



Universiteit Utrecht

Using iRODS to manage, share and publish research data

Ton Smeele & Lazlo Westerhof
ITS/ResearchIT, Utrecht University



Agenda

Profile Utrecht University

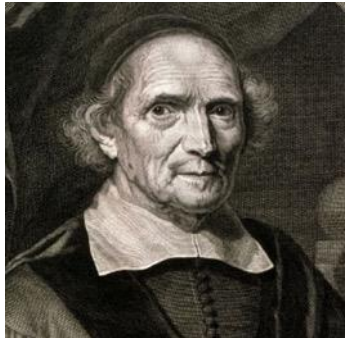
Yoda introduction and concepts

Demonstration

Challenges, issues & lessons learned



Organisation & people



ESTABLISHED

1636

PROFESSORS

550
Incl. faculties
Medicine

FACULTIES

7+2
teaching
institutes

STAFF-MEMBERS

6,960

STUDENTS

30,523



Top ranking



NOBEL PRIZES

12



SPINOZA PRIZES

15



SHANGHAI RANKING 2017



The Netherlands

1



Europe

13



World

47



4 Strategic themes - focused research



DYNAMICS OF YOUTH

- Integrating Utrecht expertise on youth development, from synapse to society



INSTITUTIONS FOR OPEN SOCIETIES

- Cooperation, Self-regulation and Collective Action
- Sustainability and Resilience
- Innovation and Economic Growth
- Equality, Inclusiveness and Social Mobility
- Democratic Governance, Citizenship and Trust



LIFE SCIENCES

- One Health
- Personalised Medicine & Health
- Regenerative Medicine & Stem Cells
- Science for Life



PATHWAYS TO SUSTAINABILITY

- Towards Industry with Negative Emissions
- Future Food: Pathways towards Healthy Planet Diets
- Transforming Infrastructures for Sustainable Cities
- Water Climate & Future Deltas



Why iRODS as Research Data Management platform

- **scalable platform**

- can manage billions of files, petabytes of data
- infrastructure/vendor neutral solution

can be used to manage large/many data collections

- **enforces data policies**

- secures sensitive data
- auditable controls

supports demonstrable research integrity

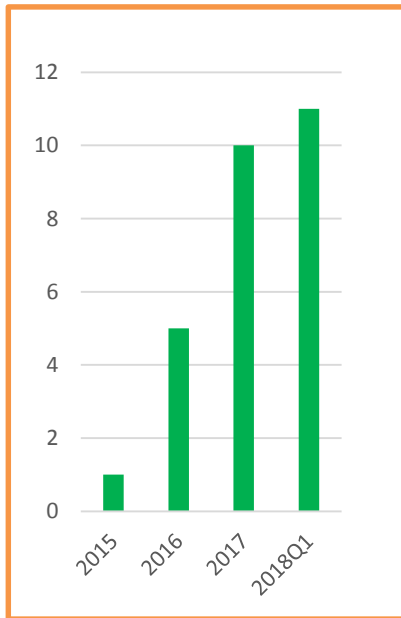
- **manages metadata alongside the data**

- metadata based data policy execution decisions
- data workflow automation

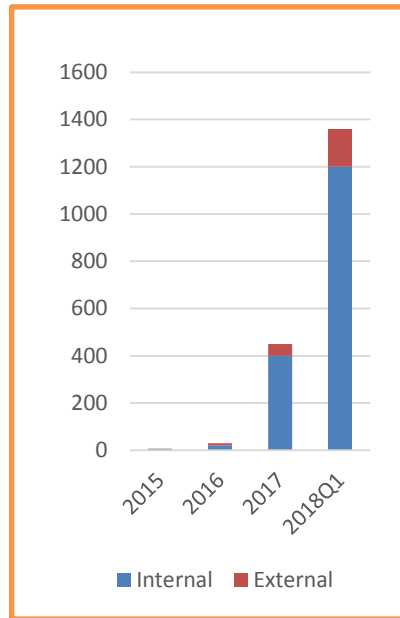
facilitates research data workflows



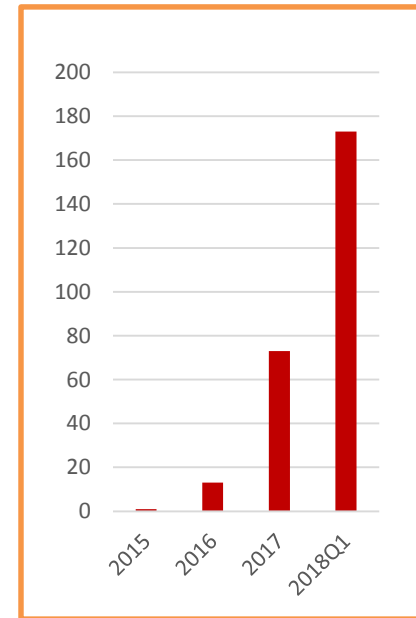
Utrecht University iRODS managed research data



11 Zones



1400 Users



180 TB Data

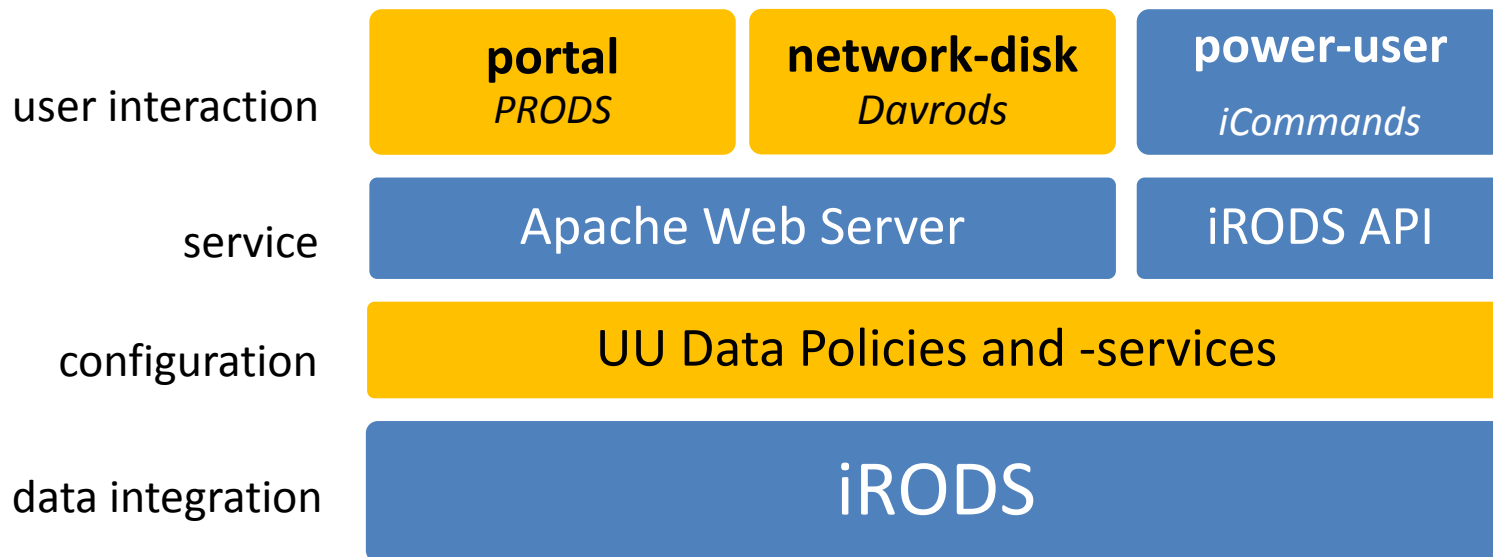
production instances only, figures are indicative



Our iRODS implementation is called "Yoda":

preconfigured iRODS based system, delivered and supported as a service

- enhanced with (graphical) user interfaces, policies and rules

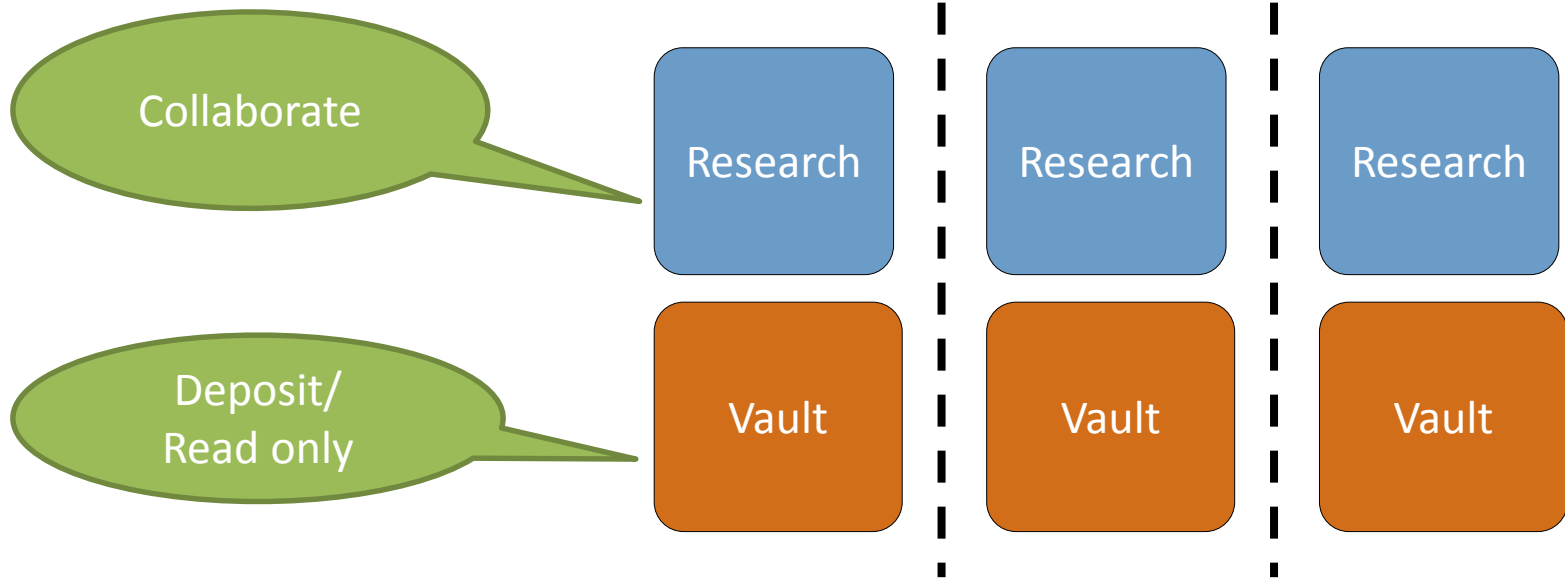


10,000 lines of rules

14 custom microservices



Yoda Data compartments



Each data compartment relates to an iRODS group

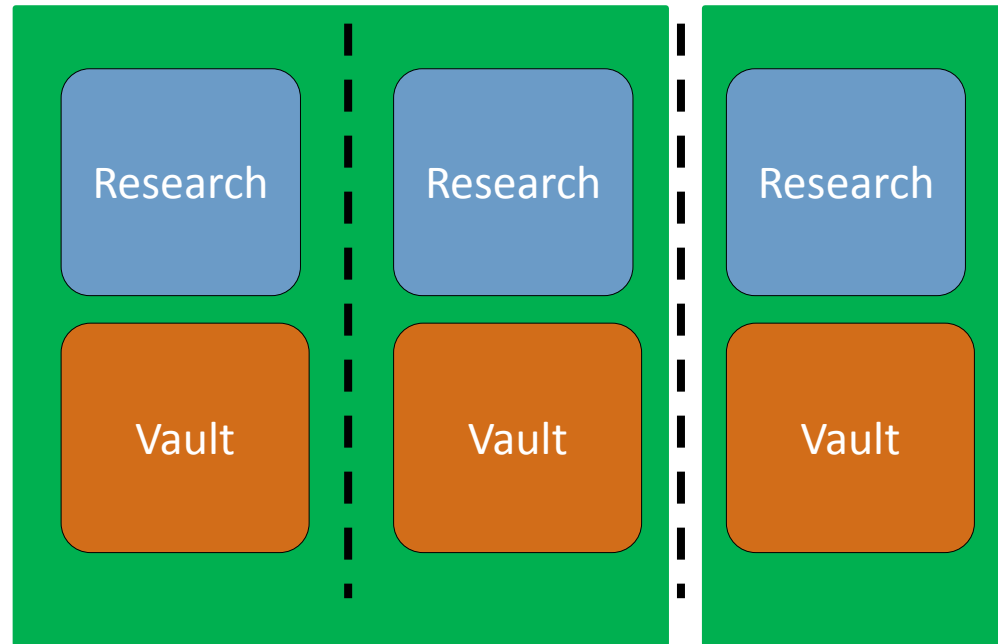


Yoda Communities ("category")

A community comprises of multiple data compartments

Per community:

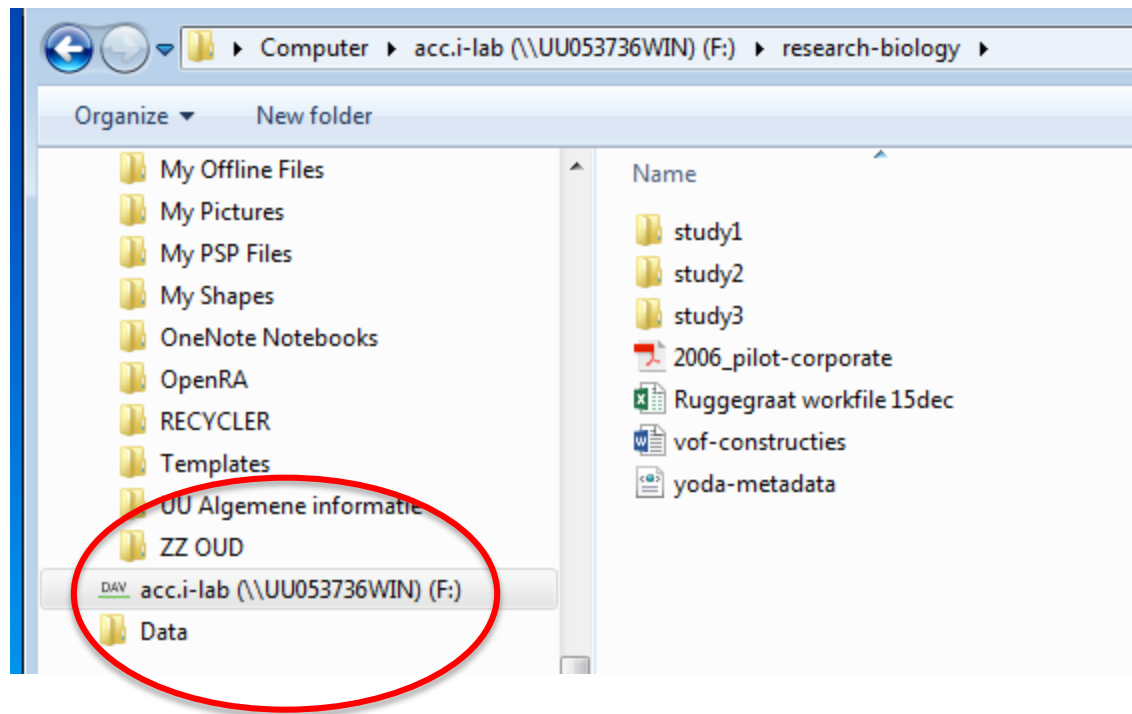
- cost calculation/invoicing
- appointed datamanager(s)
- metadata schema



Community concept implemented as metadata on iRODS groups



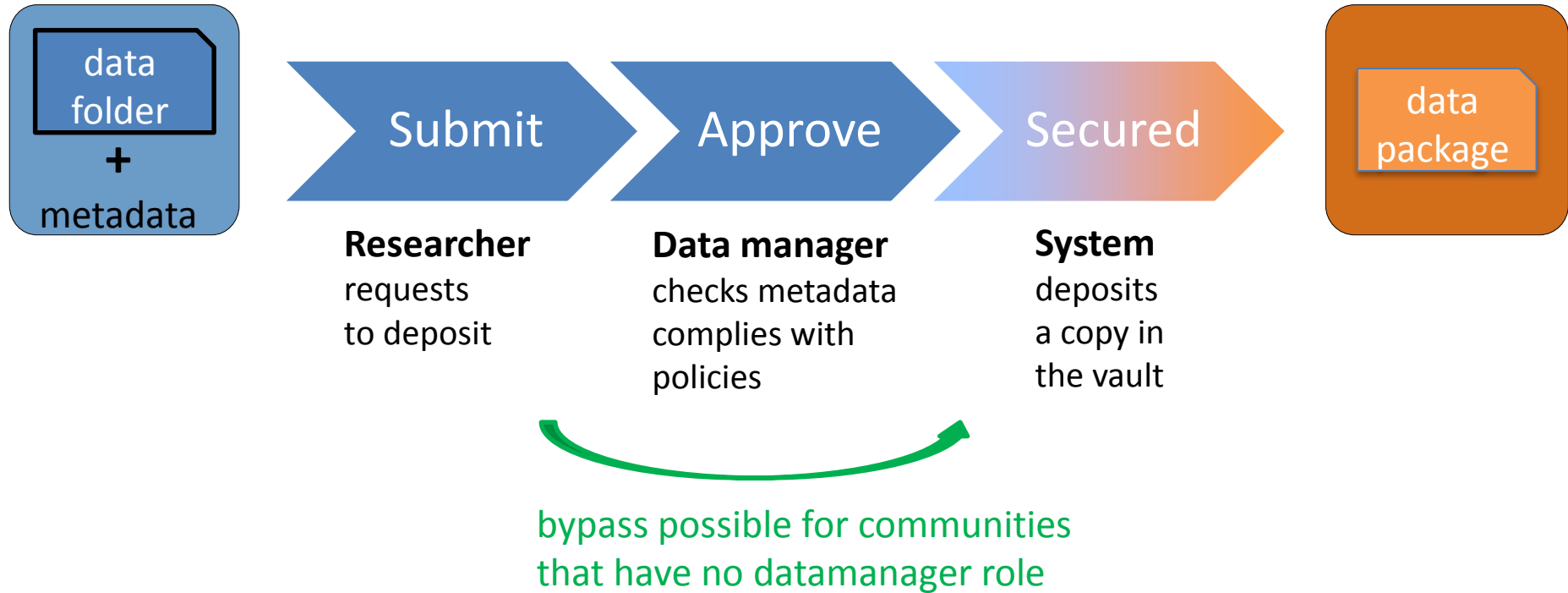
Collaborate during research via the Yoda disk



WebDAV access from anywhere on any workstation using Davrods

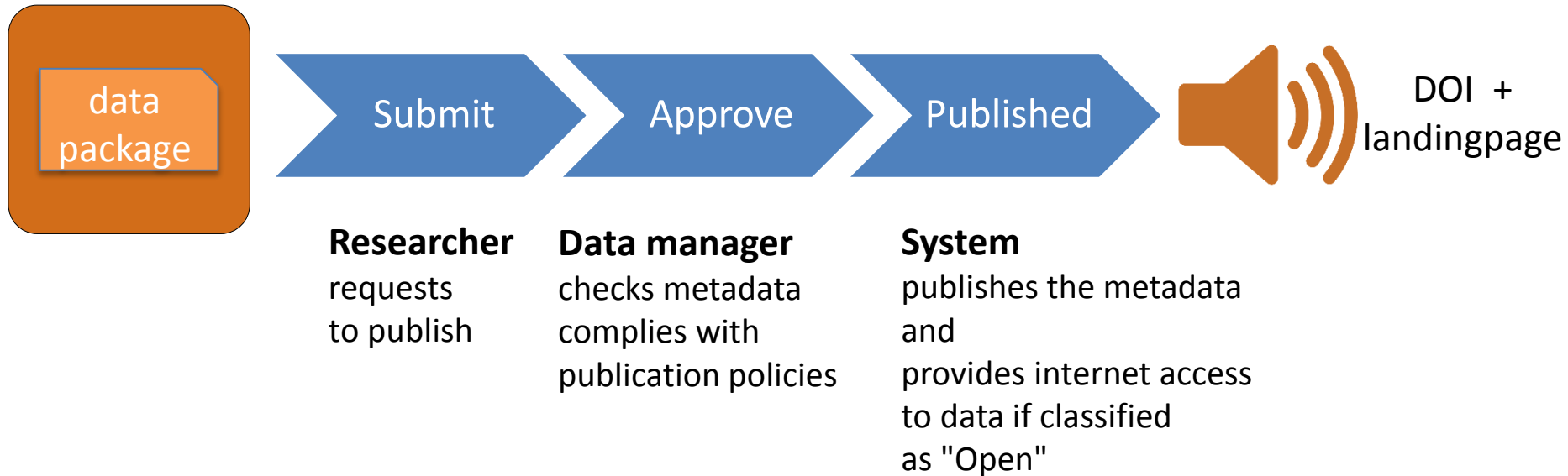


Data Deposit workflow





FAIR Data Publication workflow





'FAIR' Research Data Management using iRODS



Collaborate safely as a group ("Research" folder)



Maintain **integrity**, deposit a folder in the vault



Allow FAIR **reuse**, publish a data package





demonstration



Challenges, issues and lessons learned

- **Metadata form interaction with browser: was XML now adopting Json**
- **iRODS 4.1.11 stable and reliable except for delayed rules engine
(resolved in 4.2.2+)**
- **many components and architectural layers, need to simplify implementation and configuration**



Yoda manages data during/after research



Collaborate safely as a group ("Research" folder)
-> membership self-managed by researchers



Maintain integrity, deposit a folder in the vault
-> metadata can vary per community,
-> datamanager approves deposit



Allow FAIR reuse, publish a data package
-> datamanager approves publication, DOI citable data



Yoda is available under GPL license
at <https://github.com/UtrechtUniversity>

Thank you

More info:

Ton Smeele

a.p.m.smeele@uu.nl

Lazlo Westerhof

l.r.westerhof@uu.nl