The FAIR Principles - what these are and some examples

Parthenos Heraklion– 17/05//2017

Hella Hollander WP3 Leader

KNAW-DANS

FAIR Principles

1. Introduction work in WP3: Common Policies (5 min)
2. FAIR Principles (5 min)
3. Used FAIR to structure, connect and present (5 min)
4. Questions and Answers (5 min)

1. Introduction work in WP3: Common Policies

Hella Hollander

KNAW-DANS

Common Policies

Goal is to:

Agree on and define the concepts of Policy, Guidelines, Best practice, their objectives and target audience

Produce a coherent, authoritative, well accepted set of policies/guidelines/tools concerning the management of data lifecycle and related issues such as IPR and quality.

Parthenos Flagship Expected Results:

Guidelines on data management

Stakeholders

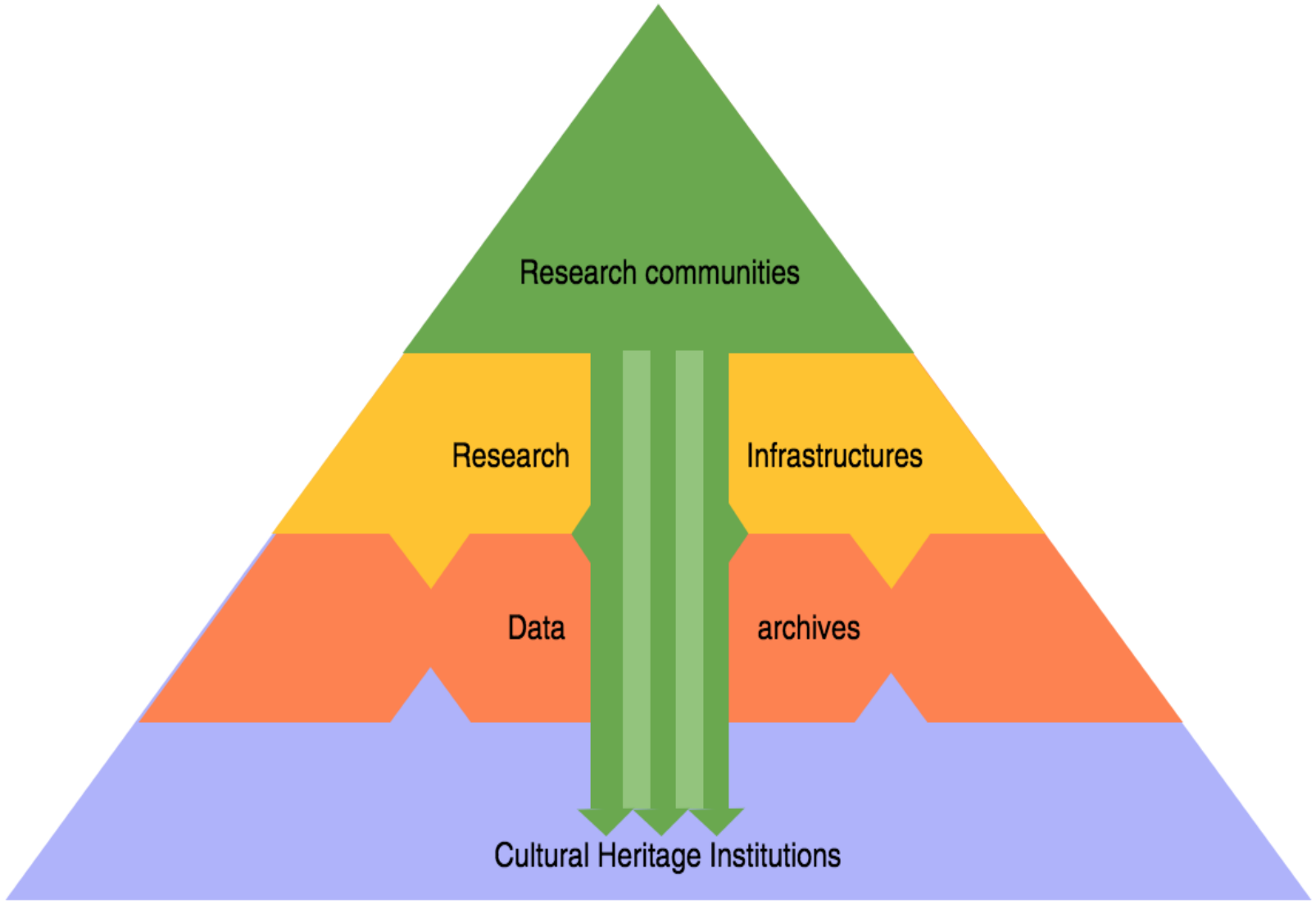
1) **History** in a broad sense: including Medieval Studies, Recent History, Art History, Epigraphy

2) **Language-related Studies** including Literature, Linguistics, Philology, Language Technology

3) **Archaeology, Heritage & Applied Disciplines** including Cultural Heritage, Archives, Libraries, Museums, Preservation / Conservation experts, Digital curation / edition / publishing

4) **Social Sciences** in a broad sense: Sociology, Political Science, Geography, Anthropology, Cultural Studies

Stakeholders



Objectives

Scientific objective

Enable research questions to be answered

More & better!

Project objective

Make research data available

More & better!

Task objectives

Provide policy and guidelines for repository, data and metadata.

Conduct foresight studies

Common Policies: Approach

The results of the effort of our work should have a **long-term impact** on common policies and guidelines on research data management, IPR, Open Access and Open data and how to **implement** them within the Humanities.

An **inventory** of existing policies from the different infrastructures has been made. (Matrix)

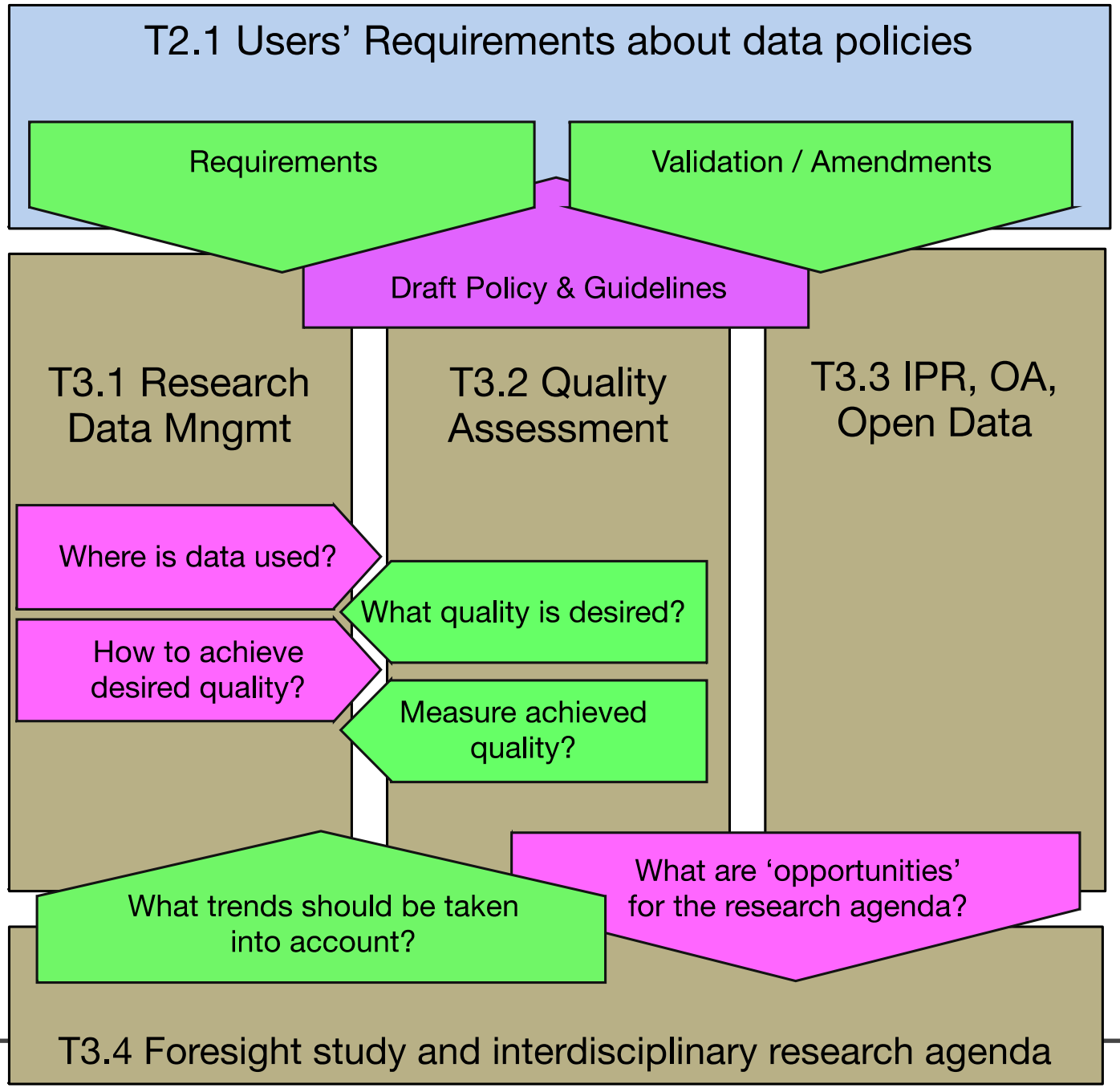
D3.1 represents the result of desk research and theoretical background giving **guidelines and case studies** to the researchers. The FAIR principles are used as a connecting backbone, making it easy to access the results.

The outcomes of this deliverable will be made more useful and reusable by creating an **interactive guide** (web tool) to present the results: The wizard.

There is a need to **define and test** the requirements for shared policies

Common Vision

- Help researchers to make their data of better quality, interoperable, sharable, findable and reusable (FAIR principles)
- Agree on and define what policies, guidelines and best practice are.
- Overview of existing policies in the Parthenos disciplines, for different data lifecycle phases
- Find the commonalities between disciplines in the humanities in terms of policies, RDM and IPR, open access
- Find the gaps: what disciplines are advanced in terms of policies and what are not
- Give recommendation and guidance to researchers
- Give recommendation and guidance to data archives
- Give recommendations and guidance to cultural heritage institutions
- Give recommendations and guidance to research infra structures


T2.1 Users' Requirements about data policies
Requirements
Validation / Amendments
Draft Policy & Guidelines
T3.1 Research Data Mngmt
T3.2 Quality Assessment
T3.3 IPR, OA, Open Data
Where is data used?
What quality is desired?
How to achieve desired quality?
Measure achieved quality?
What trends should be taken into account?
What are 'opportunities' for the research agenda?
T3.4 Foresight study and interdisciplinary research agenda

2. FAIR Principles

Hella Hollander

KNAW-DANS

Open and FAIR Data in Trusted Data Repositories

Data does not only need to be Open

Data must also be FAIR

Findable, Accessible, Interoperable, Reusable

And must remains so, and therefore should be preserved in a DSA Certified Trusted Digital Repository

Perfect Couple

FAIR principles for **data quality**

DSA criteria for **quality of TDR**

minimal set of community agreed guiding principles to make data more easily **findable**, **accessible**, appropriately **integrated** and **re-usable**, and adequately citable.

- A perfect couple for quality assessment of research data and trustworthy data repositories
- Ideally: a DSA certified archive will contain FAIR data

FAIR Data Principles

In the FAIR Data approach, data should be:

Findable – Easy to find by both humans and computer systems and based on mandatory description of the metadata that allow the discovery of interesting datasets;

Accessible – Stored for long term such that they can be easily accessed and/or downloaded with well-defined license and access conditions (Open Access *when possible*), whether at the level of metadata, or at the level of the actual data content;

Interoperable – Ready to be combined with other datasets by humans as well as computer systems;

Reusable – Ready to be used for future research and to be processed further using computational methods.

Implementing FAIR Principles

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

15 Criteria

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

Everybody loves FAIR!

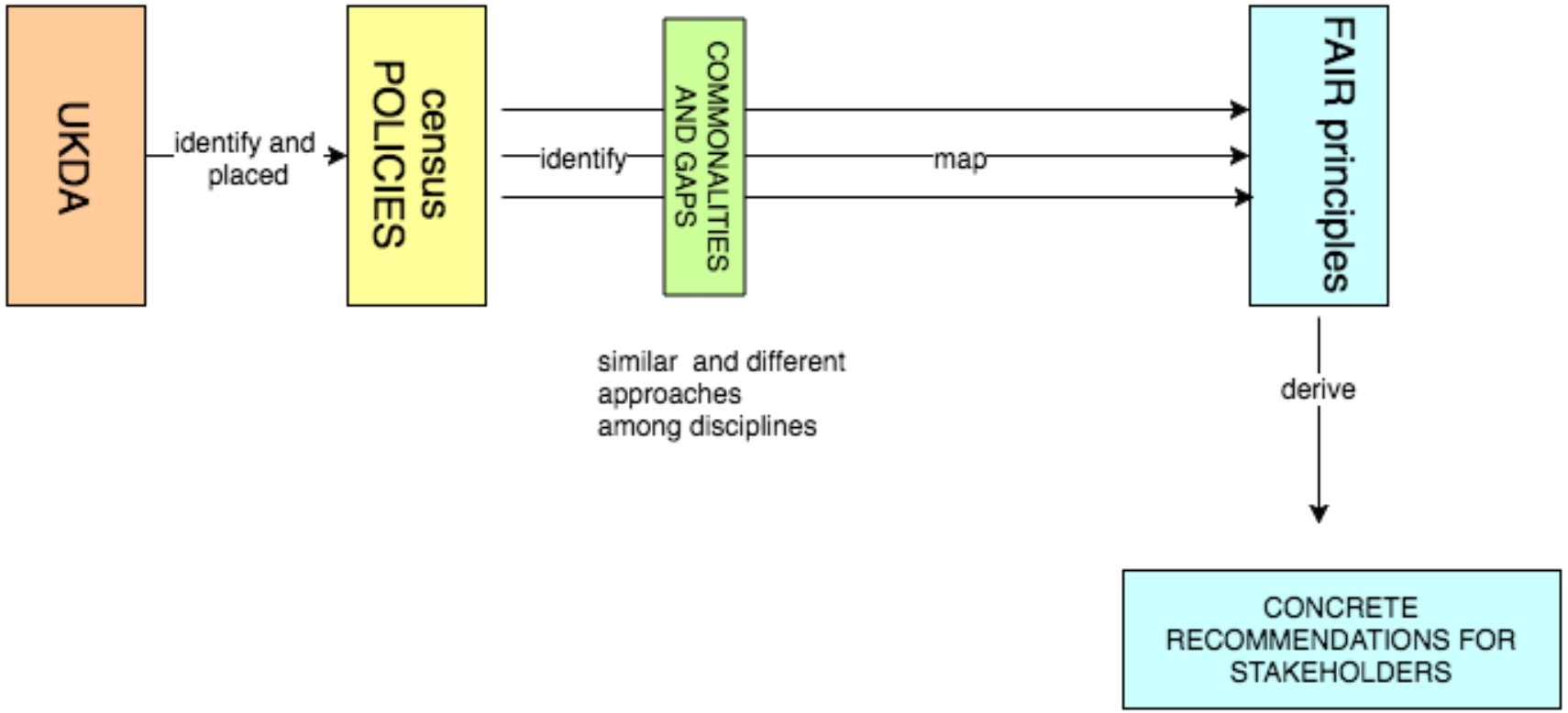
Everybody wants to be FAIR… But what does that mean?
How to put the principles into practice?

3. Used FAIR to structure, connect and present

Hella Hollander

KNAW-DANS

From census of policies to recommendations



what I suggest here is that in the end the fair principles are transformed from assessment principles to step-by-step guidelines to be used by researchers in the humanities and other stakeholders (CHI, data archives etc...)

PARTHENOS high-level recommendations

Overview of existing policies and recommendations concerning the quality of (meta)data and repositories, IPR, Open data and Open access. They are revisited and mapped onto the FAIR Principles.

The result: a set of high-level PARTHENOS recommendations

FAIR structuring principles

PARTHENOS High level principle: **Accessible**

- FAIR: Defined by long term storage and access. Well defined licence and access conditions on level of metadata and data content
- DSA criteria: Data are accessible
- Matrix: Data re-use
- Policies: Conditions of use eg of DANS
- Use case: Mary wants to use data and she finds licence policies telling her about the conditions of use.
- PARTHENOS data model: SSK toolkit (standards: DC) and training modules on the website about data and metadata

PARTHENOS high-level recommendations: Accessible

Examples:

- (Meta)data should be open as possible and closed as necessary
- Protected data and personal data must be available through a controlled and documented procedure. Information that needs to be protected, for example for privacy reasons, should not be part of the publicly accessible (meta)data but should be recorded as part of the documentation of the resource in restricted contexts.
- In order to be fully accessible, research data should be fully accessible via (free) exchange protocols.
- Maintain the integrity and quality of data. This is a general principle, that emerged in particular from the interviews with historians. It refers to the necessity to maintain the richness and the context of the data created and collected during time

FAIR Principle Accessible in Data Management Plan

2.2 Making data openly accessible	Specify which data will be made openly available? If some data is kept closed provide rationale for doing so	Specify if there are any restrictions on public accessibility and describe the exceptions to public and free access
	Specify how the data will be made available	☐ Deposition in a repository, please specify which ☐ Other, please specify ☐ I don't know yet
	Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?	☐ Component Metadata Infrastructure (ISO 24622-1) to create an environment that supports different metadata schema ☐ MAG and METS-MDI schemas ☐ Dublin Core ☐ VRA ☐ NISO ☐ MD5 ☐ METS ☐ ACDM ☐ CIDOC CRM ☐ (Qualified) Dublin Core metadata

Questions & Answers

Hella Hollander
DANS-KNAW

Hella.Hollander@dans.knaw.nl

PARTHENOS is a Horizon 2020 project funded by the European Commission. The views and opinions expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.