Developing EU-BON's site-specific portal

Yoni Gavish

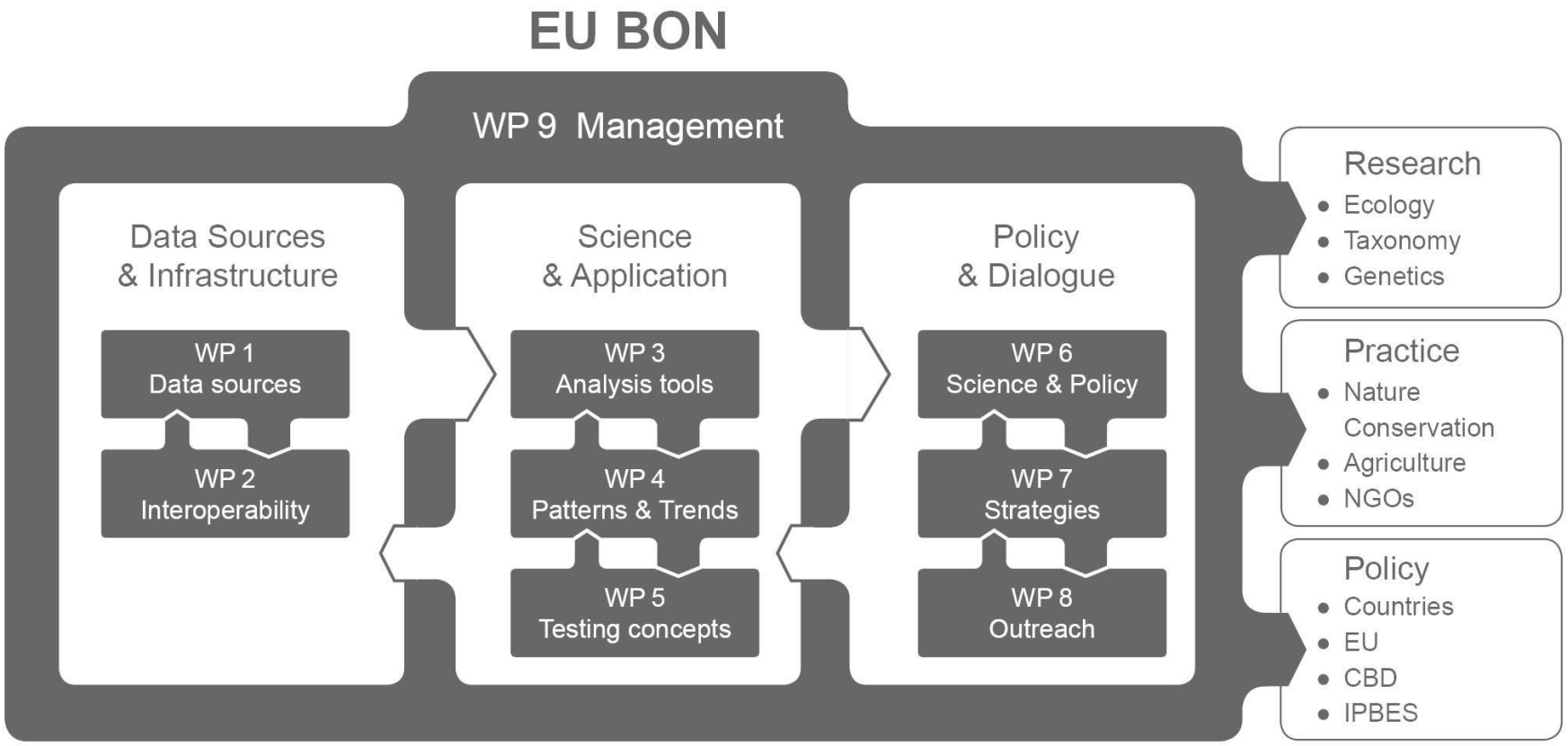
University of Leeds

gavishyoni@gmail.com

Before we start...

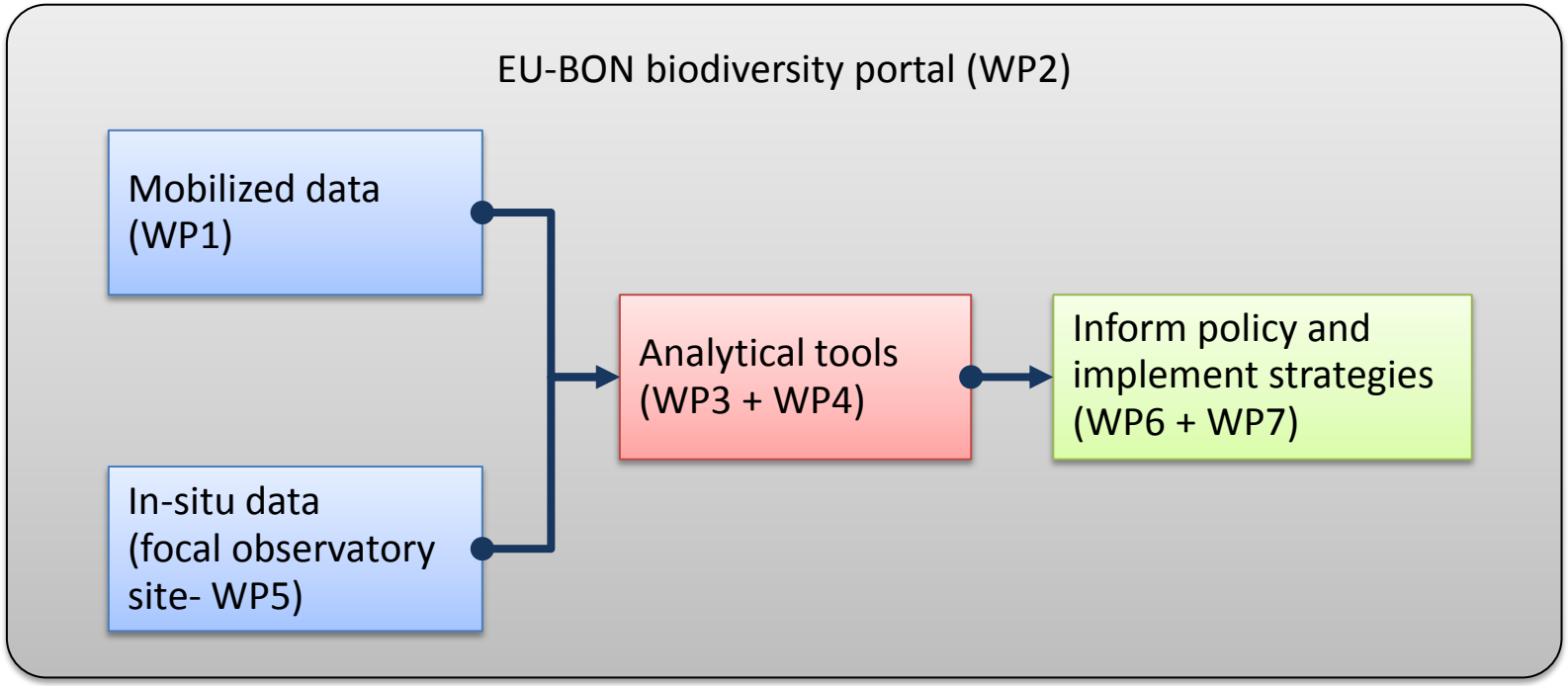
- I have no expertise in data management + web designing
  - Some (most?, all?) of the ideas are not feasible
  - However, they represent a vision for a useful portal
- I represent mainly the `researchers` user group
  - Other user groups – `site managers`, `policy makers`
  - However, it is a start…

Review + AB: more cross-axes links



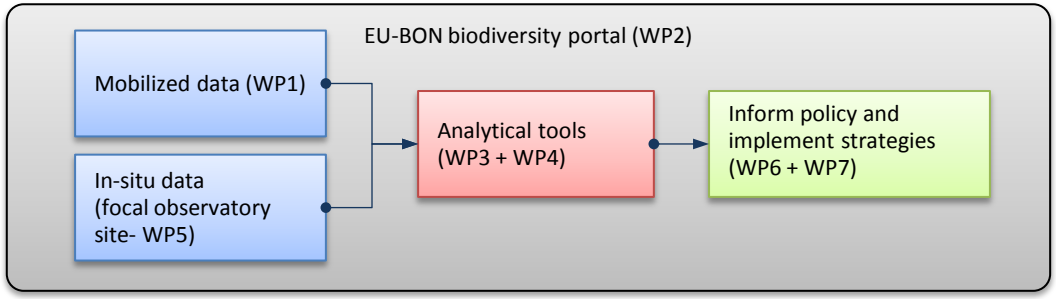
SC: Site-specific portal: original aim

- Show-case for the flow of information between the three axes:



SC: Site-specific portal: original aim

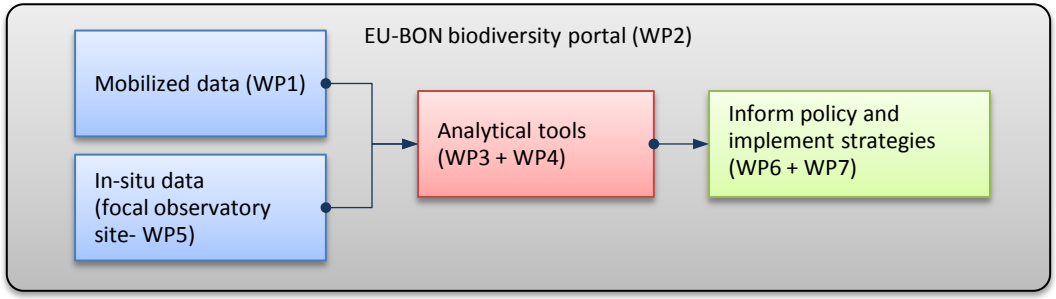
- Show-case for the flow of information between the three axes:



- Stack various types of basic / derived data layers:
  - Remotely-sensed data
  - Environmental data
  - Climatic data
  - Biotic data
- Highlight the links between the different data-layers
- Highlight the analytical tools used to create derived layers from basic layers
- If possible, relate derived layers to policy (EBVs?)

SC: Site-specific portal: original aim

- Show-case for the flow of information between the three axes:



- Stack various types of basic / derived data layers:
  - Remotely-sensed data
  - Environmental data
  - Climatic data
  - Biotic data

Not a limiting factor…

- Highlight the links between the different data-layers
- Highlight the analytical tools used to create derived layers from basic layers
- If possible, relate derived layers to policy (EBVs?)

Sierra-Nevada site

- Around 1000 data layers , not counting biotic dataset

Source	Category	Examples	# of layers
Basic	Environmental	DEM, geology, landscape units, hydrology, etc.	15
Basic	Environmental	Vegetation maps	15
Basic	Remote-sensing	NDVI and EVI from MODIS (2000-present)	stored in a database
Derived	Remote-sensing	Trends for NDVI and EVI from MODIS (2000-present)	stored in a database
Basic	Remote-sensing	NDSI (snow cover) from MODIS (2000-present)	stored in a database
Derived	Remote-sensing	Trends NDSI from MODIS (2000-present)	stored in a database
Basic	Remote-sensing	Land surface temperature from MODIS (2000-present)	stored in a database
Basic	Environmental	Spatial distribution of threatened species	100
Derived	Environmental	SDM for threatened species. Future climate projections.	100
Derived	Environmental	Yearly climate maps (1960-2010)	150
Derived	Environmental	Predicted Yearly climate maps (2010-2100)	600

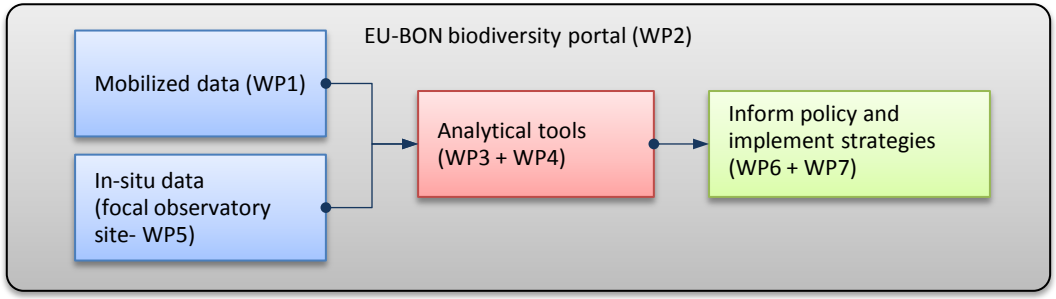
Rhine-Main Observatory

- Around 500 data layers used / produced within EU-BON

Source	Category	Examples	# of layers
basic	environmental	DEM, Geology, …	± 10
basic	climatic	Bio-climatic layers	± 20
basic	Remote-sensing	LandSat	± 35
basic	Remote-sensing	Spot 5	± 50
basic	Remote-sensing	KOMPSAT-2	± 10
basic	biotic	occasional observation (GBIF)	?
basic	biotic	monitoring programs, previous research (SGN)	?
derived	environmental	temperature time series (FEM)	± 10
derived	environmental	habitat map (UnivLeeds)	± 5
derived	biotic	ensemble species distribution models (SGN)	± 150
derived	biotic	alpha adjusted species distribution models (UnivLeeds)	± 150
derived	biotic	Community – α + β diversity models (UnivLeeds)	± 5
derived	biotic	Community – pairwise similarity (MfN)	± 5

SC: Site-specific portal: original aim

- Show-case for the flow of information between the three axes:



- Stack various types of basic / derived data layers:
  - Remotely-sensed data
  - Environmental data
  - Climatic data
  - Biotic data

Not a limiting factor…

- Highlight the links between the different data-layers
- Highlight the analytical tools used to create derived layers from basic layers
- If possible, relate derived layers to policy (EBVs?)

Two main user-interface windows

Data visualization window:

- The main interface in which all available layers are visualised, and all available and potential layers are listed

Data exploration window:

- An interface that allows the user to explore linkage between layers, their meta-data files and the analytical tools that were used to create them

Data visualization window

GIS-based section

- The actual visualization of the stacked layers
- User can control order, symbols, transparency, colours, etc.

Layer list section

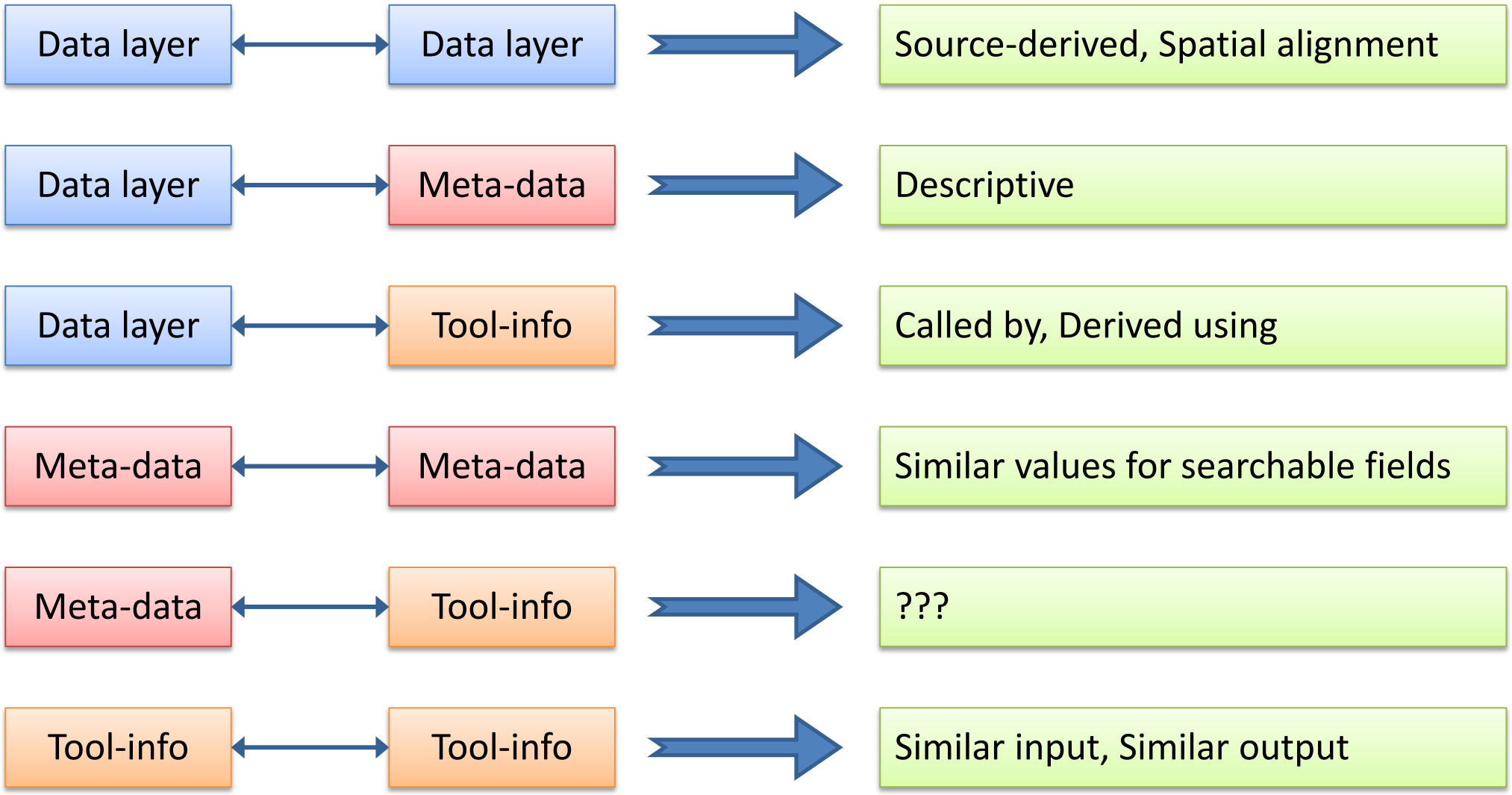
- A pivot-table style list of layers
- Variables may represent:
  - Source – basic or derived layers
  - Category – Environmental, climatic, remotely-sensed, biotic, SDMs,…
  - Availability – open access, password protected, require permission, cost money…
  - …

Data exploration window

- Dynamic network of nodes and links (Visual dynamic mind-maps)
- 3 types of nodes:

Data-layer	Meta-data	Tool-info
• Remotely sensed image • DEM • Climatic layer • NDVI, NDWI, etc. • Habitat / LC / LU class • Point observation • Monitoring plots • SDM layer • Alpha-diversity layer • Beta-diversity layer • Etc.	A meta-data file of a data-layer, with searchable fields	A file containing: • Information on an analytical tool • Link to a manuscript describing it • Link to an executable R code or R package • Etc.

Data exploration window – types of links



Data exploration window – links

- User-defined links
  - e.g., source-derived link between 2 data layers
- Machine-defined links
  - e.g., same value for a field in the meta-data file
- In general, should be quite flexible
- A ‘new’ type of meta-data file..

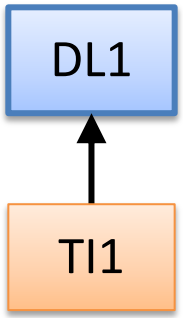
Data exploration window

- Start with data layer DL1

DL1

Data exploration window

- Expend the tool in which data layer DL1 was derived

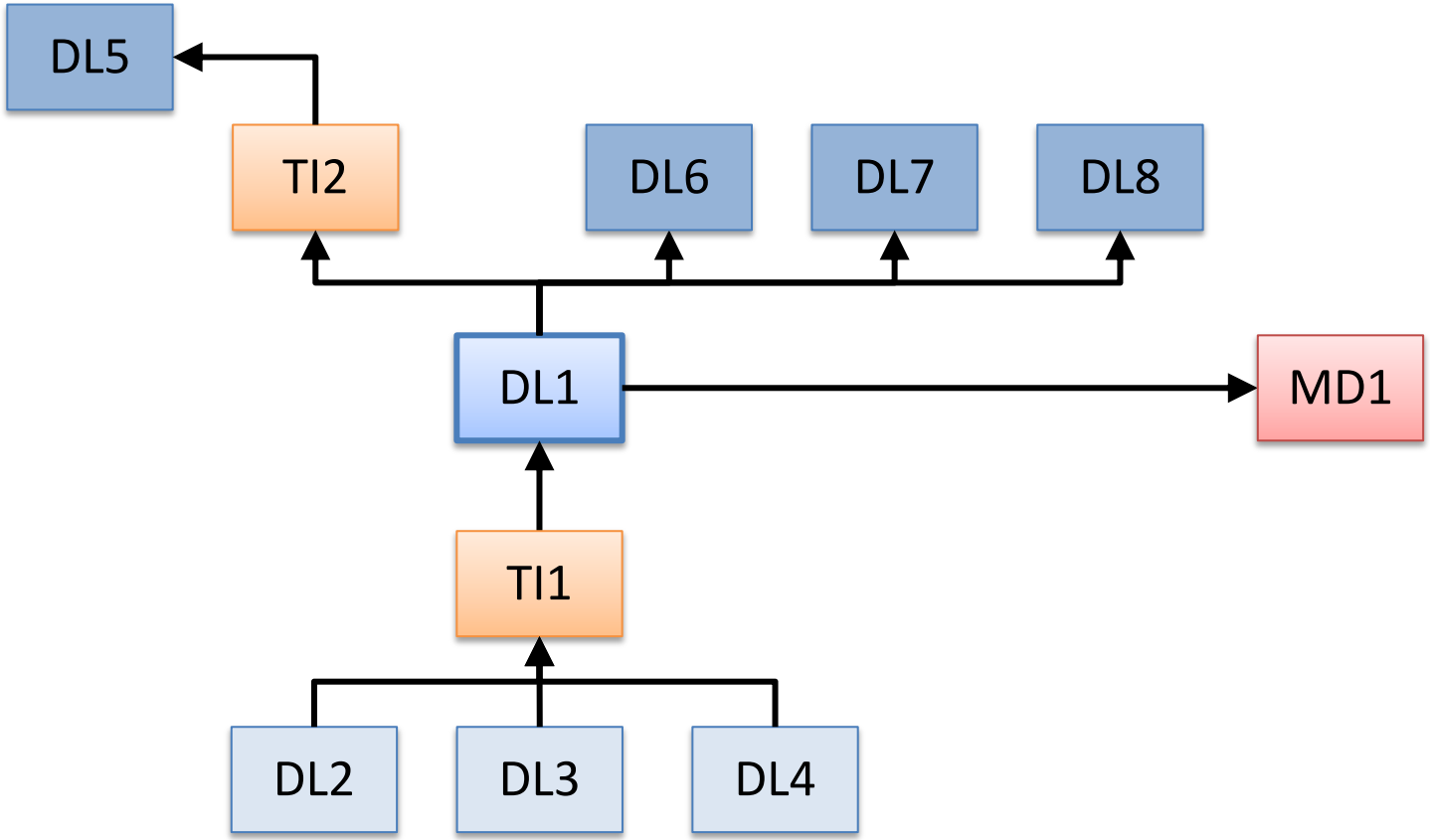


Data exploration window

- Expend the meta-data file of DL1

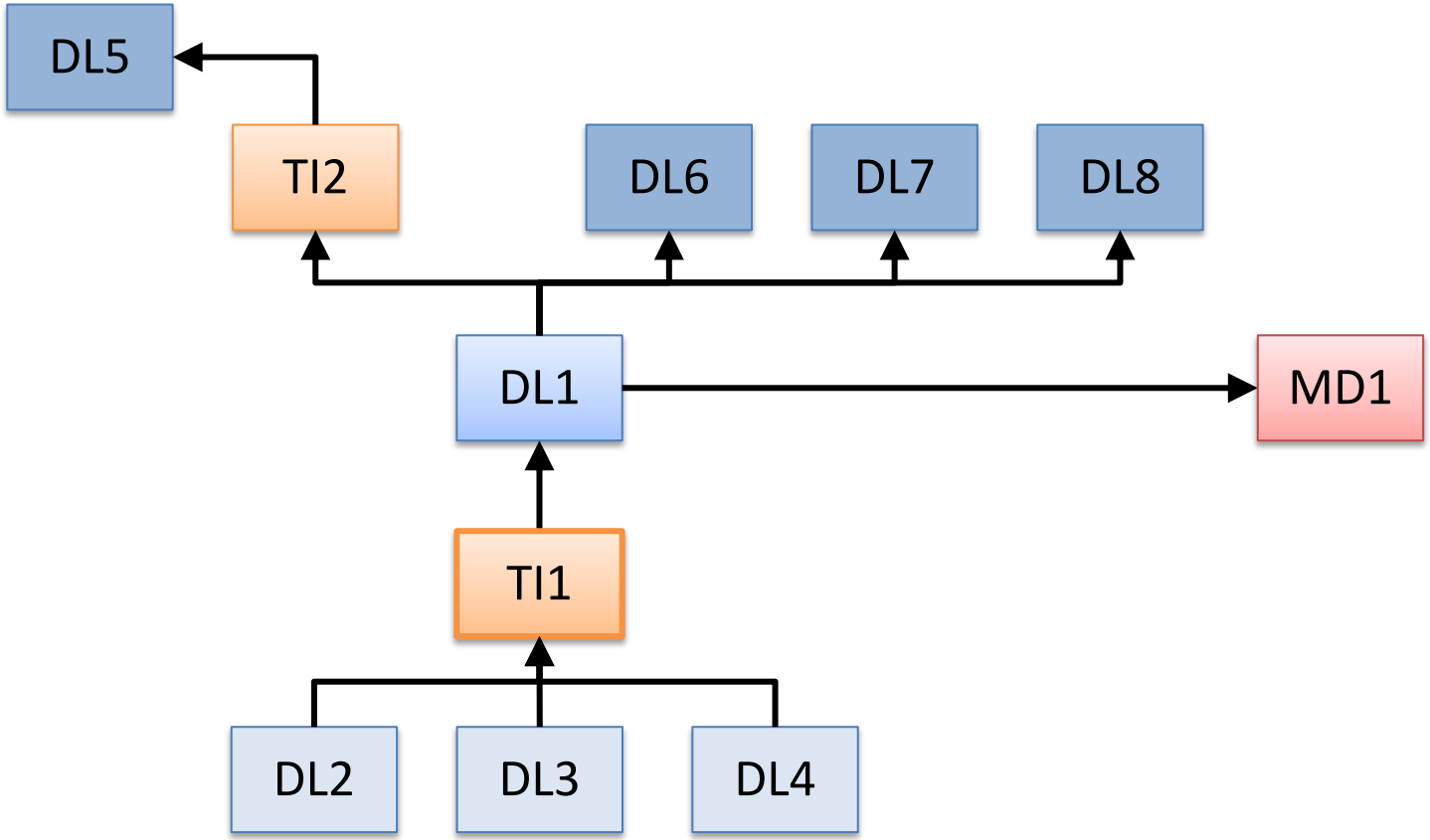
Data exploration window

- Expend all the derived layers of DL1 (with or without known tool nodes)



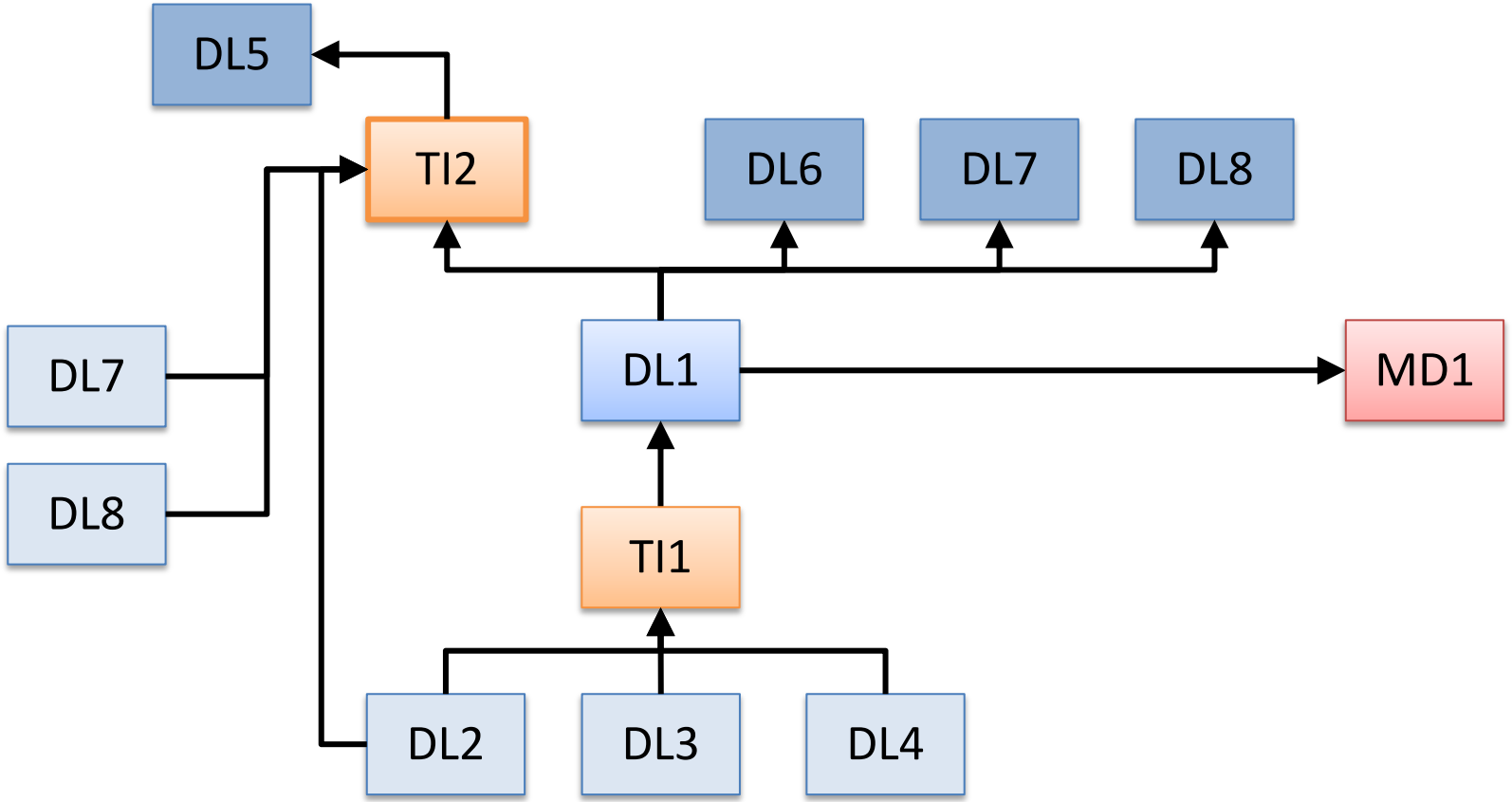
Data exploration window

- Expend all the source layers of TI1



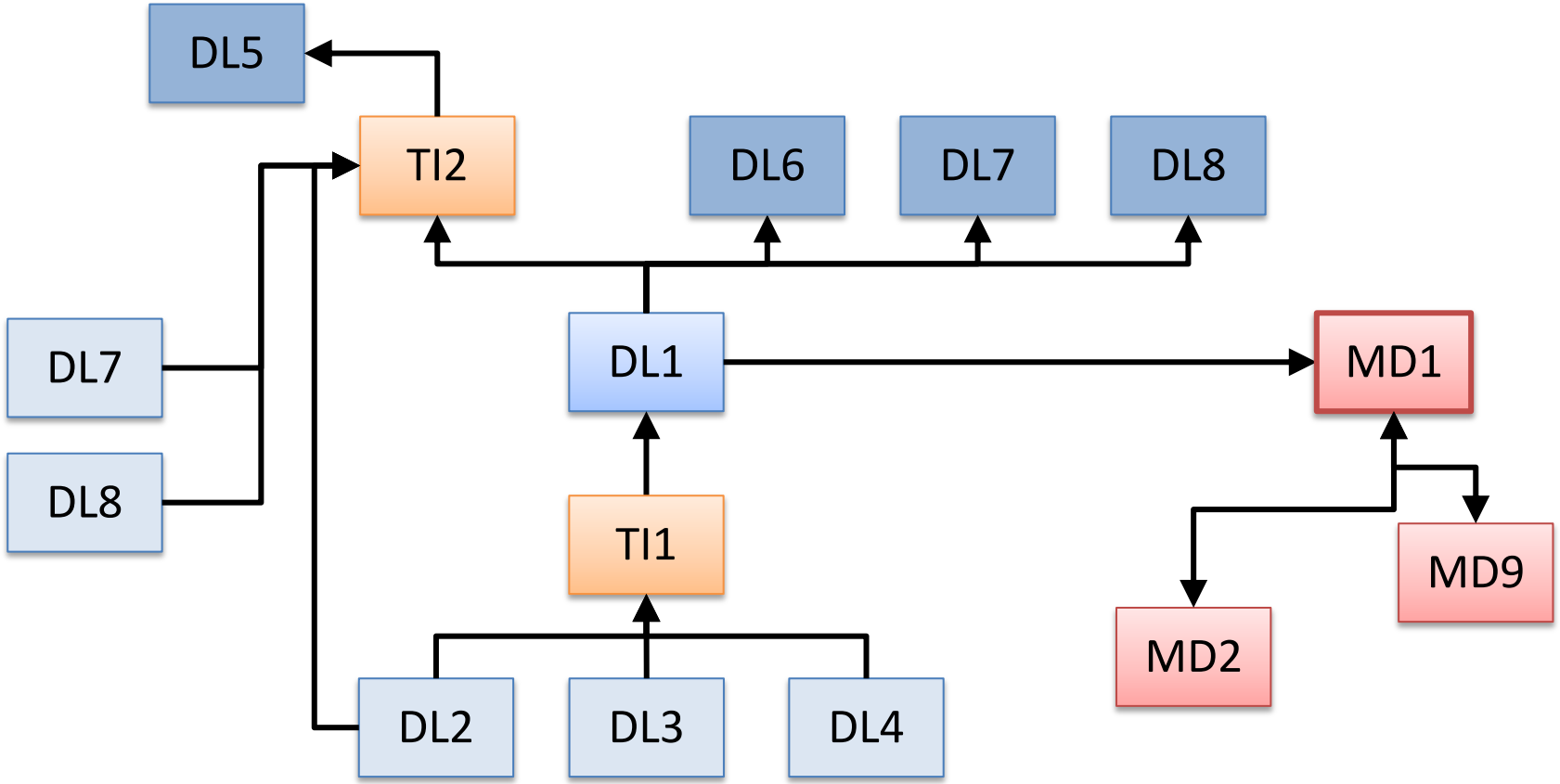
Data exploration window

- Expend all the source layers of TI2



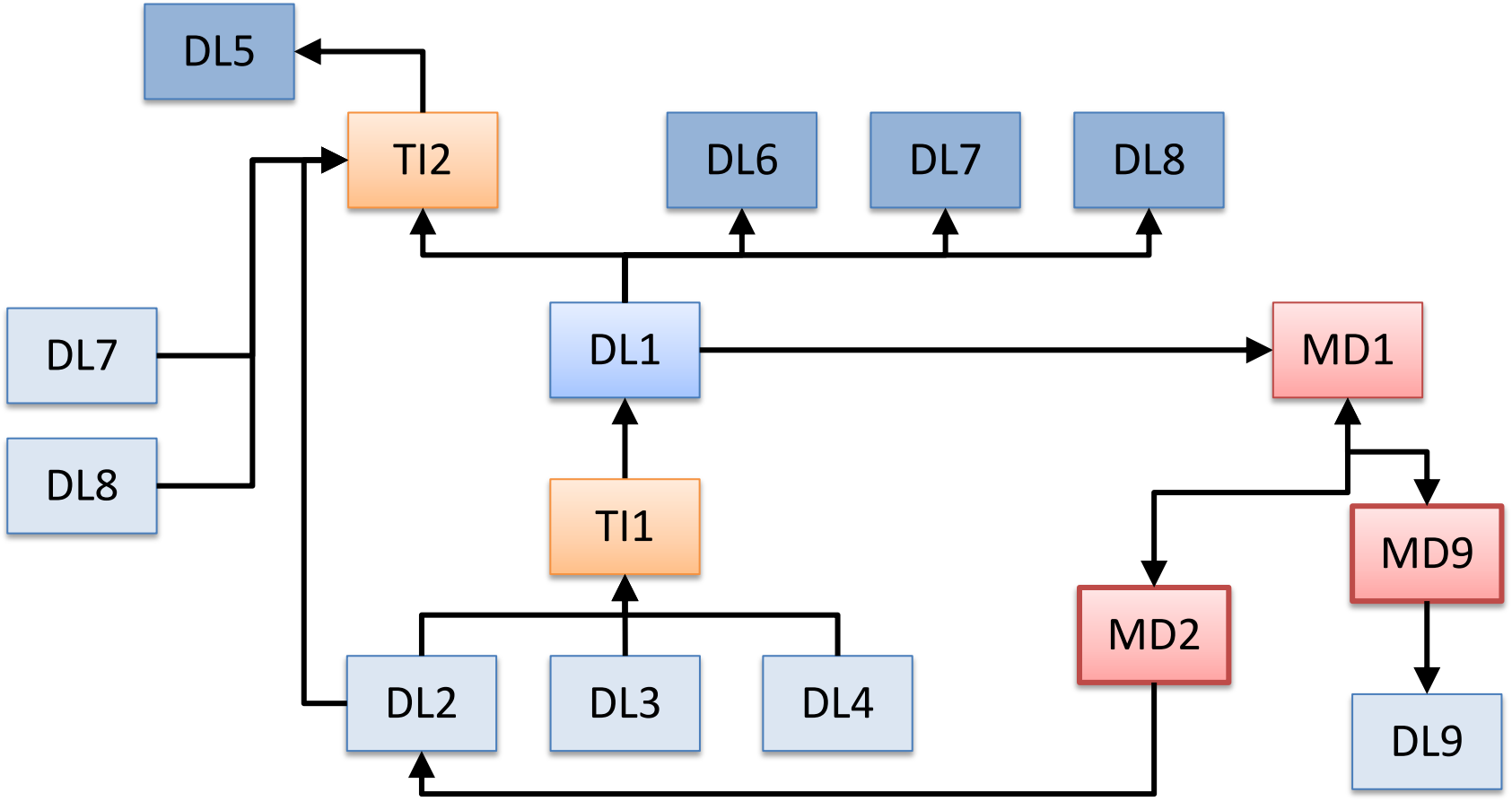
Data exploration window

- Expend all meta-data files with similar value in a field as MD1



Data exploration window

- Expend data-layers of MD9 + MD2

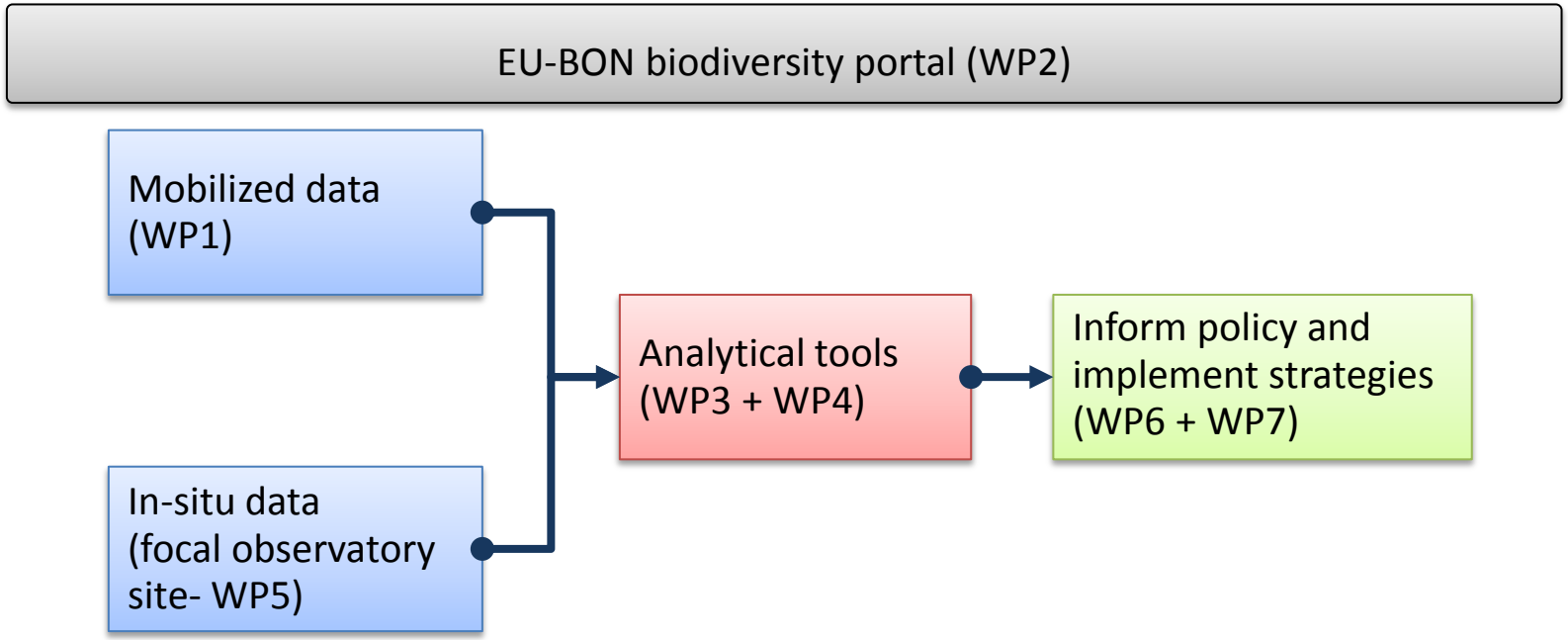


Data exploration window – advantages

- The links keep track and archive information not covered by meta-data files
- Visualization of analytical framework and data usage
- Simple, intuitive sharing of complex analyses with non-professionals
- If visually attractive, the ‘cool’ factor…

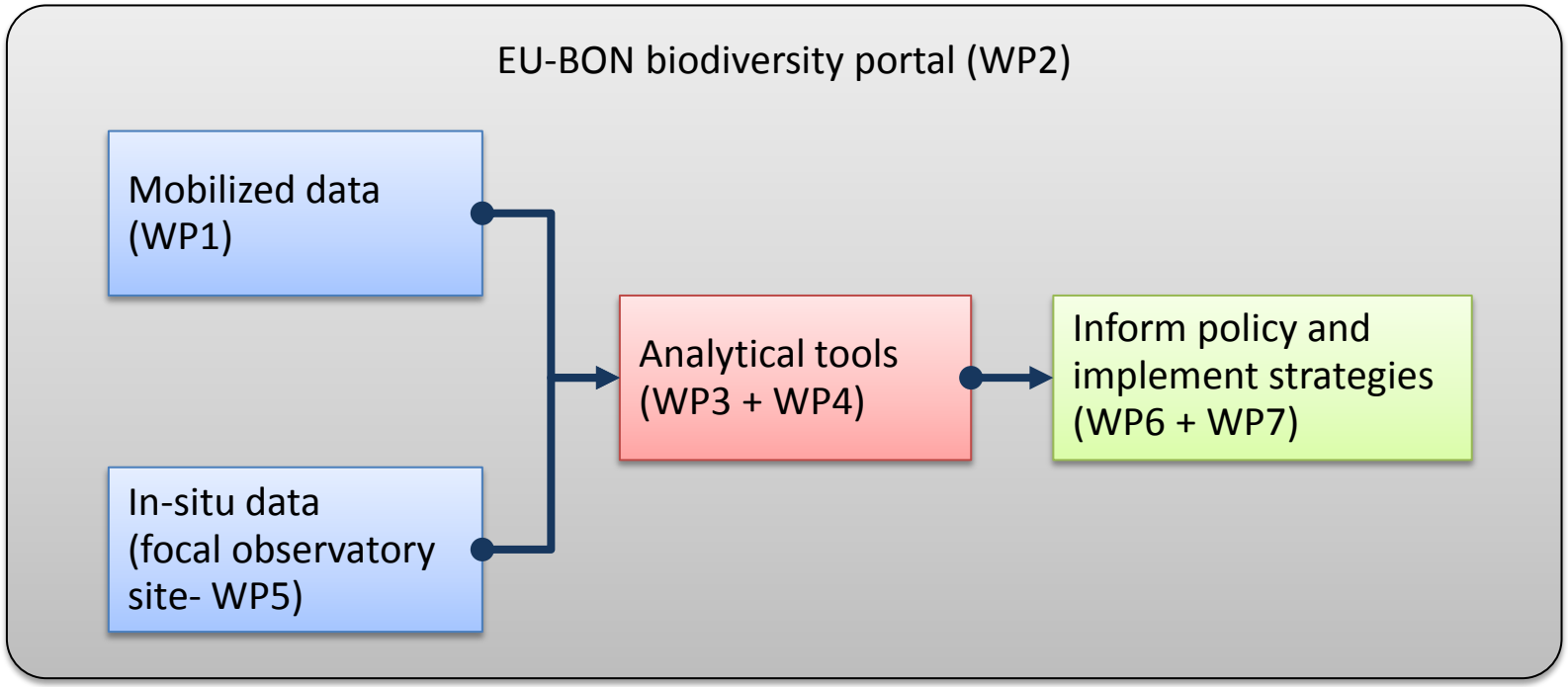
Data exploration window – disadvantages

- Creating links may be time consuming for the users
- May require considerable programming time
- Disconnected from the general EU-BON biodiversity portal



Reconnecting the site-specific and general portals

- Manage, explore and visualize biodiversity layers of different types

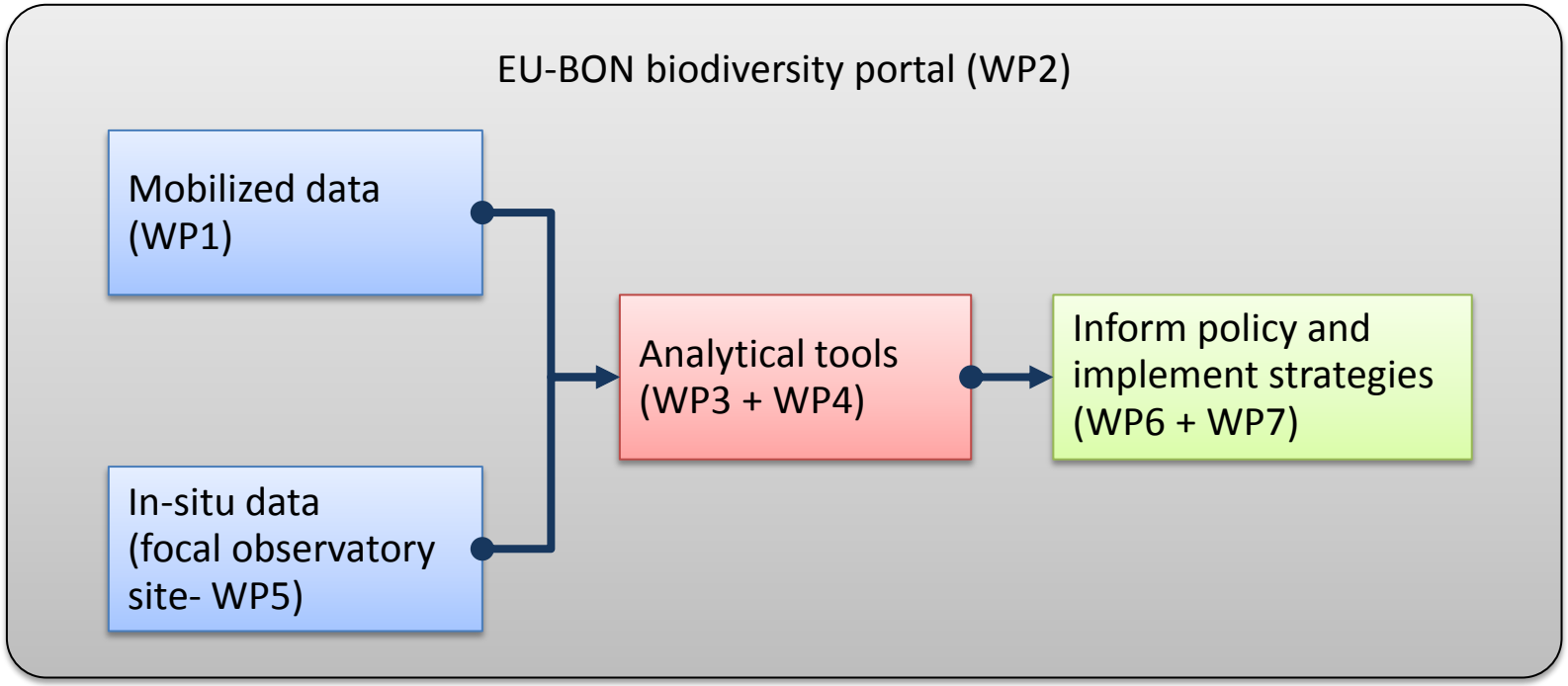


Site-specific portal – Main Aims

- Manage, explore and visualize biodiversity layers of different types
- Facilitate the acquisition of additional source layers for the focal site
- Identify relevant analytical tools.
- Bi-directional feedback with the general EU-BON portal
- Dissemination of available information
- Link to policy –
  - e.g., explore the link between EBVs derived from remote sensing (global scale) and biodiversity (on local scale)

Reconnecting the site-specific and general portals

- Show-case for the flow of information between the three axes



- Example of a user's project in a data-manager tool

What do reserachers need from a biodiversity portal?

- Most analyses → integrate data + tools from various sources
  - **Data:** Remotely sensed + earth observation data + biotic samples / plot …
  - **Tools:** R packages, GIS softwares, specific softwares …
- Data harvesting → crucial + time consuming
  - difficult to stay on top of everything
- Keeping track of data → challenging
  - working on more than one project in more than 1 platform…

Researchers need a 'data manager' tool:

Manage potential, available , processed and derived data-layer from various types

Data-manager tool – Main features

The general portal: A data 'superstore' –

- Cover various types of data
- Contains information on how each data-layer can be accessed, e.g., link to:
  - Download page, form filling window, contact details of data holders

Portal-project interactions: Dynamic + Bi-directional –

- The user can export a list of data-layers from a search into a project
- New (derived or basic) layers created in a project, can also be added to the general portal
- The user is notified if the portal contain new layers that match its search criteria

The personal project – Simple + collaborative

- Multiple users can mange the same project under various permissions
- The user does not need to understand 'meta-data language' to use the portal or the projects

Why both general portal + projects are needed

For the researcher:

- Keep track of what layers they already have, applied for, used in analyses etc.
- May integrate the work of multiple researchers to one 'project' (e.g. ongoing work on a given landscape by various institutions)
- Dynamic update of the list of relevant layers
- Prioritize acquisition of data layers

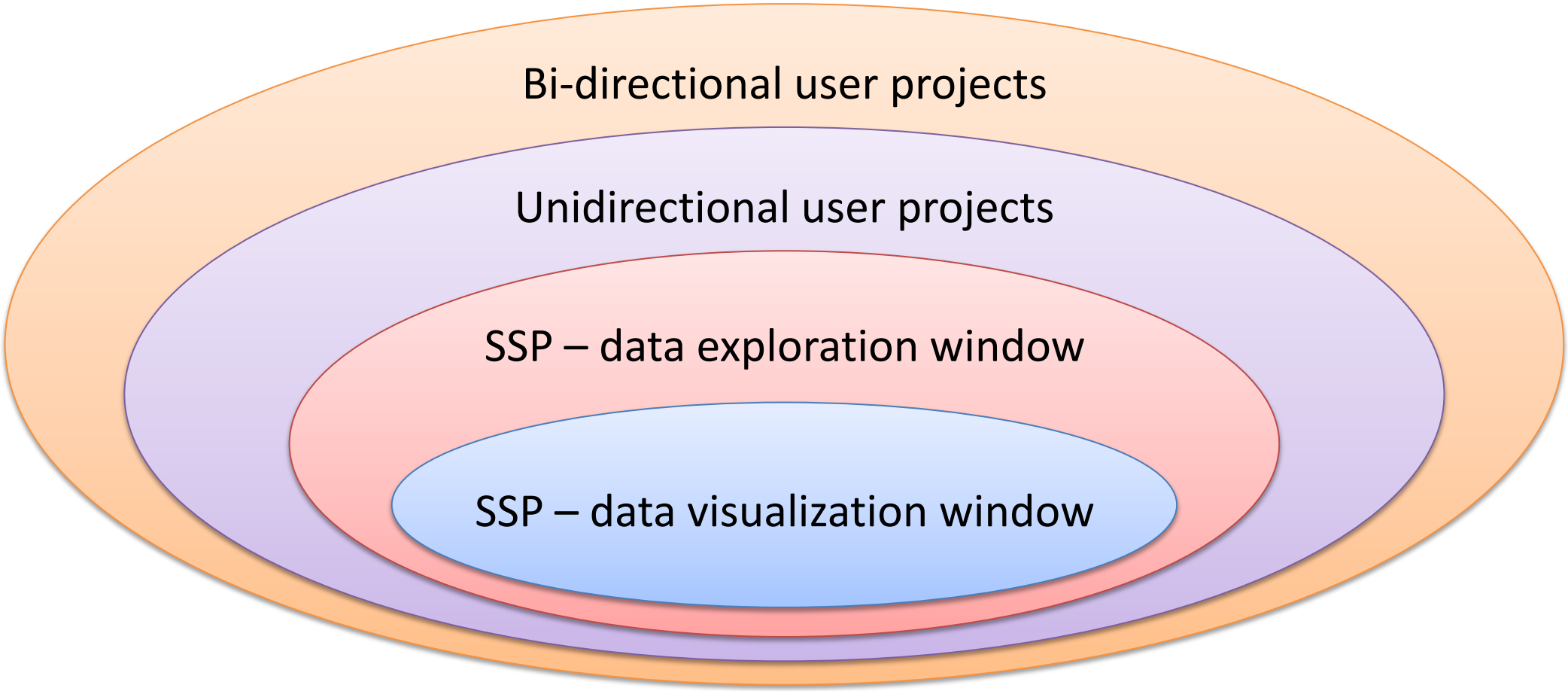
For the portal:

- Update portal content after EU-BON ends
- The users can provide quality control
- The users can add new data layers that currently goes '*under the radar*' –e.g., PhD projects
- Users may maintain it and contribute new functionalities (e.g., R and R-Forge)
- Accumulation of derived layers have yet unknown benefits… (e.g., SDMs)

Reality check...

- We cannot do it all within EU-BON
- Some probably cannot be done at all…
- First, identify impossible objectives
- Then, strategically prioritize our time investment, e.g.:
  - Drop the ‘data exploration window’ if:
    - Other aspects of the general portal are more important
    - We focus on the bi-directional user projects
  - Start with unidirectional user projects?
    - The user can download layers information from the portal, but cannot upload layers / quality controls/ etc. to the portal

Further development ...



4 (5) options for site-specific portal / project

Feature	SSP visualization	SSP visualization + exploration	Unidirectional project	Bi-directional project
Visualization of layers	yes	yes	yes	yes
User uploaded layers in project	yes	yes	yes	yes
Layer exploration tool	no	yes	Priority?	Priority?
User defined links archive	no	yes	Priority?	Priority?
Search results from portal	no	no	yes	yes
New layers notification	no	no	yes	yes
Multiple users for one project	no	no	yes	yes
Tools suggestion from portal	no	no	Priority?	Priority?
Export basic layers to portal	no	no	no	yes
Export derived layers to portal	no	no	no	yes
Quality control for data	no	no	no	yes
Potential maintenance by users	low	low	medium	high
New functionalities by users	low	low	medium	high

The site-specific portal team:

• Florian Wetzel	MfN
• Johannes Penner	MfN
• Francisco Antonio García Camacho	CSIC
• Simao Belchior	Simbiotica
• Hannu Saarenmaa	UEF
• Ruth Sonnenschein	EURAC
• Jerome O'Connell	UnivLeeds
• Mathias Kuemmerlen	SGN
• Stefan Stoll	SGN
• Francisco Javier Bonet García	UGR
• Eva Chatzinikolaou	HCMR
• Christos Arvanitidis	HCMR

Sierra-Nevada site

- Around 1000 data layers , not counting biotic dataset

Source	Category	Examples	# of layers
Basic	Environmental	DEM, geology, landscape units, hydrology, etc.	15
Basic	Environmental	Vegetation maps	15
Basic	Remote-sensing	NDVI and EVI from MODIS (2000-present)	stored in a database
Derived	Remote-sensing	Trends for NDVI and EVI from MODIS (2000-present)	stored in a database
Basic	Remote-sensing	NDSI (snow cover) from MODIS (2000-present)	stored in a database
Derived	Remote-sensing	Trends NDSI from MODIS (2000-present)	stored in a database
Basic	Remote-sensing	Land surface temperature from MODIS (2000-present)	stored in a database
Basic	Environmental	Spatial distribution of threatened species	100
Derived	Environmental	SDM for threatened species. Future climate projections.	100
Derived	Environmental	Yearly climate maps (1960-2010)	150
Derived	Environmental	Predicted Yearly climate maps (2010-2100)	600

Immediate benefits...

Stefan + Mathias (SGN)			Ruth (EURAC) + Jerome (UnivLeeds)		
Define site's needs + priority			Identify data and tools		
Need	Priority	Feasibility	Spatial resolution	Temporal resolution	RS Product

Immediate benefit for the sites...

Stefan + Mathias (SGN) Define site's needs + priority				Ruth (EURAC) + Jerome (UnivLeeds) Identify data and tools	
Need	Priority	Feasibility	Spatial resolution	Temporal resolution	RS Product
NDVI maps	high	easy	10 to 5m	monthly / seasonally	SPOT
Nutrients in agricultural areas	low	difficult	10 to 5m??	monthly / seasonally??	SPOT??
Biomass in terrestrial areas	medium	moderate	10 to 5m??	monthly / seasonally??	SPOT??
Photosynthesis/biomass in aquatic areas	low	difficult	10 to 5m??	monthly / seasonally??	SPOT??
Forest types	high	easy	10 to 5m	Seasonal	SPOT
Agricultural crops: intensity, rotations	high	easy	10 to 5m	monthly / seasonally	SPOT
Impervious surface	high	easy	10 to 5m	Yearly	Worldview 2, Kompsat, SPOT??
habitat fragmentation	high	easy	10 to 5m	Yearly	Geoprocessing based on improved layers
Shading of water areas	high	easy	10 to 5m	Seasonally	radar images & Geoprocessing
Water area extent during floods	high	moderate?	10 to 5m	circum-weekly?	Geoprocessing based on improved layers

Immediate benefit for the sites...

Stefan + Mathias (SGN) Define site's needs + priority			Ruth (EURAC) + Jerome (UnivLeeds) Identify data and tools		
Need	Priority	Feasibility	Spatial resolution	Temporal resolution	RS Product
NDVI maps	high	easy	10 to 5m	monthly / seasonally	SPOT
Nutrients in agricultural areas	low	difficult	10 to 5m??	monthly / seasonally??	SPOT??
Biomass in terrestrial areas	medium	moderate	10 to 5m??	monthly / seasonally??	SPOT??
Photosynthesis/biomass in aquatic areas	low	difficult	10 to 5m??	monthly / seasonally??	SPOT??
Forest types	high	easy	10 to 5m	Seasonal	SPOT
Agricultural crops: intensity, rotations	high	easy	10 to 5m	monthly / seasonally	SPOT
Impervious surface	high	easy	10 to 5m	Yearly	Worldview 2, Kompsat, SPOT??
habitat fragmentation	high	easy	10 to 5m	Yearly	Geoprocessing based on improved layers
Shading of water areas	high	easy	10 to 5m	Seasonally	radar images & Geoprocessing
Water area extent during floods	high	moderate?	10 to 5m	circum-weekly?	Geoprocessing based on improved layers