Data Archeogram

Mapping the datafication of work

Armelle Skatulski

January 2021

© Autonomy
autonomy.work

Industrial Internet of Things (IIoT)

Internet of Things (IoT)

The connected worker

The connected user

Data Mining

Algorithmic Governance

Digital primitive accumulation
Informational asymmetry

Behavioural predictions
User as product

Data-driven monitoring
Edge computing

Targeted marketing
Devices augmentation

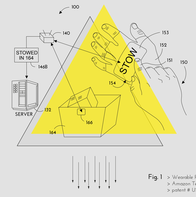


Fig.1 > Wearable RFID device
> Amazon Technologies Inc.
> patent # US9,881,955 B2, 07.17.2017

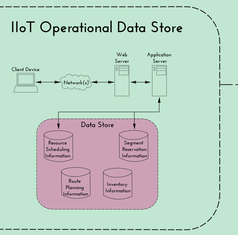


Fig.6 > Data store system for data collected from ultrasonic bracelet
> Amazon Technologies Inc.
> patent #US 9,881,076 B2, 30.01.2018

Visual Analytics
Streaming

Profiling

Data Extraction

The connected workplace

e.g. the smart warehouse or factory

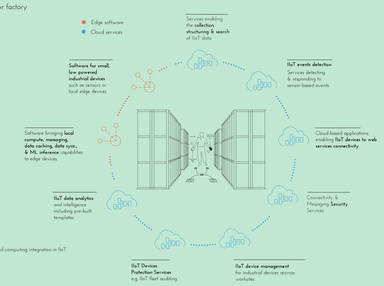


Fig.4 > Edge & cloud computing integration in IIoT

The connected social factory

e.g. the connected domestic or social sphere

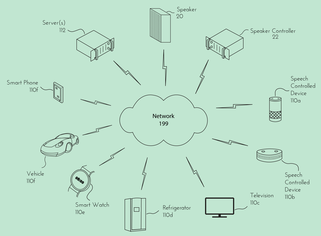


Fig.5 > Processing spoken commands
> Amazon Technologies Inc.
> patent # US9858927 B2, 02.01.2018

Employer

Data Collection

Multiple Collectors

Choice architectures
Nudging

Profiling

Edge Deep Learning on IoT devices

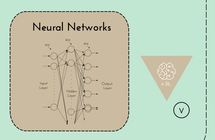


Fig.7 > Adaptive neural network speech recognition models
> Amazon Technologies Inc.
> patent # US9,153,231 B1, 06.10.2015

Internet Infrastructures

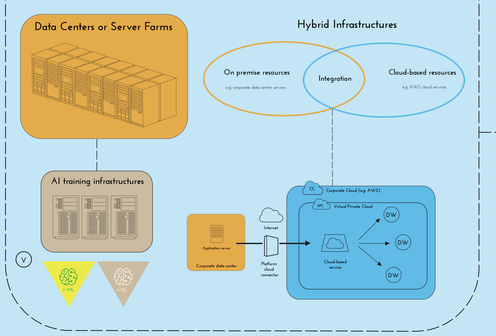
Internet Infrastructures

Infrastructures of Storage, Processing, & Analytics

Cloud computing
Analytics

Predictive modelling
AI design

Data Storage & Analytics Infrastructures



Data Landing

> Terminology obscures active strategies

Analytics

Data Warehouses
> designed for data analytics
> pattern recognition across large datasets
> data organized in tabular formats through schemas
e.g. to implement *SQAL queries
> big data analytics, full text search, and machine learning,
can access semi-structured or unstructured data

*SQL - structured query language is a standard language for storing, manipulating & retrieving data

> e.g. Amazon Redshift; Microsoft Azure Synapse; Google BigQuery; Snowflake

Data Migration

Pre-analytics:
> selection
> labeling

Data preparation:
> by hand (clickwork)
> by algorithms

Data Lakes
> a repository that can store data in their native format: structured, semi-structured, and unstructured
> can be built by customers using cloud-based services where they can run their own analytics incl. open source tools
> data analytics using programs such as U-SQL, R, Python and .NET, Apache Hadoop
> architecture elements, enterprise IT data exchange, catalog services, governance, self-service access

> e.g. Microsoft Azure; IBM & Redhat; Hadoop; AWS

Databases
(relational or transactional)
> designed to capture and store data from transactional systems
> workloads : transaction processing
> optimized for running production systems
> source : data captured as-is from a single source
> continuous write of operations for transactions throughput
> highly normalized, static schemas

> e.g. Amazon RDS, Single Store, MySQL

data normalization

Data Warehouses

Analytics

Data Marts
> a data warehouse that serves the needs of a specific team or business unit, e.g. finance, marketing, or sales

Data Lakes

Databases

Data Swamps
> ungoverned data lakes

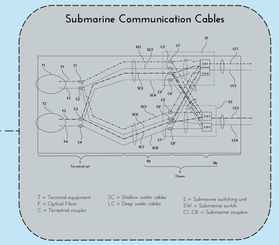


Fig.8 > Equipment for long range submarine optical communication
> Alcatel Submarine (FR), Nokia Solutions & Networks Oy (FI)
> patent # EP3710065A1, 23.09.2020

Platforms

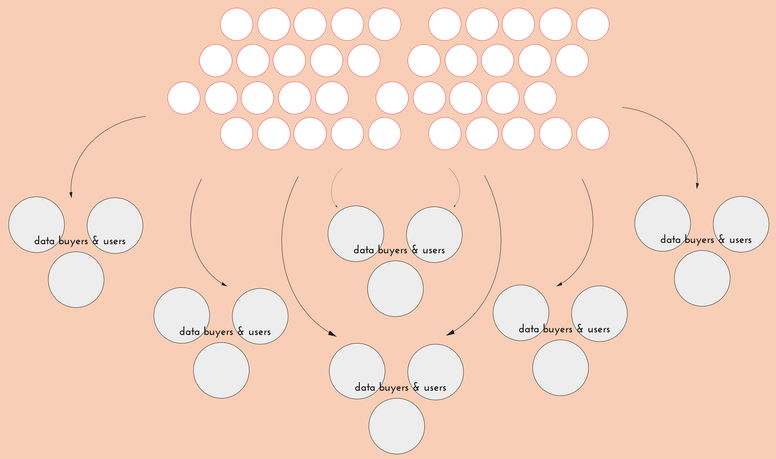
Market Leaders:

> Amazon
> Microsoft
> Google
> IBM
> Siemens
> Alibaba

Platforms

Infrastructures of Data Rent & Usage

Data Marketplaces



Data warehouses users:

> Business analysts
> Data scientists
> Data developers

Types of analytics:

> Batch reporting
> Business intelligence
> Data visualisation

Data Lakes Users:

> Business analysts
> Data scientists
> Data engineers
> Data architects

Types of anlytics:

> ML
> Exploratory analytics
> Data discovery
> Streaming

Legend

- (DC) Event of data capture
- (V) Data Valorisation
- (I) Issue warning: e.g. data rights breaches, security risks, etc.
- (T) Technology for: data capture ; sale/rent ; process ; storage ; etc.
- (I) Invisible forms of labour
- (DW) Data warehouse
- Data Capture Devices
- Data Marketplaces
- Cloud Computing
- Edge Computing
- Machine Learning
- Deep Learning Neural Networks
- Notice-and-Consent Mechanisms
- Internet Infrastructures

Autonomy