

Sustainability of Linked Open Data – A Key Challenge for Agricultural Applications

Dr. Ansgar Bernardi
DFKI GmbH, Kaiserslautern

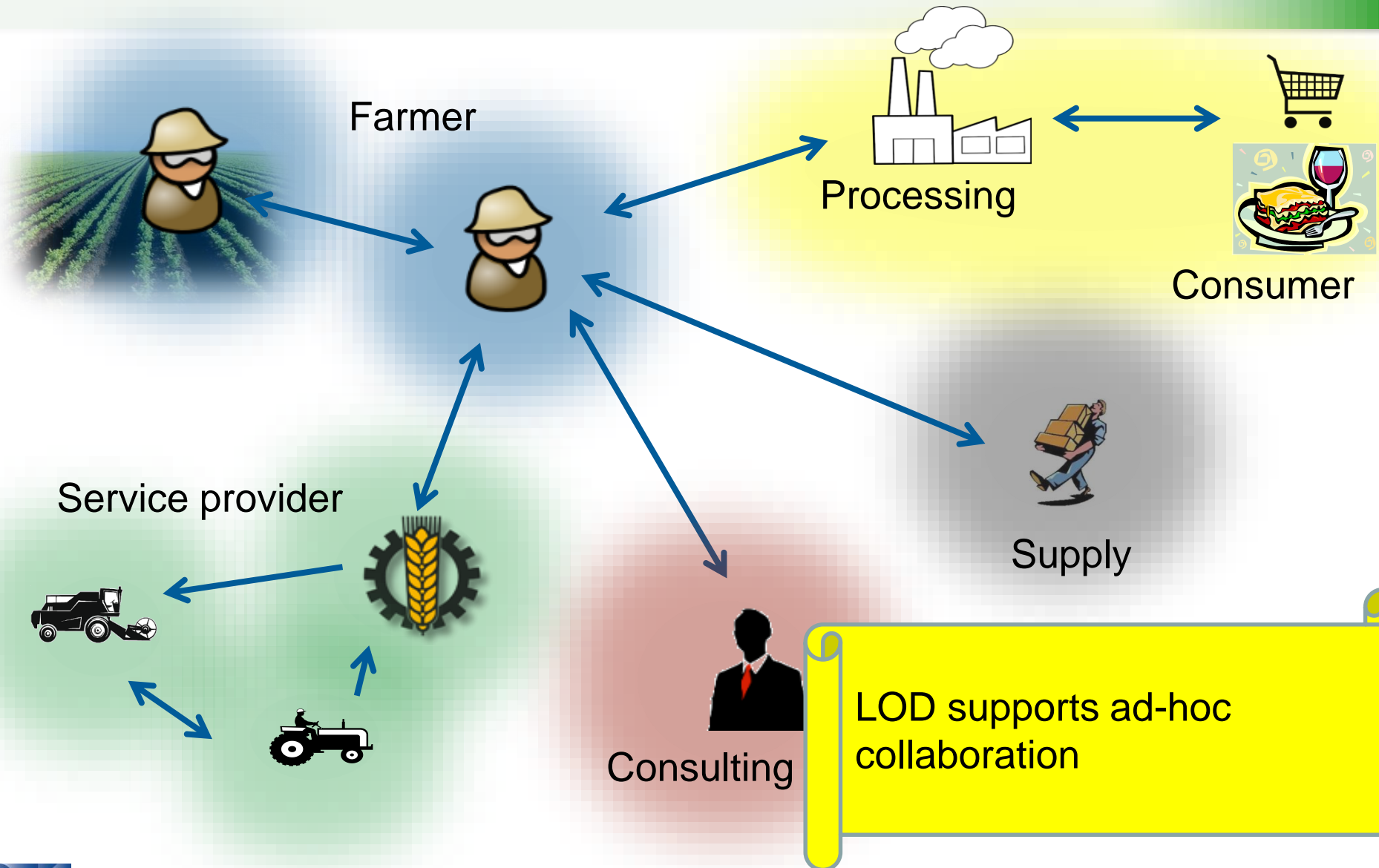
LOD: Goal

- ◆ **Data sources shall be usable**
- ◆ **Let arbitrary services read / use the data**
 - Traditionally, data are tightly coupled with „their“ application services
- ◆ **Make data machine-readable**
 - Machine readable \neq HTML!
 - **RDF**: (universal) data format
 - **HTTP**: Transfer protocol
- ◆ **Make data machine-understandable: Ontologies!**
 - RDF Schema
 - Re-use available vocabularies / ontologies
- ◆ **Connect data**
 - ... To support the need of YOUR service

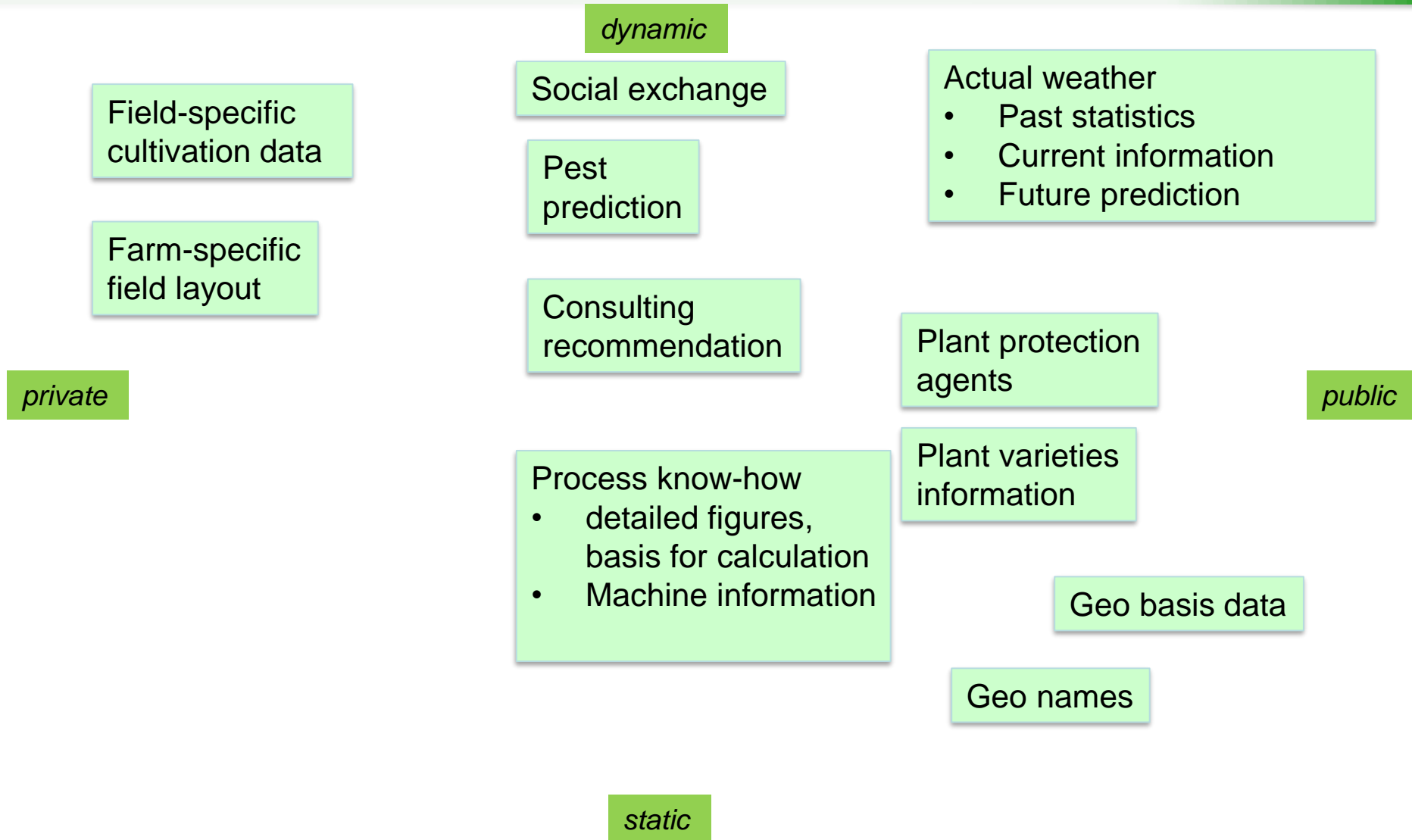
◆ “Get your data out”



Agriculture with its collaborative structures can ideally profit from linked open data solutions



Relevant data have different dynamics and origin



LOD Solutions – Community Interaction Without Formal Agreement

- ◆ Data owner
 - **Publish** data
 - Update changing data

- ◆ Solution developer
 - **find** multiple data sources
 - **Access, Understand, Adapt, Integrate** data
 - Build service

- ◆ User
 - Access actual data while using the service

Example: SmartRegio (2015-2017) – use (linked) (open) data for regional profiles

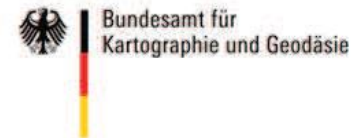
open data sources



closed data sources



public



EnergyMap.info

private



Observations & experiences show difficulties...

- ◆ Data collection seems cheap, but interpretation is a challenge!
- ◆ Finding relevant data is not trivial
 - few open data portals or search engines (*SINDICE* & *SIGMA* disappeared)
 - Raw data not published, only specific condensations
- ◆ Integrating multiple data sources is difficult
 - Different formats & structures
 - Different representation means
 - Different granularity

... and suggest caution!

- ◆ Data suppliers are not aware of their own data, the challenges of distributed settings, or data quality
 - (usually) No structured overview
 - Competent seniority not available
 - „there is no correct geo data“
- ◆ Organizational and legal aspects are a crucial challenge
 - Who is responsible?
 - Reliable anonymization is difficult!
- ◆ Formats, interfaces, access conditions change any time!

LOD is shaky ground for reliable services

- ◆ Individual LOD sources offer no guarantees for accessibility, maintenance, or long-term availability
 - Individual availability, interfaces, access conditions change any time
- ◆ Data actuality, correctness, reliability are questionable
 - Dynamic changes and community developments
 - Who offers guarantees for correctness? Damages possible!
- ◆ Well-balanced business models difficult
- ◆ Understanding multiple data sources is difficult
 - Semantic grounding (e.g. ontologies) can help, but not common place
 - Can the service verify the correctness of data?
- ◆ Data sources can drop out temporarily or permanently
 - Agricultural applications in the field need to cope with unreliable networks

Suggested solution approach principles

- ◆ Dedicated organizations for crucial basis information
 - Reliable publication & maintenance with guaranteed availability
 - Data models oriented towards agricultural use cases
 - Data collected/managed by public institutions: “guaranteed” LOD as public obligation
- ◆ Clear maintenance processes for dynamic data
 - Continuous adaptation and development
 - Clear versioning concepts
 - Collaborative public contributions
 - Control mechanisms to avoid spam and errors
- ◆ Private data need balanced business models
 - cost-vs-benefit of LOD
 - Public funding
 - Pay-per-use (not “open” anymore!)
- ◆ Employ local caching wherever possible
 - guard against temporary un-availability.
 - Maintenance procedures shall manage asynchronous parallel development.

Thank you for your attention!

Dr. Ansgar Bernardi
DFKI GmbH
Trippstadter Str. 122
67663 Kaiserslautern
Tel. 0631-20575-1050
E-Mail: ansgar.bernardi@dfki.de



SmartRegio was funded by



**Bundesministerium
für Wirtschaft
und Technologie**

aufgrund eines Beschlusses
des Deutschen Bundestags