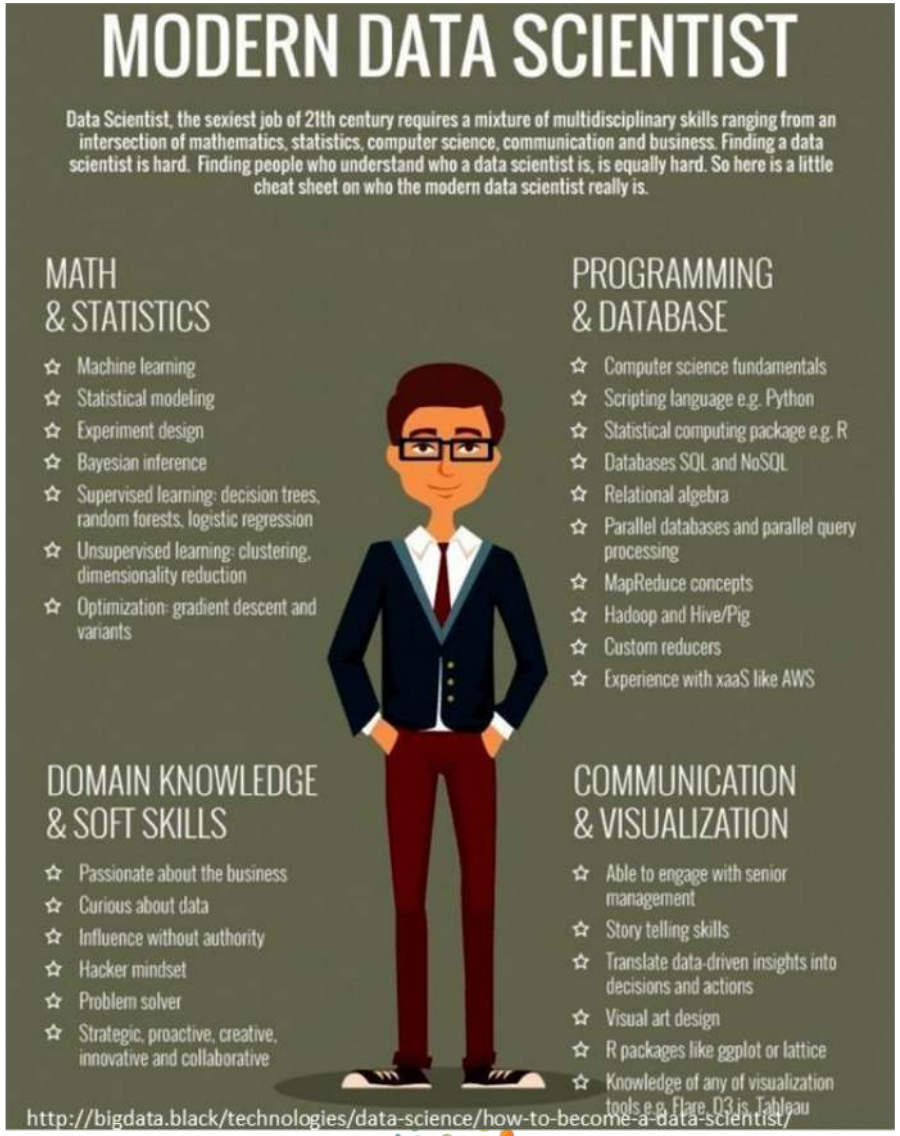
Desired aspects of the system

- Modular architecture
- Transitional databases (for data mining and QA)
- Metadata generation throughout the process
- Parallel processing (Hadoop?) according to processing time requirements
- API for dissemination
- Data transmission standards (SDMX, JSON, XML, …)

- White box / Grey box / Black box
- Maintenance agreements (SLA)

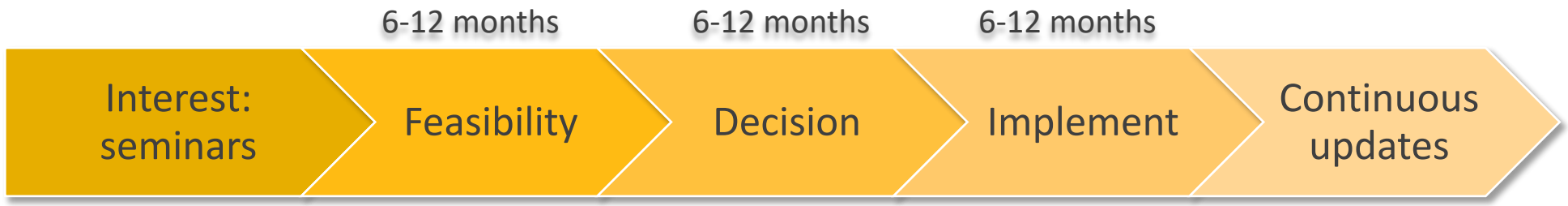
Required Skills

- Tools for running processing algorithms (e.g. Java, Python, R)
- Supporting tools for QA of the data and visual analyses (e.g. GIS tools such as Esri, MapInfo, QGIS, R, and Excel, Tableau, Qlik, etc.).
- Skills of developing software for the parallel processing of big data – data processing software, ETL (Extract, Transform and Load) developer with Python, Java, PIG, Hive, Spark, SQL, etc.
- Skillset for PostgreSQL/PostGIS, Oracle or any other database system (including GIS data) is required.
- System architect and system operators
- Statistical data processing methodologists with big data experience.



http://bigdata.black/technologies/data-science/how-to-become-a-data-scientist/

Desired Timeline



Cost of the system

Seminars, workshops

+

Pilot Project / feasibility assessment

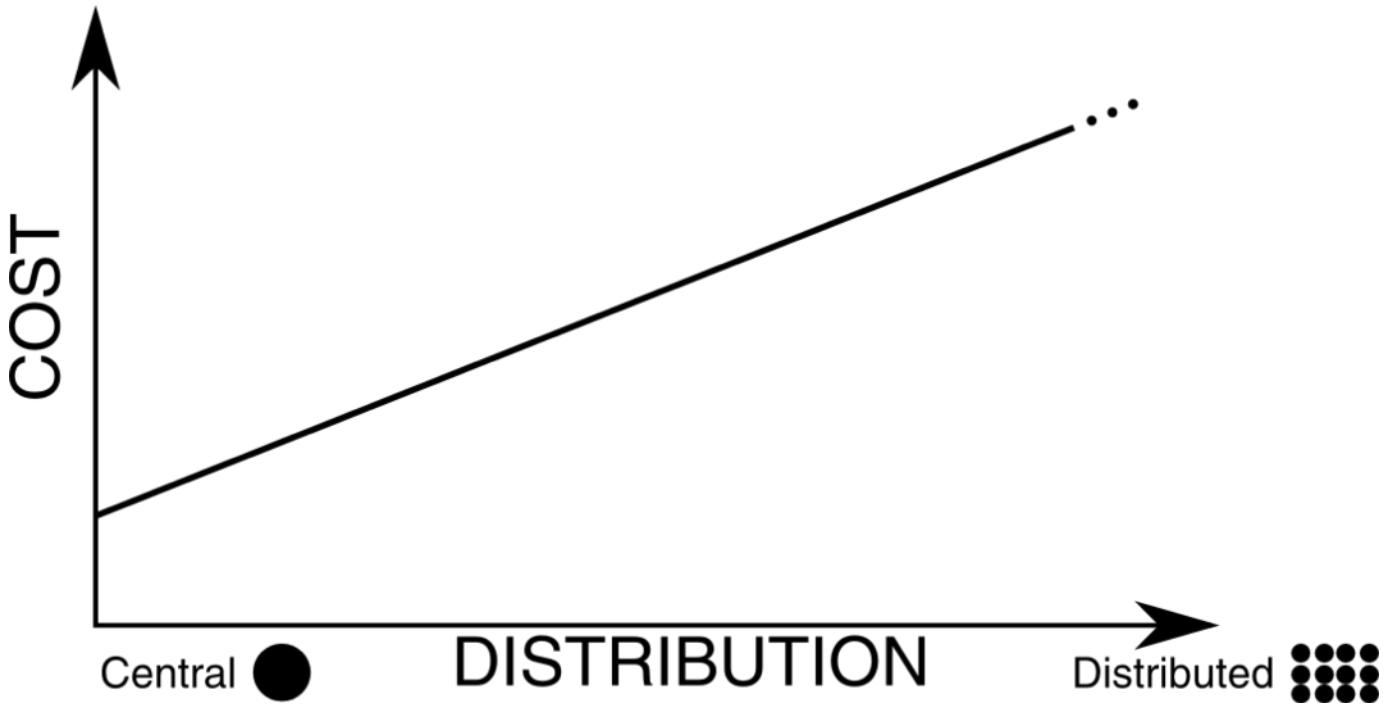
+

Implementation cost

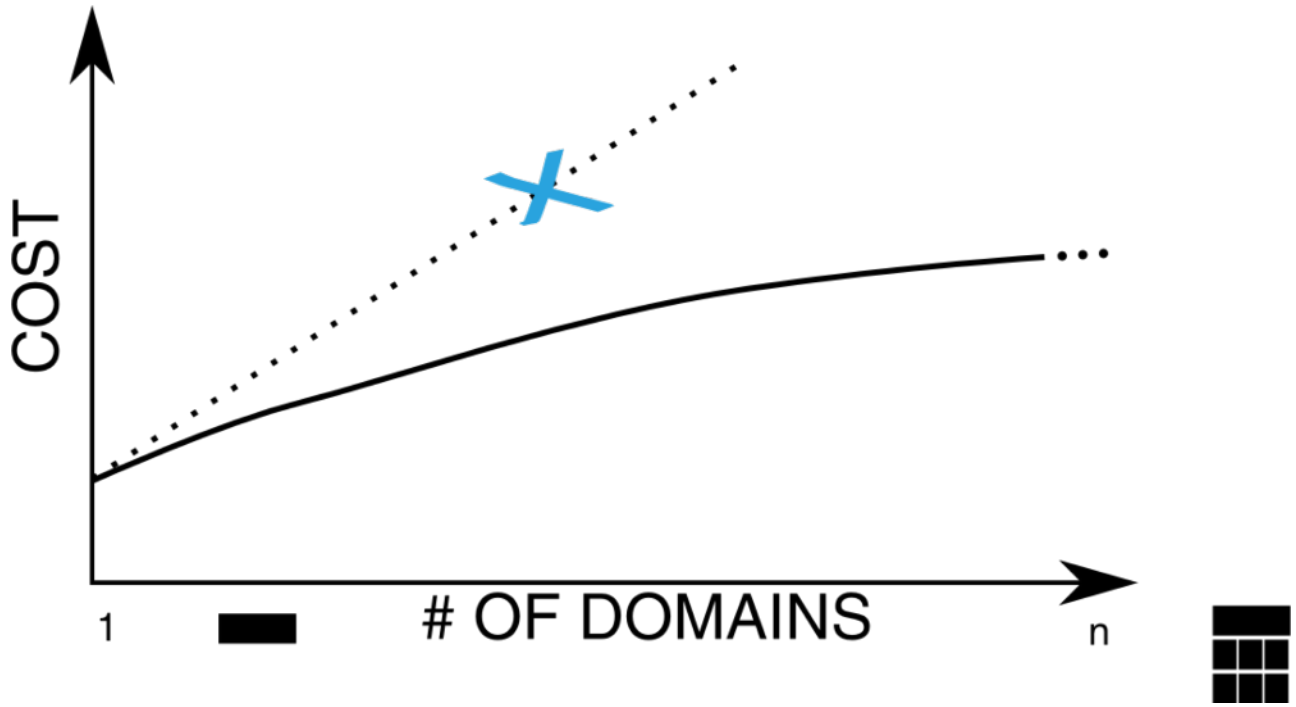
+

Maintenance (SLA)

Cost of the system

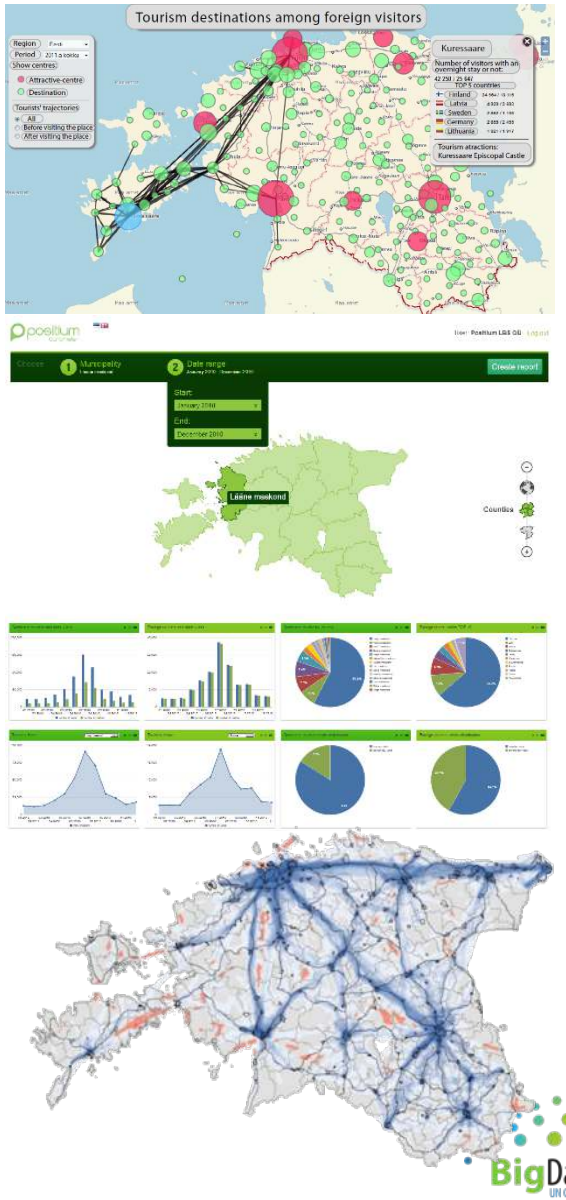


Cost of the system



Funding for Mobile Data Analytics comes from Value

- Population statistics
- Everyday commuting
- Transportation planning
- Urban planning
- De facto population
- Tourism statistics
- Epidemiology
- Economics
- Natural disasters, proximity of population to risks
- Safety and Security
- ICT statistics
- Spatial marketing
- Scientific research in many areas



Homework

- What are the first priority projects in your country for the use of mobile phone data?
  - Personal opinion – Gathered at the workshop
  - Consolidated NSO opinion – To do after returning back to office