

Science Mesh beyond science

perspectives for adoption in a wider business context

Marcin Sieprawski
Head of Big Data Lab
Software Mind





Software house focused on building **dedicated teams** to extend **product engineering** and **digital transformation** capabilities

4

R&D Labs
in Poland

HQ in Cracow
Branches: Warsaw,
Rzeszow Bielsko-Biala



Branches

and representations

USA, Australia, Singapore,

20

Years of experience
established in 1999

Big enough to scale, small enough to care

Software Mind – Agile Soul

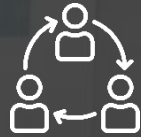


Software Mind Toolbox



Agile for outsourcing

Our teams will feel like part of your organization



Agile scope management

How to start fast with almost no requirements



Agile reporting

Always be sure what you get from your teams



Automated testing

Frequent releases without army of manual testers



Code craftsmanship

Great craftsmanship is foundation for great products



Short production cycles

Don't ask about the progress ... see it for yourself



Security

It's not only responsibility of InfoSec team



Automated deployments

Release new features quickly with a DevOps mindset



Right communication

Good interaction is key to your success

Selected case studies

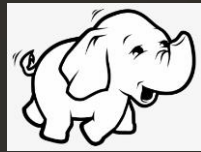


Startup: Electronic Identity Protection Platform

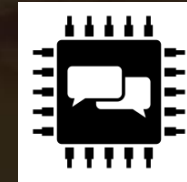
Web-scale Semantic Web startup (2005-2008)
commercializing new technology



Bare-metal cluster



Hadoop



Natural Language Processing

Main Facts:

Customer

UK Startup

Project

Building new product
from scratch – electronic
identity protection
platform

Operating Model

Dedicated Teams,
MVP approach, R&D
with universities (University
of Southampton, University
of Sheffield)

Challenges:

- Transfer of innovative technologies (Semantic Web, NLP, Graph Databases) from universities to real business
- Scaling the niche technology to fully blown commercial solution (e.g. browsing 4 billion of web pages)
- Building Big Data solution even before this term was defined

SETA

Big Data - large scale smart mobility management platform (2017-2019)



Private Cloud



Cloud - native

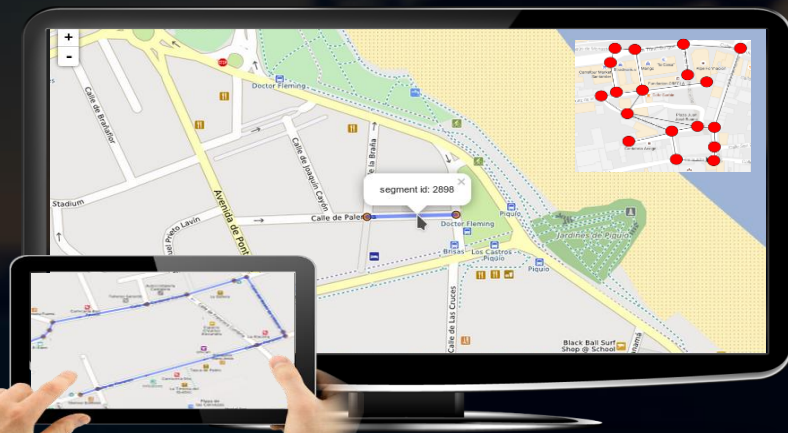


Apache Spark



Data Science environment

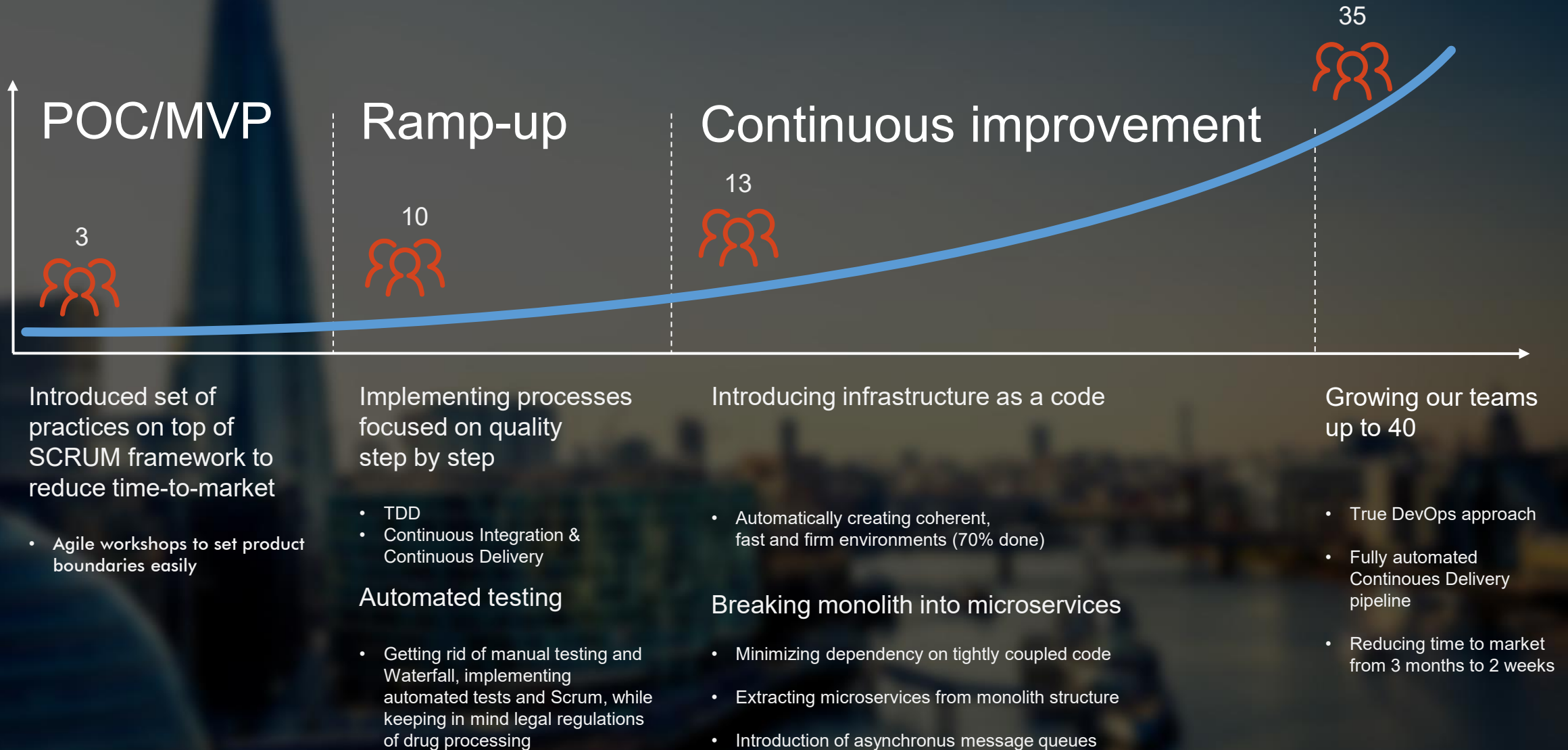
- Big Data technologies for monitoring and managing mobility in large metropolitan areas.
- Solution based on data from millions of citizens, thousands of connected cars, thousands of city sensors and hundreds of distributed databases



Main Facts

- Management of huge geospatial and spatiotemporal data
- GPU-based acceleration of geospatial indexes
- Privacy by Design
- High performance geo-located event processing engine
- Scalable backend for mobile apps
- Integrated machine learning components
- Data Science environment built into Agile software development process

[XXX] Software Mind Digital Transformation Services

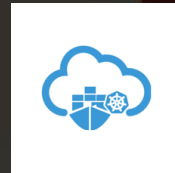


ScienceMesh

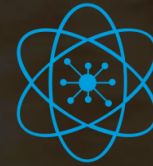
CS3MESH4EOSC project



Cloud inter-operability platform



cloud-native



Distributed Data Science environments



JupyterLab extension (CS3 APIS)

- **Leading tasks on**

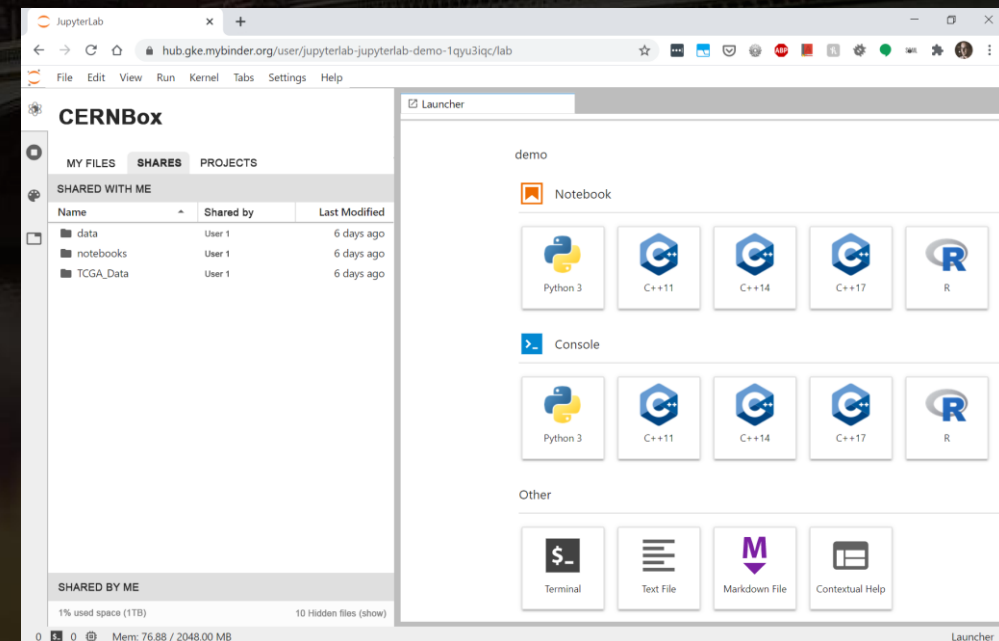
- Reference inter-operability platform
- Distributed Data Science environments

- ScienceMesh Inter-operability platform

- make cloud storage and application providers inter-operable, via the CS3 APIS

- JupyterLab extension (Cs3Api4Lab)

- Integration with ScienceMesh IOP (CS3 APIS)
- replaces the default file manager
- new UI elements for share functionalities





To cloud or not to cloud?

Cloud

Main factors: public cloud

- Scalability, high availability, low latency
- Cost optimization – IaaS
- Global infrastructure and services
 - including global databases
- Managed services beyond IaaS
 - managed databases, FaaS
- Cloud automation
- One shop – contract negotiation
- Startups: OPEX + free tier (evaluation/prototyping)
- Uniform and complete offer of services

On premise

Main factors

- **Data security: highly sensitive data**
 - [UK startup] Identity protection platform
 - BMD, Companies House + data collected from the Internet
 - Alerts for illegally traded data
 - Even search can be dangerous (AOL 2006)
 - SETA
 - GPS tracks, Floating Car Data
 - Privacy by Design
 - GDPR and other regulations
- **Response latency / data transfers**

Cloud interoperability

Main factors

- **Hybrid / multi- cloud**
 - Preventing vendor lock-in
 - Cost optimization (private cloud)
 - Managing sensitive data (Privacy by Design)
 - Supporting digital transformation (a proces: multiple environments)
- **Distributed cloud computing**
 - location of cloud-delivered services - part of its definition
 - Important in distributed data science environments
- **Main factors of cloud adoption**
 - Integration skills (hybrid cloud -> connections and integration points)
 - native-cloud skills
 - Interoperability tools

Thank you for your attention!

Marcin Sieprawski

Head of Big Data Lab

marcin.sieprawski@softwaremind.com

Headquarters

ul. Życzkowskiego 20
31-864 Kraków

Tel.: + 48 12 252 34 00

web: www.softwaremind.com

Branches in Poland

Warsaw, Rzeszów, Bielsko-Biała

Worldwide

Australia / Sydney, Brisbane

Singapore / Singapore

USA / Boise, Dallas

