

Teilnehmer Punkte I

Was wurde kommentiert/angeregt?

Teilnehmer Punkte I

- WG/Igs auf dem Wege Trust zu verbessern
- DOIs vs. PID/Handles, DataCite hat MD Anforderungen, die weiterreichen
- Hat nichts über Lizenzen gehört
- helfen Resultate um Prinzipien umzusetzen, sind sie in Audits eingebunden
- Produkte sind noch nicht reif für Produktion
- Wie könnte man Dinge zu Services machen – dringende Frage
- Bibliotheken sind nicht die Orte wo DM gemacht wird
- EUDAT hat bereits service für jeden, der DFT Model umsetzt
- AG Primers – irgendetwas in a nutshell
- Erfolgsgeschichten dringend wie passen Komponenten zusammen etc. inkl. Zeichnungen

Teilnehmer Punkte II

- CLARIN/Thorsten
 - SSO, zugreifbarkeit, verteilte dienste,
 - Datenpersistenz – archivierung
 - Oai-pmh harvesting, joint suchen
 - Workflow infra
 - Modulare metadaten
 - Paper Package (hat PID, set von Daten das zu Paper gehört, inkl. Metadaten, R scripts etc)
 - Wer darf packages versionieren, wann wird package registriert, veränderungen führen zu neuen PIDs
 - Arbeiten an CLARIN zugeschnittenen DMP tool
 - Virtual collection builder tool
 - RDA
 - Hilfe bei DMP

Teilnehmer Punkte III

- Humanities/Kai
 - Infra für daten in humanities (management, curation, stewardship)
 - Viele datentypen
 - Komplexe Strukturen, Anwendungen halten oftmals Daten zusammen
 - Common trends ok, layers of techno nicht ok wegen anwendungen
 - Im moment keine WFs, federations sind schwierig wegen overhead
 - Open data in hum oft nicht möglich
 - Simples model einfodern durch service angebot, kurzfristiges funding
 - Machen training on DMP
 - Oft fehlt geld um MD, semantics etc zu registrieren
 - Haben info system über daten
 - Haben viele praktische probleme – kein AAI, kein archiv, RZ beschissen
 - RDA
 - RDA angebot argumentatives futter
 - Kooperationen fressen zeit
 - Nachhaltigkeit
 - Unterstützung der praktiker vor Ort

Teilnehmer Punkte IV

- Re3data/robert
 - Repreg global and across disciplines
 - Haben schema, Angaben werden reviewed
 - Re3data wird eine AG von dataCite
 - RDA
 - Plattform für Austausch global und coordination
 - Wie kann man konkret komponenten verwenden
 - Koordination mit industrie, government, W3C,
 - Verweis auf BP von W3C, IETF, etc

Teilnehmer Punkte V

- Bibl/Jens
 - In Breite vertreten, MD expertise, technologie wechsel Erfahrung, sichtweisen wandeln,
 - Laut HRK sollen sich in DMPs engagieren
 - Sammeln, aufbewahren, verfügbar (Texte, eRessourcen) – transformieren dies auf FDaten
 - Museum für Papier?, Unterscheiden kaum zwischen Datentypen, wie kriegen wir Forschungsdaten in existierendes Schema
 - Optionen: Basis-Infrastruktur, Spezialisierung auf best. Disziplinen und dann erfolgreiches Datenzentrum, Eintreten in forschungsprozess
 - RDA
 - Resultate vermitteln, nicht verwenden
 - Kooperation und fortbildung ist nötig (Angebot via LIBER?)
 - Seminarreihe?

Teilnehmer Punkte VI

- Geo/Peter
 - Coverage entsprechend OGC, arbeiten mit Datenwürfeln (Cubes)
 - Haben earthserver
 - Mengung der Daten von verschiedenen Zentren, großes Problem Übertragung
 - Welches Integrations Paradiga (XML, JSON,
 - RDA
 - Einheitlicher Info Raum
 - Zusammenbringen von abstrakten Datentypen (graphen,
 - **Referenzen innerhalb DO**
 - Data/MD divide aufheben
 - Schnittstellen basiert auf formal languages
 - Trust und/oder fitness for purpose
 - **Federated registries**
 - Data qualität

Teilnehmer Punkte VII

- Medizin/Harald
 - Sensitive daten/schicksalhafte daten anonymisierung
 - Löschen von daten zwingend
 - Planung von Datennutzung bekannt zu anfang
 - Wenig standardisierung / kein Repositorium
 - Diskussion im wesentlichen national
 - Aufgaben: VRE, Identitätsmanagement, qualität, MD harvesting/mapping,
 - SFB Inf Projekte in silos
 - RDA
 - DMPs unklar
 - Bei konkreten Punkten eigene lösungen – aber systemic
 - Guidelines/principles
 - Backup-reusable archive
 - Data sharing noch schwer

Teilnehmer Punkte VIII

- Wasser/Christoph
 - Rat zu allen wasseranträgen etc
 - Wasserstände, temperatur, strömung etc
 - DM & qualität (volumina, heterogen, div typen, modellierungsformate,
 - Zusammenarbeit und vereinheitlichung komplex
 - Druck durch INSPIRE (GDI-DE´) und GovData
 - Weg (SOA, iso & OGC), interne Aufbearbeitung der Daten aufwendig
 - RDA:
 - Prinzipien werden unterstützt, eigener laden muss
 - Trifft alles zu, Bisher nicht konkret genug für wasser
 - absegnung von ministerium nötig – nicht einfach
 - Hilfe bei einstieg in RI
 - Bisher kaum WFs – zu aufwendig, semantics komplex
 - CSW statt OAI-PMH

Teilnehmer Punkte IX

- ENES/Tobias
 - Integration von randgruppen / anbieten von komponenten
 - Too much effort for DM as a
 - RDA compliant workflows
 - Prefabricated policy modules

Teilnehmer Punkte X

- Geschichte/Katrin
 - In der community zurück zu formalen methoden und daten
 - Data sharing schwierig (Monografien, sammelwerke, etc.)
 - Datenautoren nicht nur zitieren
 - Planen ein nationales geschichts repositorium
 - Dauerhaftigkeit – trust fehlt also zögern (gab fachportale)
 - Haben verschiedene Levels von aktionen bez Daten definiert
 - Gib zu viel wildwuchs – zentralisierung ist wichtig
 - Dokumentation ist wichtig zusätzlich zu metadaten

Teilnehmer Punkte XI

- EPIC/Ulrich
 - RDA
 - Heftigst: auffindbarkeit, mangelnde nachnutzbarkeit
 - Binding Task von DO/PIT

Agenda Vorschlag

		Chair: Rainer
11.30	Was als nächstes von RDA	Herman
11.45	Was bietet RDA EU /Adoptionen etc	Peter
12.00	Fragen/Diskussion	
12.30	Essen	
13.15	Zusammenfassung der Punkte	Peter
13.35	Diskussion	
14.15	Schlussbemerkung	Rainer



RDA Outlook Collaborations

Peter Wittenburg

research data sharing without barriers
rd-alliance.org

Perspectives

- EU Work Programme 16/17 (as it seems!)
 - Every project should/could have a chapter on RDA participation
 - Collaborations are intended
 - RDA will be supported by CNECT and RTD
- **RDA EU 3 is discussing a change in its setup to reserve funds for collaborations from September 2015.**
- **Senior/Junior team in RDA ready to give support from September.**

Contact Project Management Office (Herman or me)
- **D Funding – not yet clear**

Data Science is Changing

- Accepting/converging on trends

- Increasing volumes & complexity
- Increasing anonymous relationship
- Increasing sharing & re-use
- Increasingly federated landscape
- Reduce D...



Senior people:
It is time to change, our data work is too inefficient.
But need guidance and clarity for investments.
Need to accelerate solution finding.

- ...services broadly
- need to give guidance across borders
- Drivers: scientific projects requiring infrastructures
- **Is there a glue for all like for TCP/IP? – RDA a Chance**

Funders want changes

- let's use the G8 formulations – data should be

- **searchable** -> create useful metadata
- **accessible** -> deposit in trusted repository
- **interpretable**

Funders:

There is a need and will to change globally and cross-disciplinary.

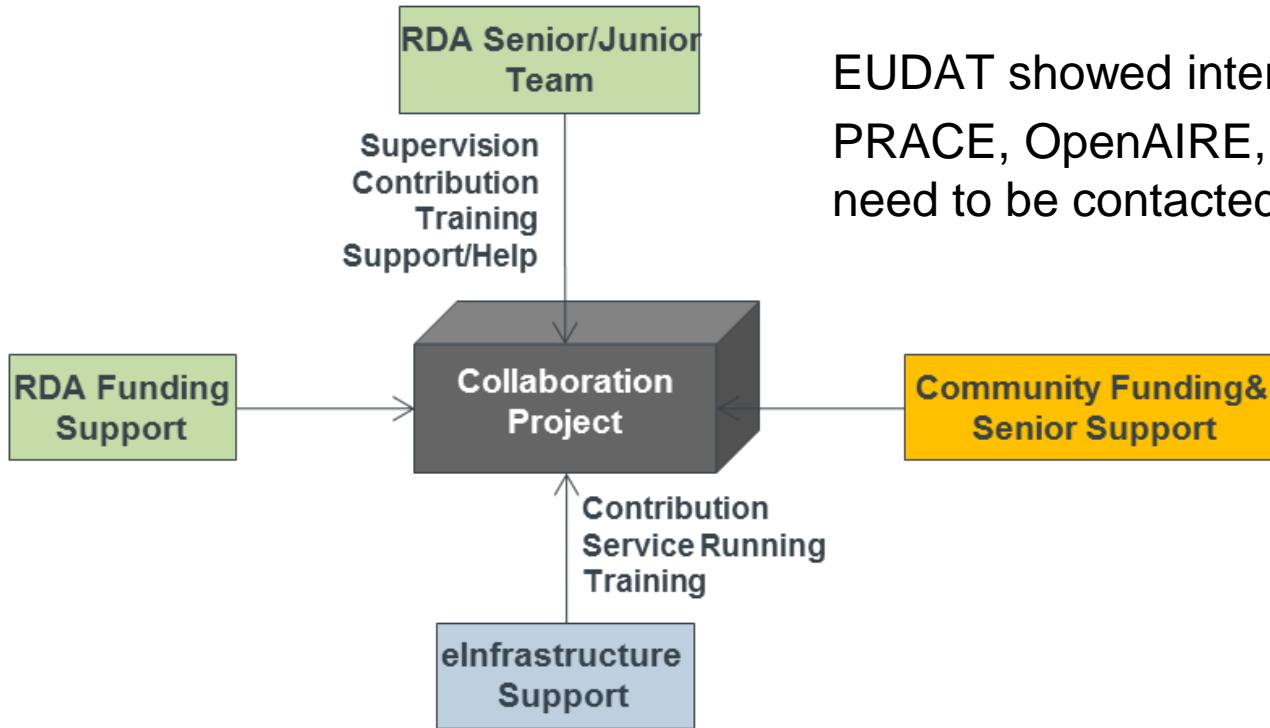
RDA is a chance – not more.

Management Plans?

What are the consequences of these principles?

- How to design the necessary infrastructure?

Plan: Joint projects



EUDAT showed interest and has a call.
PRACE, OpenAIRE, EGI, Helix Nebula etc.
need to be contacted.

- Community interest in RDA result uptake or related activities
- e-Infrastructures interested in offering funds, services and knowledge
- RDA EU will offer limited funds, support/knowledge, training etc.

RDA Collaboration Ideas

- Scope of collaborations (effort: 3 – 6 PM from RDA)
 - Co-organizing domain workshops or training courses
 - Adoption of a specific RDA output and infrastructure integration
 - Adoption of a specific software stack to be RDA compliant
 - Writing RDA Guidelines tuned to domain specifics
 - Make something from community generic (widely accepted)
 - Etc.
- Conditions
 - Visible co-funding (in-kind, person)
 - Added value towards sharing and re-use – thus RDA goals
 - Joint supervision
 - Participation in RDA events to disseminate knowledge
 - Preference for early-career involvement

RDA Collaboration Ideas

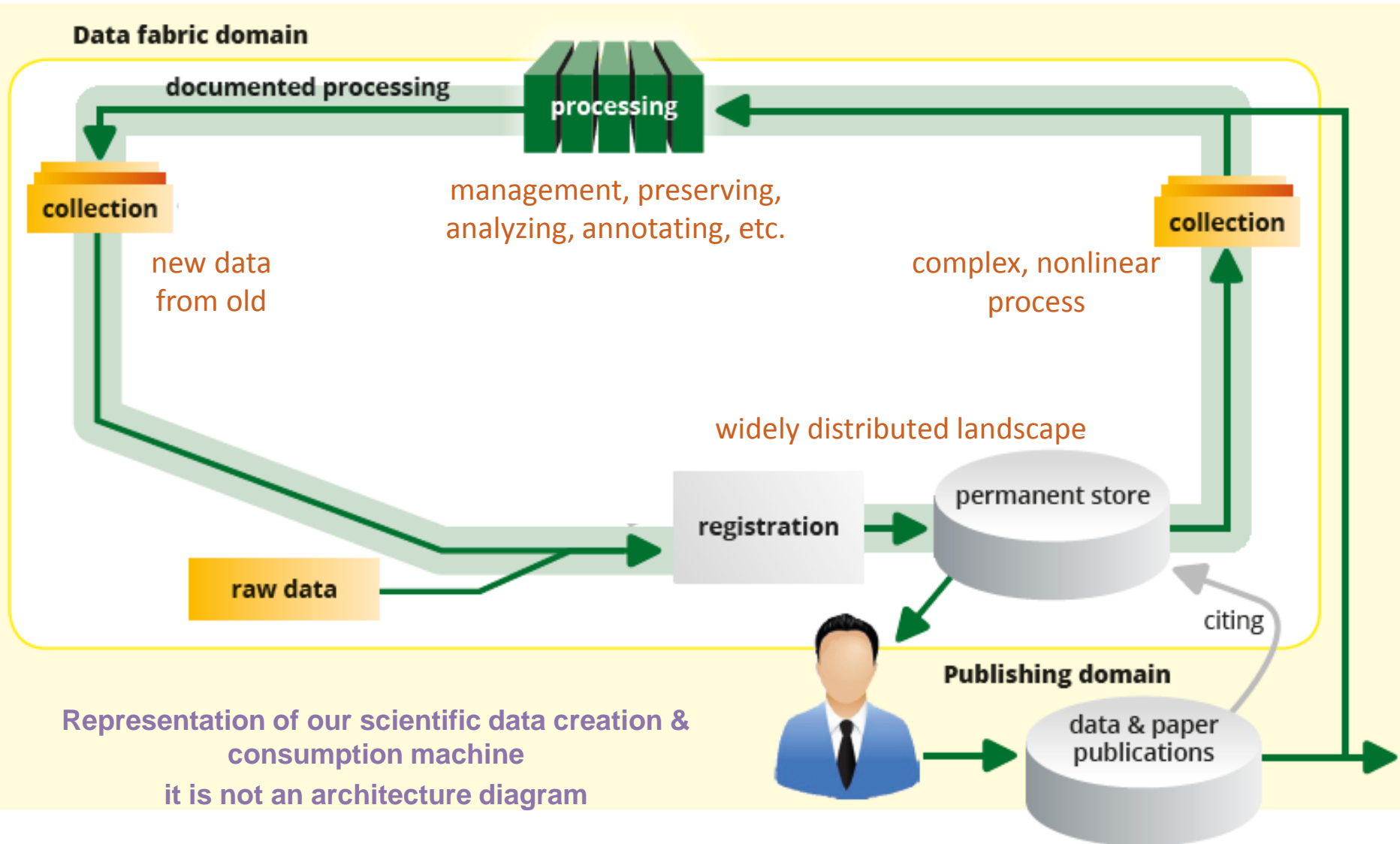
■ Process

- Open Call (synchronised with e-Infra calls)
- Advice on proposals from advisory board
- Board of Directors to finally decide
- RDA Plenaries & Data Summits as platform for presentation and evaluation

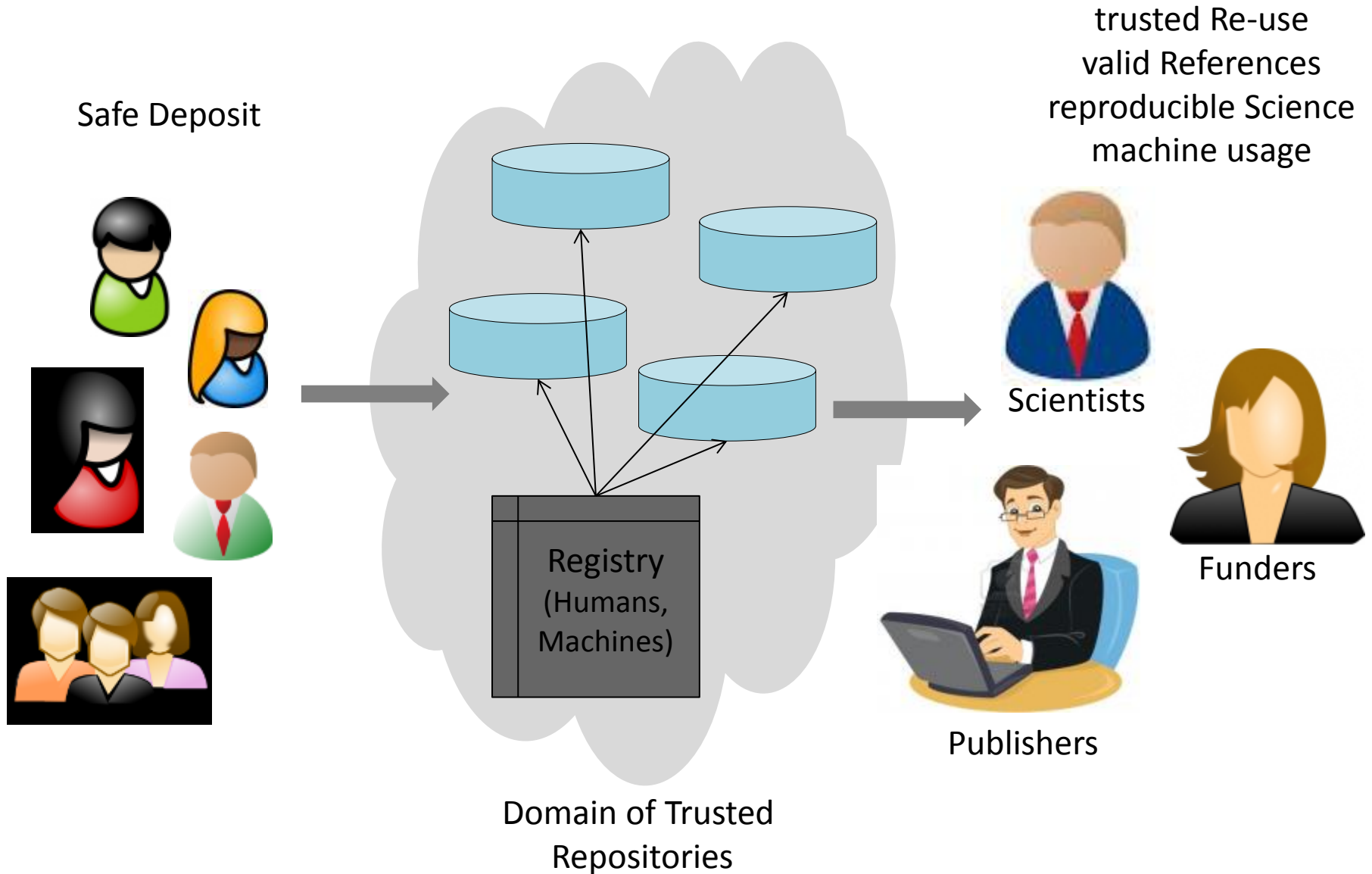
■ Timing

- Call: September 10th
- Deadline Submission: September 30th
- Evaluation: October 30th
- Start from November 15th

Abstract Data Cycle in the Labs



Example: Repository Registry (Spec&test)



Thanks for your attention.



<http://www.rd-alliance.org>

<http://europe.rd-alliance.org>

Actions – you do 😊

- Add Use Cases to Data Fabric Wiki to allow all of us to do abstractions towards common components and interfaces

Data Fabric: <https://rd-alliance.org/group/data-fabric-ig.html>

Data Fabric Wiki: <https://rd-alliance.org/node/44520/all-wiki-index-by-group>

- Comment on Paris Paper

Data Management Trends, Principles and Components - What Needs to be Done Next? V6.1:

<http://hdl.handle.net/11304/992fe6a0-fe34-11e4-8a18-f31aa6f4d448>

<https://rd-alliance.org/node/44520/all-wiki-index-by-group>

Use Cases are the basis!

	Name	Institute	state
1	Language Archive	Max Planck Institute NL	in operation
2	Geodata Sharing Platform	Academy of China	In operation
3	Datanet Federation Consortium	RENCI US	In operation
4	ADCIRC Storm Forecasting	RENCI US	In operation
5	EPOS Plate Observation	INGV/CINECA Italy	In operation
6	ENVRI Environment Observation	ETH Zurich Switzerland	In operation
7	In design
8	...	ETH Zurich Switzerland	in testing
9	Climate modeling	DKRZ Germany	In operation
10	LIGO Gravitation Physics	NCSA US	In operation
11	ECRIN Medical Trial Interoperation	U Düsseldorf Germany	In testing
12	VPH Physiology Simulation	U London UK	In operation
13	Species Archive	Nature Museum Germany	In operation
14	International NeuroI Facility	INCF Sweden	In operation
15	Molecular Genetics	MPI Germany	In operation

Participate by submitting Use Cases
via Data Fabric Wiki

Position Paper “Paris.doc”

- Recently paper a number of colleagues engaged in RDA Data Management Trends, Principles and Components – What Needs to be Done?
- Co-authors don't claim to own any ideas – but kick-off a broad discussion
- Need to accelerate solution finding and convergence process

8 Common Trends	Partly stable, some still in debate
G8+ Principles	Widely agreed
Consequences of Principles	Not really thought through
19 Components	To be discussed now
Organizational Approaches	To be discussed now

Doc: <http://hdl.handle.net/11304/992fe6a0-fe34-11e4-8a18-f31aa6f4d448>

Data Fabric Wiki: <https://rd-alliance.org/node/44520/all-wiki-index-by-group>

Components - Position Paper

1. PID System
2. Actor ID System
3. Registry S for Trusted Repositories
 - OAI-PMH, ResourceSync, SRU/CQL
 - Workflow Engine & Environment
 - Conversion Tools
4. Metadata S
5. **Much already out there**
6. **Why is it not used broadly?**
7. **Why is there a hesitation to broadly invest?**
8. **Need a systemic approach!**
9. **Need to start from Use Cases!**
10. Training Modules
11. Authorisation Record Registry S

Actions I – RDA kickoff

- General points: short-term - long-term / Report in E?
- Meetings/Aktivitäten
 - RDA Global: Plenaries + WG/IG Sitzungen
 - RDA EU: workshops/Tutorials/Webinars/etc.
 - RDA D: November Meeting Potsdam (2 days)
 - Plenar (BMBF, DFG, Keynotes)
 - Breakouts/Tutorials/Seminare – mit Experten
 - Summaries
 - PC: Rainer, Jens, Ralph, Peter, (Hans?)

Actions I – RDA kickoff

- Topics für Workshops/seminars/sessions
- Low level / nontech
 - Offered services (EUDAT, etc.)
 - Hilfe bei Einstieg in RI – kleine Abteilungen
 - Licensing (where to move to, CC for data?, etc.)
 - Productive DMP (Support, Guidelines)
 - Aufbau eines Repositories (Daten-org, Software)
 - Serie für Bibliothekare (via LIBER?)
- advanced
 - DTR Usage
 - CS-Practitioner Austausch (concret: PID Info Types)
 - Advanced PID usage with concrete cases/configurations
 - Advanced X
 - MD Komponenten!!!! – was zu PIDs
 - WFs / Prefabricated WF module/automatische Orchestrierung/prov
 - Virtual Collection Builder

Actions I – RDA kickoff

- Documents
 - Usage stories / „ architectures“ / components cooperating
 - Primers on WGs
 - „Manifesto“ zu Fundern (synched mit EU?)
 - Prinzipien/Commons (synched mit EU)

Actions I – RDA kickoff

- New WG/IGs
 - Repository Registry (different flavors)
 - Distributed Content Search / Data Mining
 - Virtual Collection Building
 - State Core Definition
- Others
 - Fertige code snippets
 - Kollaboration mit W3C etc.
 -