TNO innovation for life

Maastricht University

DITECT

PLOUTOS

DATA STANDARDS, DATA SHARING AND FOOD SAFETY:

USING FAIR DATA PRINCIPLES IN A GLOBAL FOOD SYSTEM

| CHRISTOPHER BREWSTER

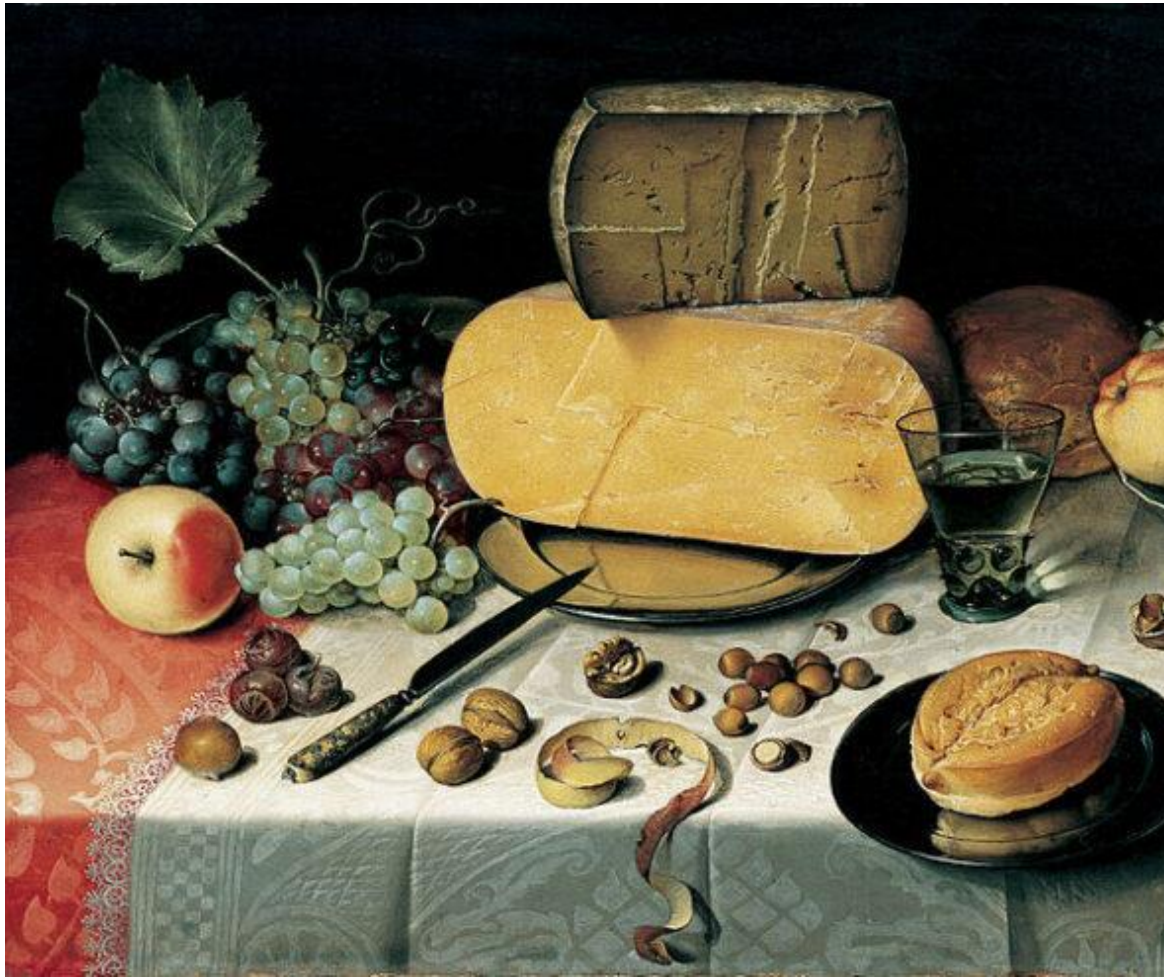
25 November 2022

DRIVERS - PAST AND PRESENT REALITY OF FOOD SAFETY

- **Past reality** was (mostly) short supply chains, with face to face interactions
  - Not a utopia as **The Laws of Manu** (*Manusmṛiti*) and many other regulations in history show.
  - Friedrich Accum (early 19th century) demonstrated the importance of data and the need for data for food safety
- **Present reality:** Ordinary people go shopping for food, and **trust** the food is “fit for consumption”, safe and unadulterated.
  - People trust/or believe that that our regulatory framework and retailor self-interest is enough to ensure this.
  - **BUT:** in reality there is a continuous struggle to ensure **quality, safety and absence of food fraud** (both by regulators and food companies)
  - The business models around food production and retailing encourage food fraud and adulteration -- but that is a topic for another day

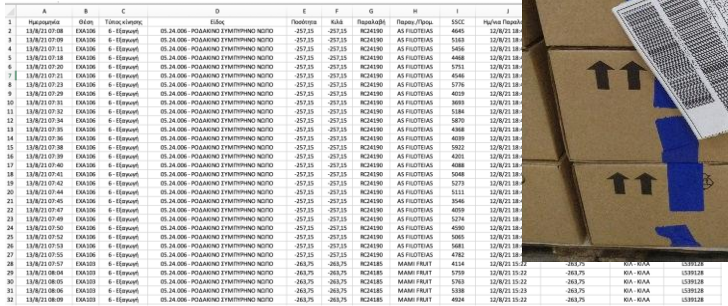
DRIVER - DATA SHARING FOR FOOD SAFETY/INTEGRITY

- Assumptions
  - There is information/data that can help identify an actual or emerging risk
  - This information/data could be shared along the supply chain/web and analysed to provide early signalling
- Two fundamental types of data
  - **On-chain data** i.e. tracking/tracing data, location, processing steps, environmental parameters etc.
  - **Off-chain data** – mostly laboratory analysis data, certification data, historic data etc.
- **The challenge lies in getting hold of that data!**



DRIVER - AN EXPLOSION OF DATA – ON CHAIN

- The gradual move from paper-based one-up, one-down records to extensive electronic records
  - **Automatic data capture**, via barcodes, RFID, IoT, integrated into ERP systems (e.g. Returnable Transport Items with RFID or even IoT devices)
  - **On the farm:** weather stations, sensors in the soil, on livestock, in machinery, from drones and Earth Observation.
  - **In the food processors:** ERP systems capturing GS1 EPCIS events, processing, packaging, transforming
  - **In the logistics steps:** GPS tracking, sensors monitoring (e.g. temperature, humidity)

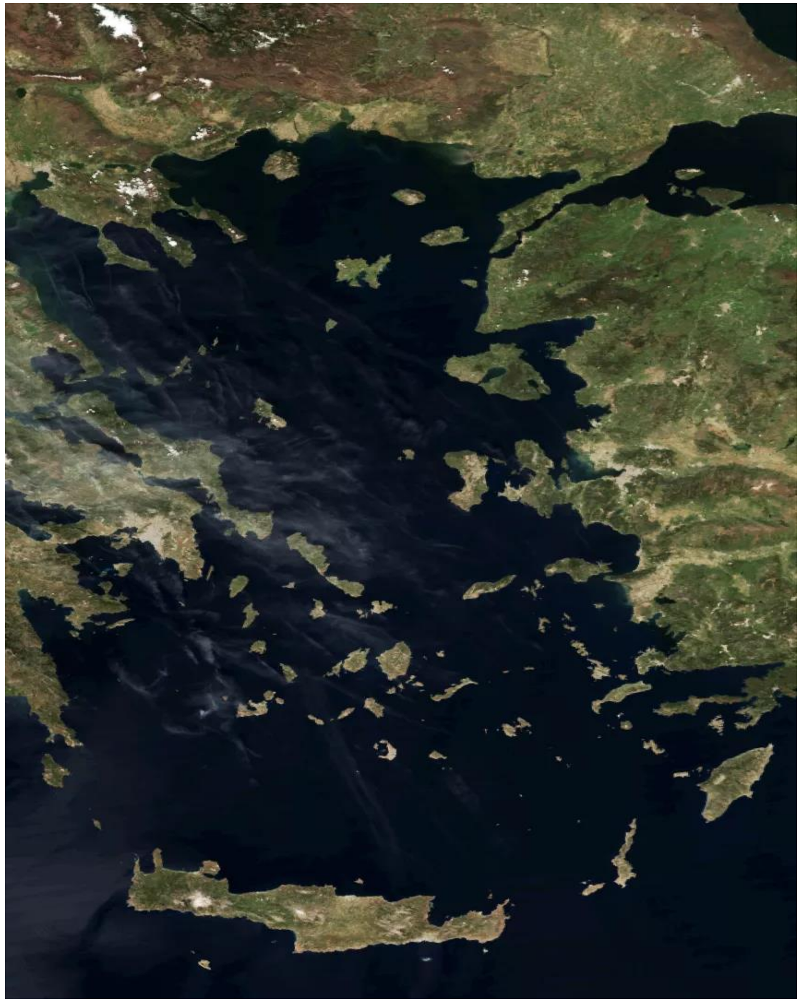


DRIVER - AN EXPLOSION OF DATA – OFF CHAIN

- Two areas:
- 1. **Food safety data collection**
  - Traditional methods i.e. **sample based microbiological** laboratory analysis (“finished product testing”)
  - More recently **non-destructive methods** – so called “process analytical technology” (Nychas et al. 2016) (e.g. FTIR spectra)
  - Lots of mathematical modelling, growing use of **machine learning and deep learning** (Nychas et al. 2021) – <u>**LOTS OF AI HERE**</u>
  - Here we are not concerned with limitations/ capabilities of these approaches only to say **they generate a lot data – but where is it? Who has access?**
- **2, Food certification**: for quality features of food (GlobalGAP, organic, Fairtrade, and many other certification systems)
- **In all these cases data resides with different organisations and institutions but is not available to the community.**

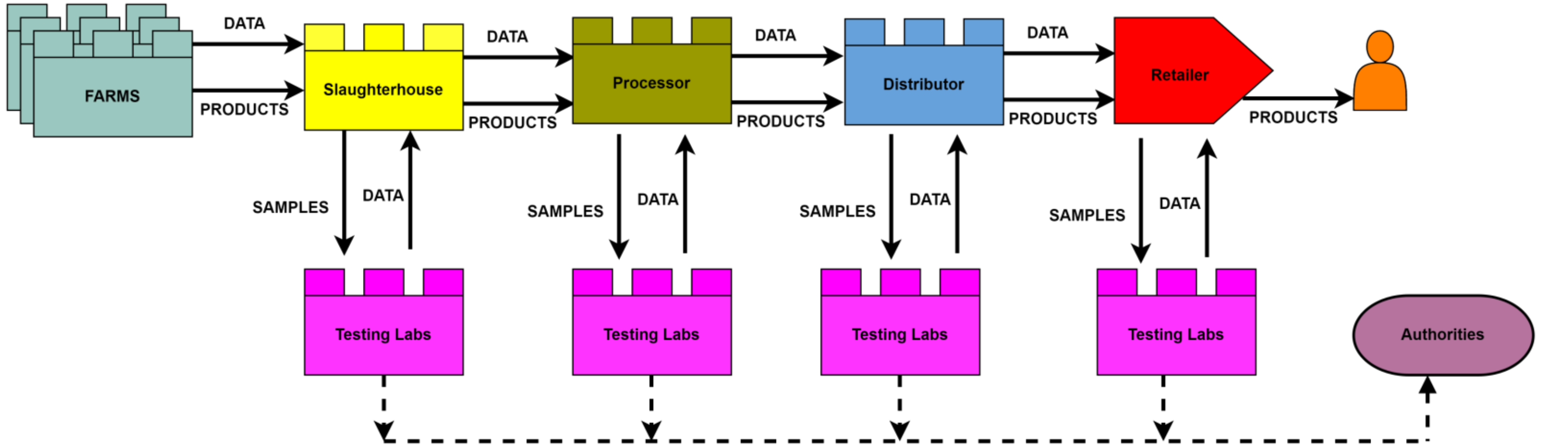
BARRIERS - KEY CHALLENGES ACCESSING THE DATA

- **Lack of data standards**
  - Each instrument produces different data in different formats
  - No common data model or “ontology”.
- **No commonly agreed architecture** for sharing data across multiple stakeholders (food producers/processors, and food analysts)
  - Essential to allow stakeholders to control data
  - BUT also allow data to be shared for contractual or research purposes
- **Socio-technical challenge:**
  - Lack of culture that data sharing is a good idea
  - Lack of understanding of the benefits
  - Lack of principles – governance, data ownership, permitted reuse etc.



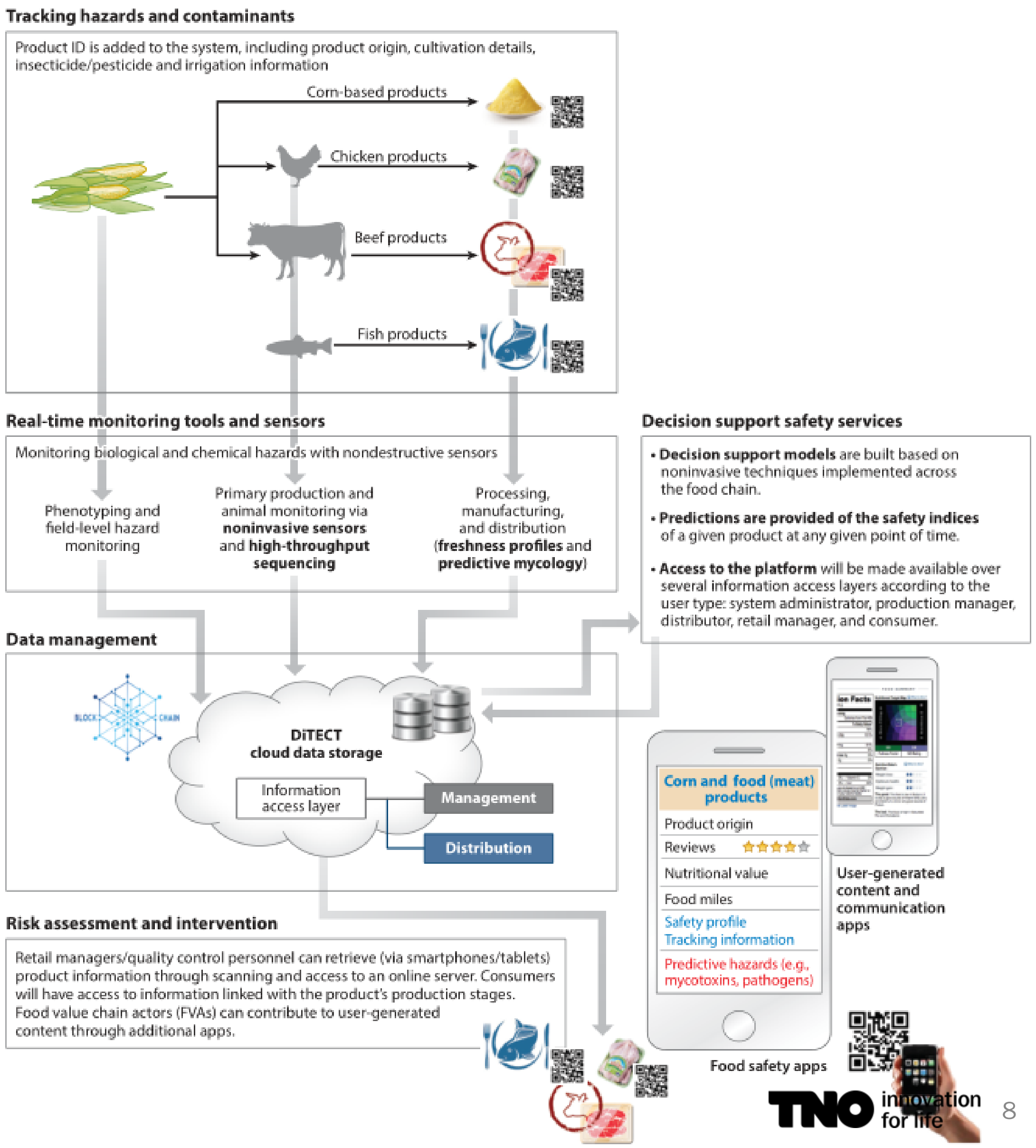
BARRIERS - SILOS IN SUPPLY CHAIN + FOOD HAZARDS NETWORKS

- Great resistance to data sharing in food supply chains/webs – primarily fear third party willuse data against the stakeholder whether competitor or government agency
- Add the complication of food laboratories (research and commercial)
- More data gets siloed – partly through privacy/confidentiality, partly lack of infrastructure for data sharing
- AND each laboratory/instrument does things differently



IMPLEMENTATION - DITECT PROJECT VISION
[HTTPS://DITECT.EU/](https://ditect.eu/)

- “DIgital TEChnologies as an enabler for a conTinuous transformation of food safety system”
- H2020 project collaboration between EU and China for sustainable food security call with 21 EU and 13 CN partners. Started Nov 2020.
- Objective is to “develop an integrated framework for real-time detection, assessment, and mitigation of biological, chemical and environmental contaminants throughout the food supply chain”
- Project is large with lots of activities in four pilot areas on corn, chicken, beef and fish. Focus is on the use of non-destructive real-time monitoring techniques
- BUT it is a research project – so from a data perspective we are (initially) working with a centralised approach.



DITECT APPROACH TO DATA 1. – FAIR DATA PRINCIPLES AND TOOLS

- The FAIR Data Principles
  - Origins in health/life sciences,
  - Data is often commercially, or ethically sensitive hence "accessible",
  - Data often needs to be reused (or that is the ideal)

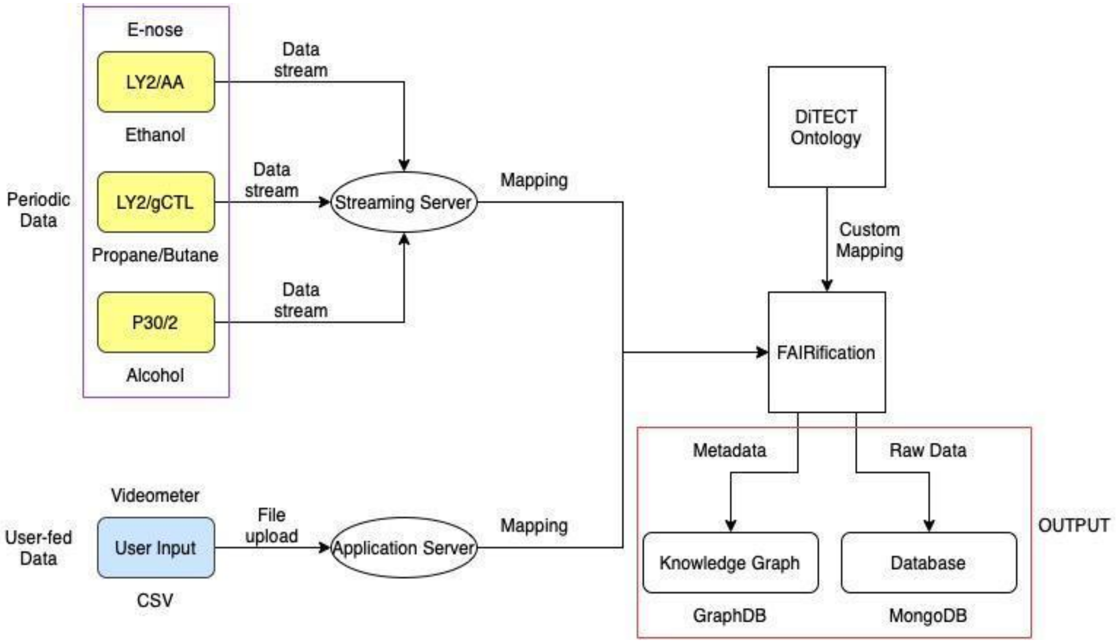
    FINDABLE – "Machine-readable metadata are essential for automatic discovery of datasets and services"

    ACCESSIBLE – "How can the data be accessed?" Metadata specifying access, authorization and authentication

    INTEROPERABLE -- Data (and metadata) uses standard formats and vocabularies/ontologies

    REUSABLE– Metadata makes clear usage rights in standardised format

  - A FAIRification framework to extract, transform, attach rich metadata, and load sensor output to a central repository (for now).
  - Part of challenge is to build a suitable data model – an ontology – to annotate the data.

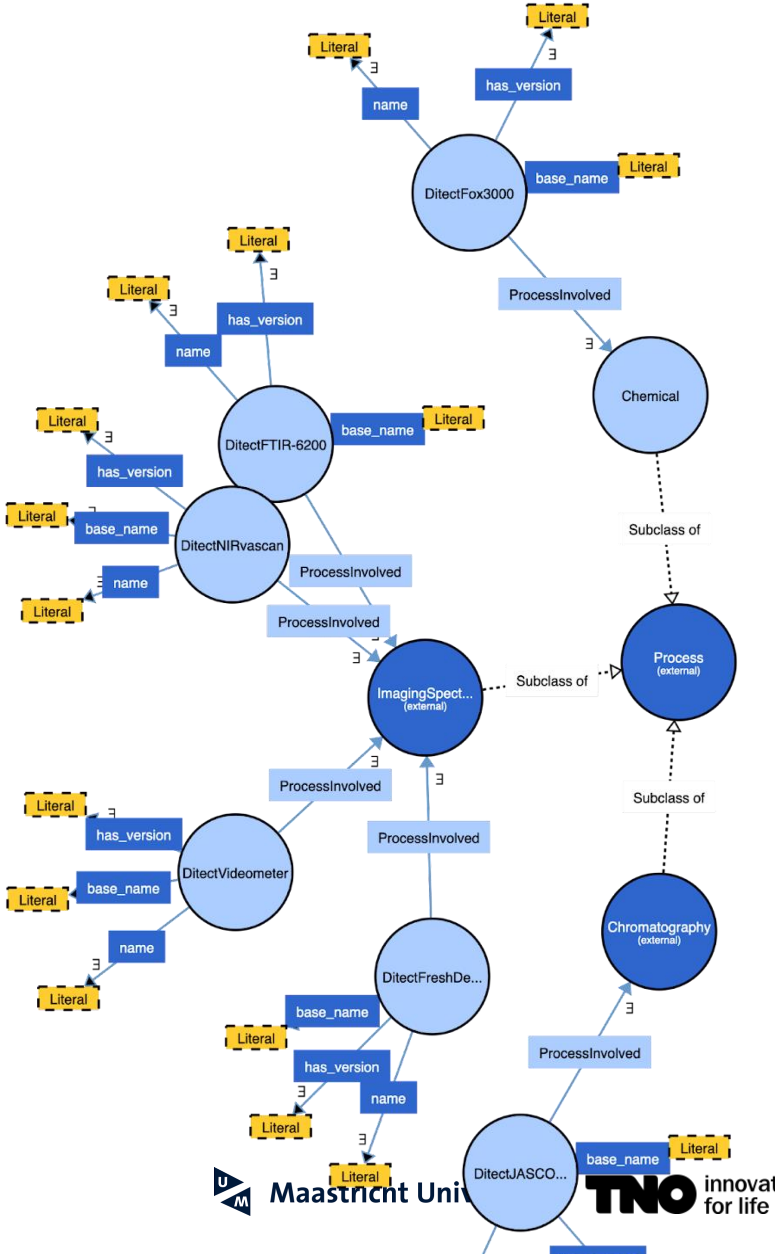


Data ingestion and FAIR-ification process – adding metadata to data sets

DITECT APPROACH TO DATA 2: DATA STANDARDS FOR METADATA

AN ONTOLOGY FOR FOOD SAFETY

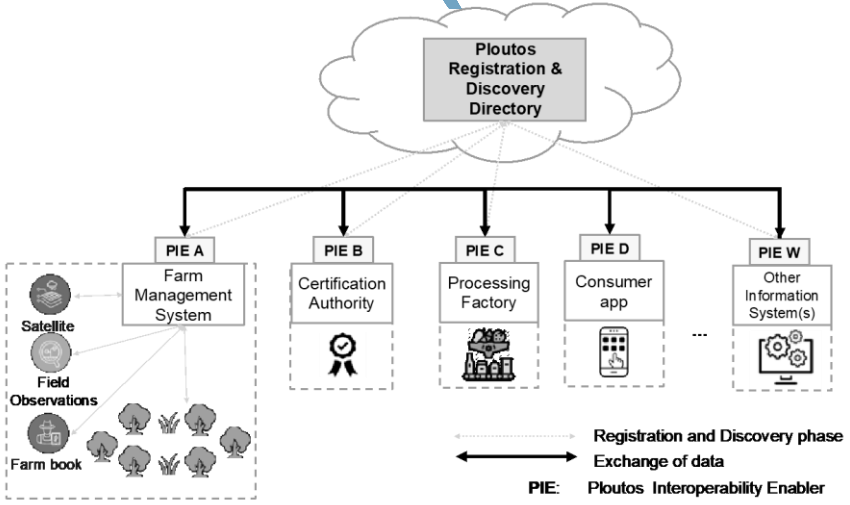
- We need a data standard which will provide the metadata to describe each data point
  - Each step in the supply chain (of different products – for DiTECT four pilots – but needs to be generally applicable)
  - At each step an analysis may be undertaken
  - At each step the different kinds of sensors used, and all the data parameters involved.
  - Complex process because many sensors/types of laboratory tests
  - Completely different ways of representing the data
- A formal (OWL language) ontology has been developed using concepts/classes from existing widely used standards/ontologies, such as: AGROVOC, SSN/SOSA, SAREF, FOODON, etc.

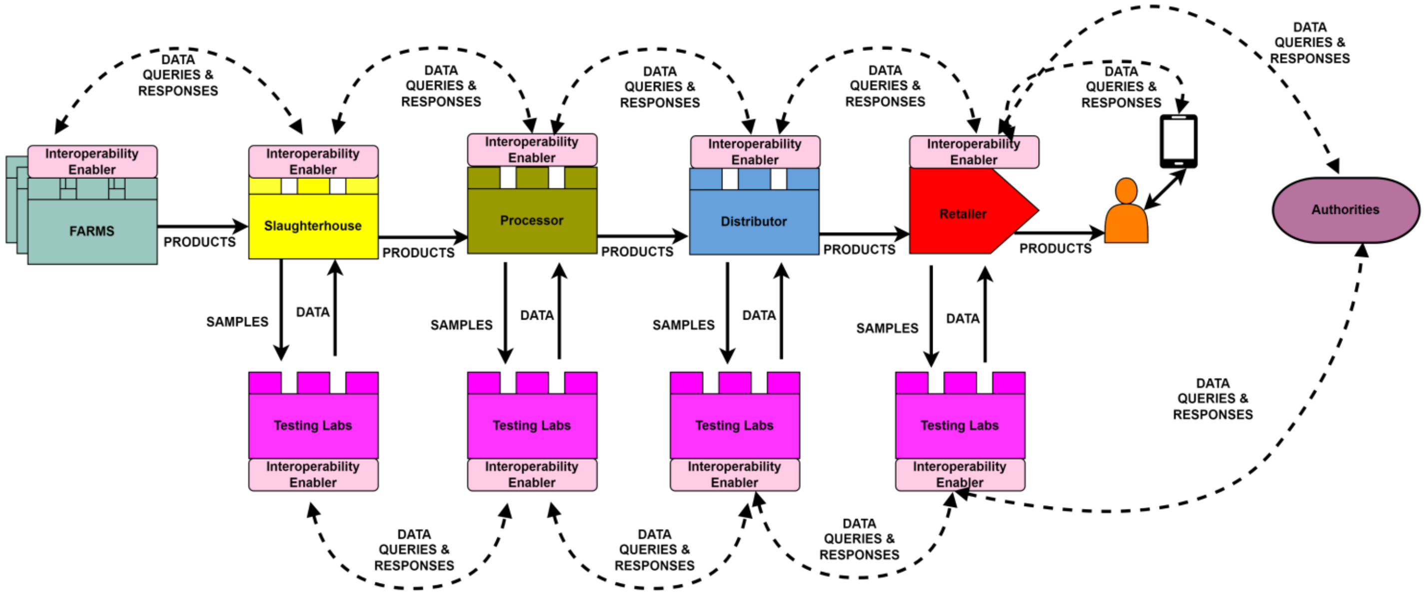


Maastricht University TNO innovation for life

DITECT ARCHITECTURES FOR DATA SHARING (THE PLOUTOS APPROACH)

- Building on the work of the Ploutos project (https://ploutos-h2020.eu/) developed for data sharing across a supply chain.
- An architecture which enables controlled and technically sound flow of data among the various information providers and consumers **without at the same time disturbing the current operations of the underlying systems.**
- **Intended to work with legacy systems using "interoperability enablers" and lots of <u>semantic technologies</u>.**
  - **An Ontology – Reasoning – querying across distributed networks**
- In DiTECT we are designing such an architecture for food safety sector.





Maastricht University TNO innovation for life

POLICY IMPLICATIONS

- Technical work is not that difficult
  - The community needs to adopt the FAIR data principles ← **needs pollical and financial incentives**
  - Data standards (ontologies for metadata description) need to be agreed or widely adopted ← **regulations help ….**
  - An architecture for distributed data sharing needs to be seen as a priority – that takes into account the FAIR data principles, and respects access control/security issues ← **again regulatory impositions help, as well integration in agrifood data spaces initiative**
- The real challenge is to change a culture – so part of the work in DiTECT (in the data management WP) is to **develop guidelines and best practice advice.**
  - Changing a culture to make data sharing with collaborators, with the food safety authorities, and with researchers normal will be a major achievement (**and may take time**). ← **not to be left to the commercial sector alone**
- Integrating both supply chain data (on chain data) with off-chain food safety data should be the eventual objective..

IMPACT: ABSENCE OF TRUST IS EXPENSIVE → ENERGY AND ENVIRONMENTAL COST

Economic

- + more efficient management of food hazards/crises
- + Digitalisation may reduce *individual* stakeholder costs and risks
- ------------------------------------
- - Greater overall cost to the sector
- - Probable barrier to small entrants in food system

Social

- + Greater trust if fewer food crises arise
- ------------------------------------
- - More dependence of digital infrastructure – computer said so syndrome
- – danger of deskilling citizens and professionals
- - Sticking plaster over a problematic globalised agrifood system – may be solving the wrong problem

Environmental

- + Reduce food waste or avoidable losses
- ------------------------------------
- - Infrastructure and modelling has a scope 2/3 energy costs
- - May lead further consolidation/reduction in food choices – because only certain items are in the system.

Jan Bruegel the Elder - Temptation in the Garden of Eden

THANK YOU

CHRISTOPHER.BREWSTER@MAASTRICHTUNIVERSITY.NL

CHRISTOPHER.BREWSTER@TNO.NL

TNO innovation for life

Maastricht University