Ultra Tendency

- Big Data Consultancy in Germany and Europe since 2010

- Plus 130 Experts with global footprint

- We are the **Data Excellence** company

  - Centralised Data Platforms

  - Large-Scale data driven applications

  - Data Mesh

  - ML / AI Ops

  - Cloud Migrations

  - Managed Service

- References in Telco, Automotive, FSI, Utilities, Manufacturing and more …

Agenda – Data Mesh

- Definition

- Benefits

- Challenges

- Project Execution

- Live Reference

Data Mesh Principles

- ## Domain-driven ownership

  - Share directly with data consumers

- ## Data as a Product

  - Share the data most valued by consumers

- ## Self-serve data platforms

  - Increase utilisation by producers and consumers alike

- ## Federated computational governance

  - Security, Compliance, Availability, Quality, Entity Standardization, Provenance, Usage

  - Utilise centralized systems and processes

Data Mesh aims to overcome the limitations of traditional centralized data management

- It combines several important trends in data management:
  - domain-driven design
  - multiple domain-specific systems, each specialized according to its functions and uses.
  - data as a product
  - decentralize the ownership of ingestion, processing, and serving of data

A Data Mesh is based on a <u>distributed</u> architecture for data management

Data Products are made easily available to business users – no added complexity
Data Products are discoverable, understandable, trustworthy and usable

Data Mesh Architecture
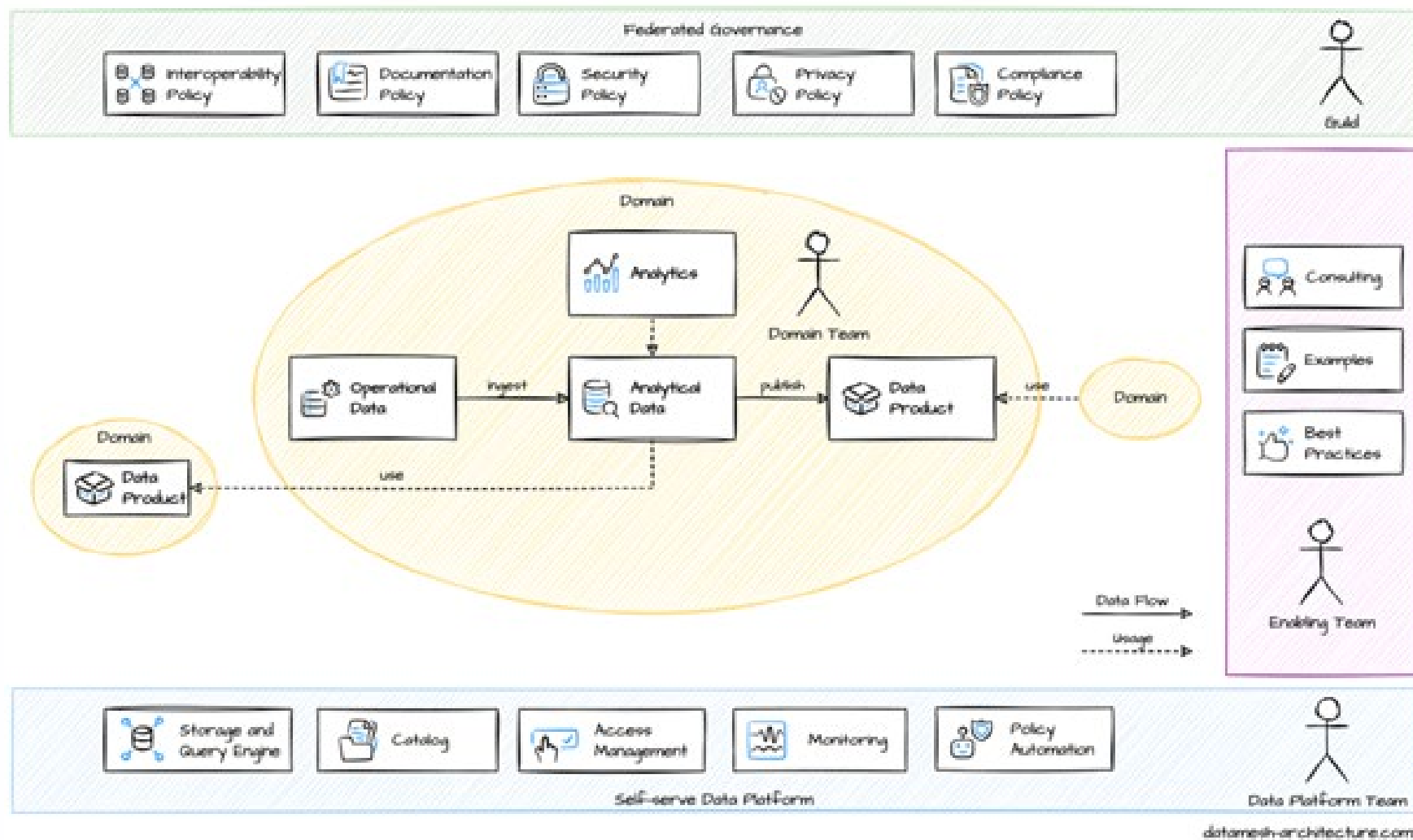
Why Data Mesh?

- Current trends in the market conditions:
  - Increased data volume
  - Increased expectations of nearly instantaneous data access
  - Increased need to respond efficiently to changes, due to:
    - Market Volatility
    - New Business Models
    - Competitive Pace
  - Overcome technology restraints of Domain Installations
    - Make it simple to provide and consume data

On the plus side...

- Empowered & distributed domains

- Reducing Cost for Data Analytics

- Improved data quality

- Highly scalable

- Reduced complexity

- Improved cross-functional collaboration

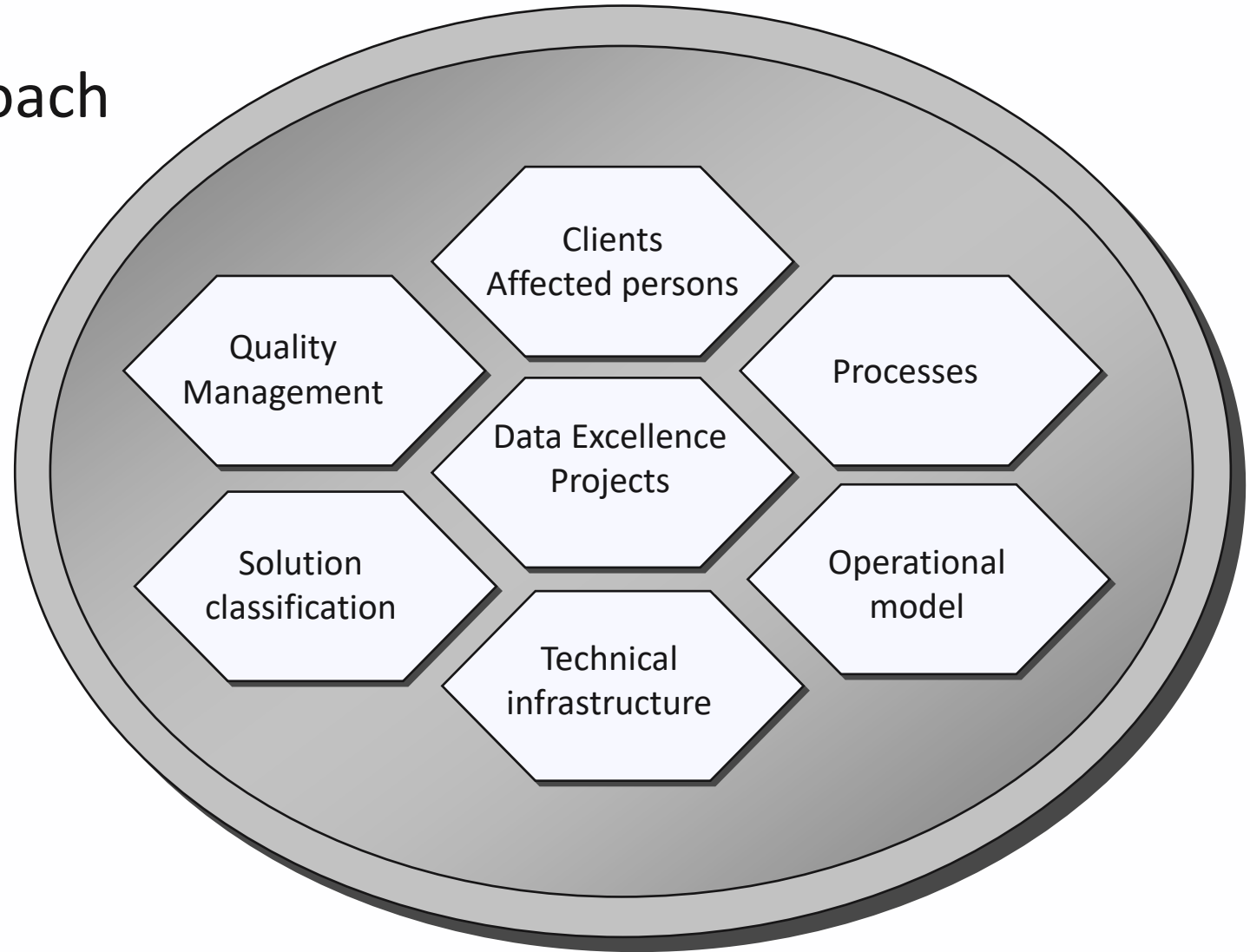- Enabling teams to work autonomously → greater innovation

On the other side…

- Failure to automatically update data catalogues
  - Documentation can be out of date
- Recognise the value of your current team
  - Data lake, warehouse, products all need to now work more, these are the good people!
- Data mesh requires extra work by domain teams
  - Have enough manpower to manage the priorities
- Shadow data analytics
  - Making your own reports may lead to own interpretation thereof
- Data products need to evolve
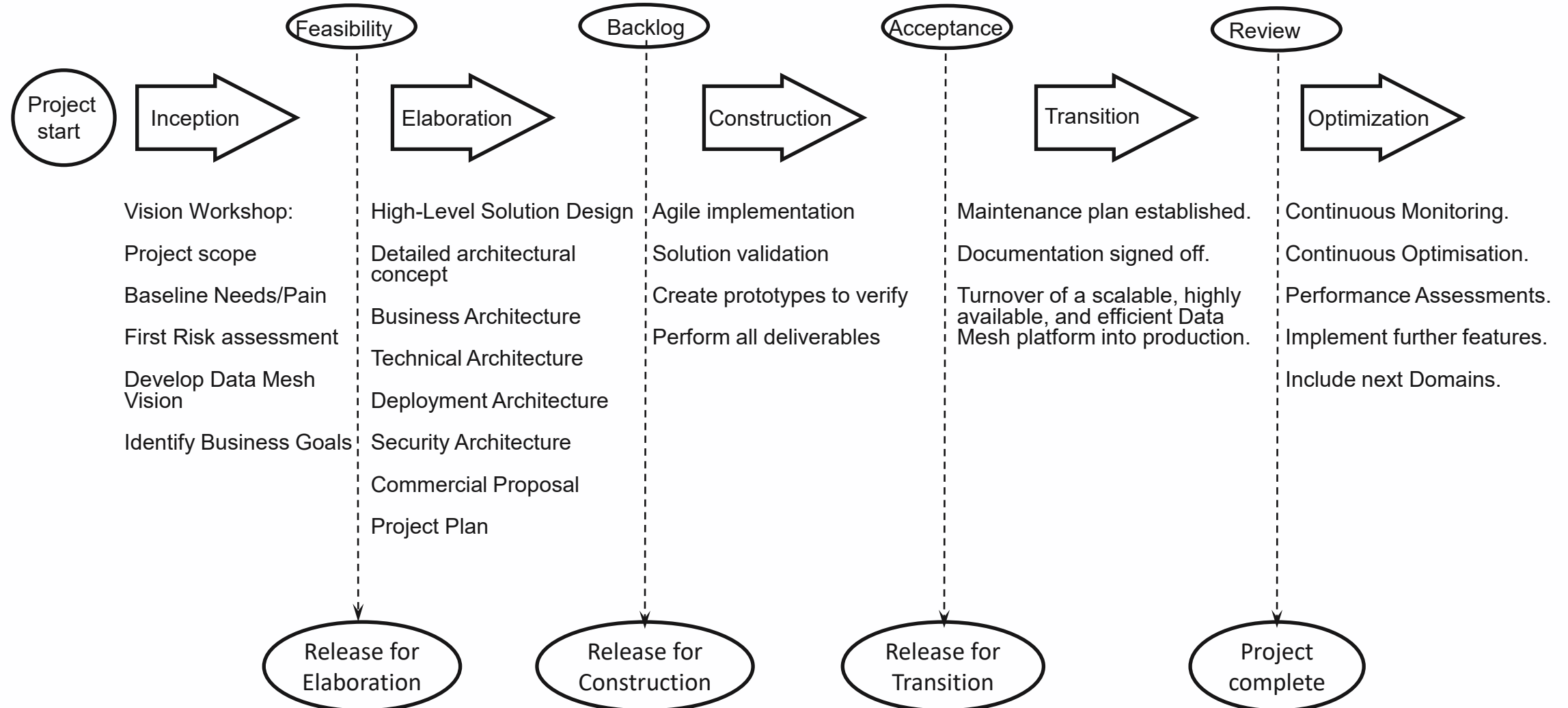 - Things change, meaning of data changes, yesterday's report may not be accurate for today's challenge

Wholistic Approach

Technology Partners

Feasibility   **Backlog**   **Acceptance**   **Review**

Project start

Inception   Elaboration   Construction   Transition   Optimization

Vision Workshop:

Project scope

Baseline Needs/Pain

First Risk assessment

Develop Data Mesh Vision

Identify Business Goals

High-Level Solution Design

Detailed architectural concept

Business Architecture

Technical Architecture

Deployment Architecture

Security Architecture

Commercial Proposal

Project Plan

Agile implementation

Solution validation

Create prototypes to verify

Perform all deliverables

Maintenance plan established.

Documentation signed off.

Turnover of a scalable, highly available, and efficient Data Mesh platform into production.

Continuous Monitoring.

Continuous Optimisation.

Performance Assessments.

Implement further features.

Include next Domains.

Release for Elaboration   Release for Construction   Release for Transition   Project complete

International Data Spaces

Mission

- The International Data Spaces Association (IDSA) is on a mission to create the future of the global, digital economy with International Data Spaces (IDS), a secure, sovereign system of data sharing in which all participants can realize the full value of their data.

SOLUTION

- Created a Reference Architecture for Data Mesh and Data Brokerage (https://internationaldataspaces.org/offers/reference-architecture/)

CUSTOMER SUCCESS

- Gaia-X - an initiative that develops, based on European values, a digital governance that can be applied to any existing cloud/ edge technology stack to obtain transparency, controllability, portability and interoperability across data and services.
- Catena-X - the first collaborative, open data ecosystem for the automotive industry of the future, linking global players into end-to-end value chains - as simply, securely and independently as never before. The shared goal: a standardized global data exchange based on European values. The claim is data sovereignty.
- DIH – see next slide

Deutsche Telekom: Data Intelligence HUB (DIH)

SUMMARY

- Ultra Tendency built Europe's largest data brokerage and self-service analytics platform for Deutsche Telekom IoT GmbH (monetization of data without loss of sovereignty). Hereby, Ultra Tendency created a Data Mesh before the term was known.

CHALLENGE

- High degree of innovation, diverse range of functions of the platform (especially in the area of data analysis), as well as requirements for data sovereignty (first reference implementation of International Data Space IDC connector)

SOLUTION

- Ultra Tendency architected and built the product end-to-end (Requirement Engineering to Front-end). The DIH is based on a microservices architecture on Azure. It leverages Azure-native components like AKS, ADLS, Power BI, as well as JupyterHub, Rstudio Grafana, Apache NiFi, Apache Kafka, Apache Zeppelin, H2O, IBM Watson Studio

CUSTOMER SUCCESS

- Largest data marketplace in Europe that offers third parties access to secure data exchange and data monetization
- First data marketplace to meet the strict security requirements of the International Data Spaces Association
- Over 1,000 users from a wide variety of industries

Thank you

Mike Görlich
VP International Sales
Tel: +49 (0) 176 4550 4132
Mail: mike.goerlich@ultratendency.ag

Ultra Tendency Book of Standards

We've learned a lot. We've written it down.

- How do you build high quality software?

- How do you adhere to requirements and scope?

- How do you build on time and budget?

Ultra Tendency Academy - USPs

- Customized trainings based on your individual demand to get the best learning effects for your team

- Practice-oriented trainings and coaching for state-of-the-art technologies (Big Data, Cloud, Container, IIoT)

- Training can span across technologies from different vendors

- Hands-on activities on customer environments (e.g., Dev / Test) to train within customer context (if possible)

- Learn from industry experts with extensive real-world experience

- Address Technology & Methodology

- Introductory and deep dive trainings

- Target groups: Architects, Developer, Administrator, Management

Curriculum

Cloud
Data Driven Applications on Azure /AWS; Operating Cloud-based applications

Big Data
Kafka, Data Engineering, SPARK, Flink, MLOps, CDP introductory

Cross Functional
Infrastructure as Code, Continuous Delivery, Enterprise Application Architecture

Container
Docker / Kubernetes, Microservice Architecture, Modern DevOps approaches

IIOT
IIOT Application Design, IIOT Application Development and Operations

Ultra Tendency

- First ever provider of Big Data Consultancy in Germany and Europe
- Plus 130 Experts make us the most experienced company dealing with any data related challenge
- We are the **Data Excellence** company
  - Centralised Data Platforms
  - Large-Scale data driven applications
  - Data Mesh
  - ML / AI Ops
  - Cloud Migrations
  - Managed Service
- References in Telco, Automotive, FSI, Utilities, Manufacturing and more …

Backup slides

Data Products in Data Mesh

1. Key qualities of an effective data product:
   - Discoverable
   - Addressable
   - Trustworthy
   - Self-describing
   - Interoperable
   - Secure
2. Recommendations for creating data products:
   - Do it agile
   - Straightforward Access
   - Designed with consumers
   - Imperfection
   - Starting small
   - Single platform
   - Reuse & Interoperability
   - Skills

Data Products in Data Mesh

3. Recommendations for maintaining data products:
   ➢ Iterations and Feedback
   ➢ Communication of changes
   ➢ Well defined Lifecycle
   ➢ Re-use
   ➢ Additional Domains
   ➢ Federated computational governance
   ➢ Telemetry and usage information

4. Characteristics of a mature Data Mesh
   ➢ Simple interface
   ➢ Interoperability
   ➢ Straight forward processes
   ➢ Governance and regulatory Standards
   ➢ Access control
   ➢ Enablement of Data consumers through Self-serve data platform
   ➢ Low technical knowledge required
   ➢ Shared ownership and motivation

Data Mesh Optimization

1. Culture
   - Data ownership
   - Data consumption
   - Embracing a data-driven mindset
   - Data enablement
2. Process
   - Agile frameworks
   - Iterative processes
   - Domain-level roadmaps
   - Standard engineering practices
3. Technology
   - Reduce cognitive load on creators
   - Interoperability
   - Simplicity
   - Performance: focus on time to value, seamless experience, and business logic
   - Optionality (avoid vendor lock-in)

Data Mesh Feasibility Pre-Analysis (Organisation Profiles)

When reviewing the principles of Data Mesh, some organisations may be more suitable for an implementation:

1. Large enterprises with multiple business units and a wide range of data sources and consumers.

2. Organisations that deal with sensitive or regulated data, where data ownership and governance are critical.

3. Companies with a strong focus on innovation and experimentation, where fast and flexible access to data is essential.

4. Organisations that have already invested in modern data infrastructure and tools, such as cloud computing and data lakes.

5. Tech-focused companies that are willing to take a more radical approach to data management and are willing to experiment with new ideas and methodologies.

Data Mesh Feasibility Pre-Analysis (Industrial Sectors)

Some verticals are particularly well-suited for a Data Mesh implementation.

Financial Services deal with vast amounts of sensitive data from a wide range of sources.
Benefit: ensure compliance with regulations and data privacy laws.

Healthcare organizations are faced with a rapidly increasing volume of data from electronic medical records, IoT devices, and other sources.
Benefit: turn it into actionable insights that can improve patient outcomes.

Retail organizations deal with large volumes of customer data from multiple sources, including in-store, online, and social media.
Benefit: deliver more personalized and targeted customer experiences.
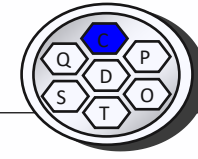
Manufacturing companies generate vast amounts of data from IoT sensors and other sources.
Benefit: optimize production processes and improve product quality.

Technology companies often deal with complex data ecosystems and have a strong focus on innovation and experimentation.
Benefit: enable faster time-to-market for new products and services.

Clients / affected Persons

- Data Producers, Consumers
- Profiles / Identification features (AuthN, AuthZ)
- Client processes
- Service targets
- Installation environment, applications
- Processing / Download times
- Represented systems availability
- Services to be provided by the end user
- Needs, Pain, Requirements
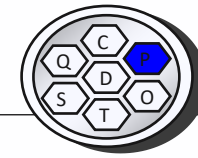
Processes

- Process identification and definition
- Definition of service levels   service depth
                                        bandwidth

- Service level agreements
- Lead and processing times
- Process control   priorities
                                        escalation

- Process measuring points
- Processes include
    - Data Provisioning, Escalation, Incident Mgmt, Downtimes/Maintenance
   - 1st, 2nd and 3rd Level for products and Project Support
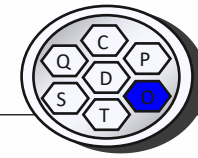
Operational Model / Operating Concept

- Target groups of Data Excellence infrastructure:
  - Tool/app/solution responsibility;
  - operational responsibility.
- Support services of the operational units:
  - client/server hardware;
  - operating systems;
  - communication system;
  - applications;
- Administration of Data Excellence deliverable:
  - Users – admin, consumers, data scientists…
  - Applications, data lakes, DWH, tools…
  - Database(s)
- Data backup/recovery, documentation, archival.
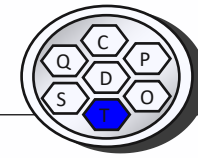
Technical Infrastructure

- Client/server infrastructure and platforms

- System interfaces

- Data Excellence products

- Policy / Model

- Technical Architecture

- Services provided by the customer

- Services provided by 3rd parties
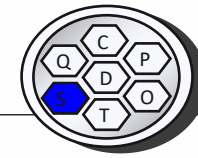
Service Classification

- Documentation of all data products to be supported:
  - List of business cases
  - Use case
  - Pain Points
- Prerequisite for quality assurance:
  - analysis of causes and effects
  - Which services / cases are out of scope
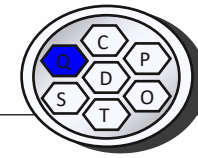
Quality Management

- Determination of the quality policy and its targets

- Cultural change: readiness to service / TQM

- Quality parameter system (QPS)

- Quality control: QPS / SLA

- Reduction of costs caused by errors

- Training measures

- Conceptual introduction of new products

- TCO comparison

Overview

1. **Agile Requirements Engineering**

   - UML / BPMN to model and engineer requirements

2. **Best-in-class Software Architecture**

   - Derivate of TOGAF (Business, Technology, Deployment, Security)

3. **Agile planning, development, and testing**

   - Scrum, SAFe, LeSS, Kanban

4. **Quantitative, metrics-based QA system (ISO / IEC 25010:2019)**

   - Product quality constantly monitored; findings incorporated into sprints

5. **Secure and professional deployment and operations**

   - Automated, repeatable and auditable

Standards & Practices

Ultra Tendency complies with the following standards and practices:

- **ISO 27001**   Information Security (all offices are certified)

- **TISAX**   Information Security (all offices are certified)

- **ISO 22301**   Business continuity (Organization is certified)

- **ISO 9001**   Quality Management (H1 2023 certified)

- **ISO 31000**   Risk Management (is applied)

- **TOGAF 9.2**   For all architectural modeling

- **BPMN / UML**   Requirements engineering

- **DMBoK 2 / ISACA DMM**   Data maturity assessments

- **Project management**   PMI, HERMES, PM2, CRO, CCRO

The UT Approach to Software Architecture

Business Architecture

Targets business stakeholders and communicates use cases via UML use case diagrams or BPMN process diagrams

Technology Architecture

Targets system architects and software developers by providing detailed component diagrams that explain how a system is composed out of software components

Deployment Architecture

Targets developers and system administrators, and describes how logical software components are deployed to physical or virtual infrastructure nodes

Security Architecture

Targets CISOs and Security specialists and describes in detail how the system complies with general and corporation-specific security requirements

Definition of Done (DoD)

Very important in Scrum is the Definition of Done (DoD). It defines when a user story is done. The Definition of Done includes the following criteria:

- Code that implements the user story is fully developed and compiled error-free.

- Unit tests are written and provide test coverage of at least 80%.

- The branch, in which the user story was developed, was integrated into the higher-level version branch

- Checkstyle and Findbugs run without errors through the new code

- SonarQube analyzes used libraries and ensures that no copy-left-licensed libraries were used

- The technical details of the user story are documented in a Wiki (e.g. Confluence)

- The increment is deployed to the test environment and is available to the customer for review.

- Upon acceptance, the increment is deployed to high stages, such as pre-prod and prod

If all user stories in the sprint meet the DoD, the increment resulting from the sprint can be prepared for deployment in production. For the customer, this means that a new product version is available for production at the latest after each sprint (e.g. after two weeks).