The role of objective trees in preservation planning

Hans Hofman

DCC/DPE/DRIVER/Nestor Joint workshop

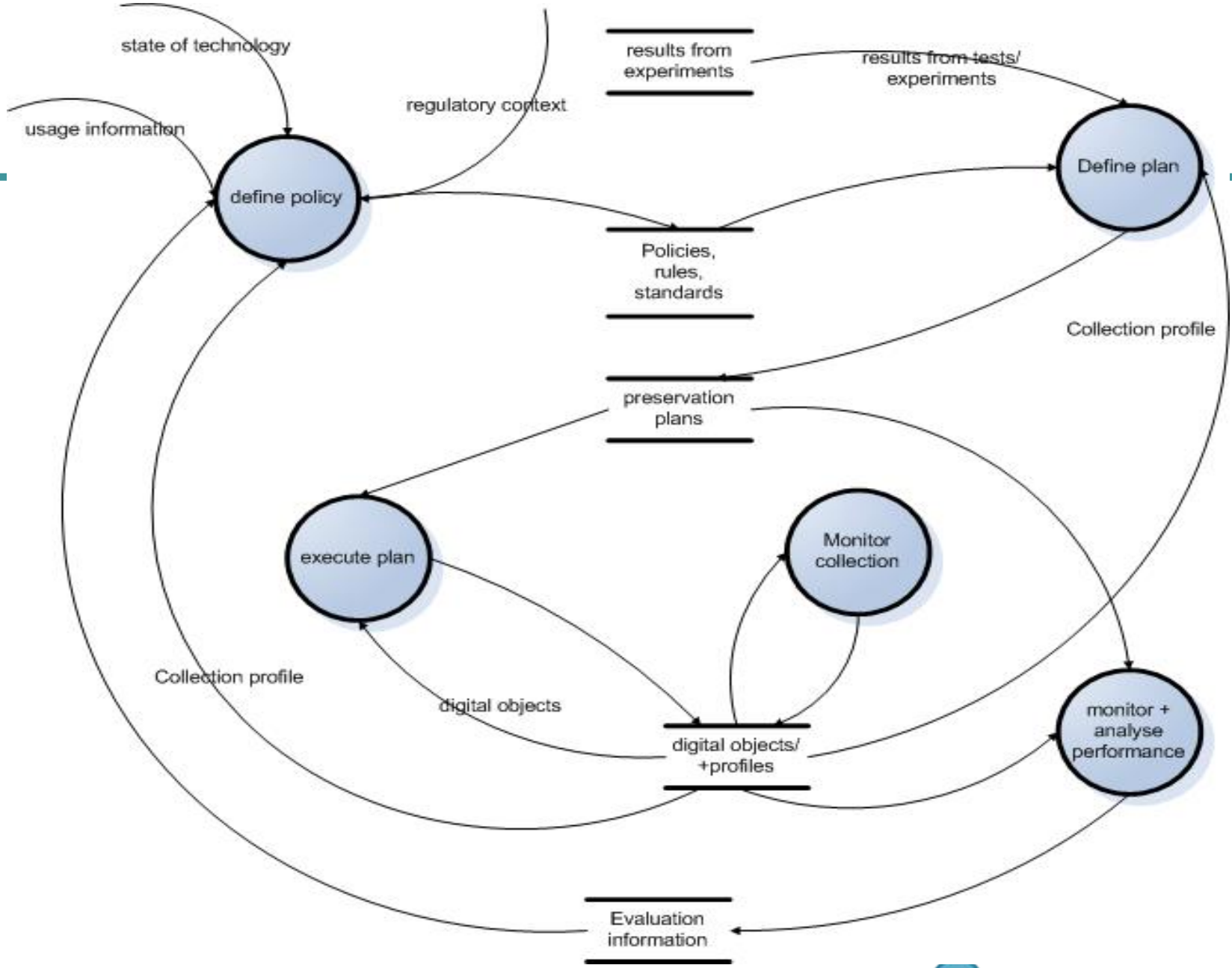
Berlin, 28 November 2007

Evaluating preservation actions

- Variety of solutions and tools exist
- Each action has unique strengths and weaknesses
- Requirements vary across (organisational) settings
- Decision on which solution to adopt is complex
- Documentation and accountability is essential

- Evaluation of preservation actions on representative sample content according to specific requirements
- Part of decision making in preservation planning

planets nationaal archief

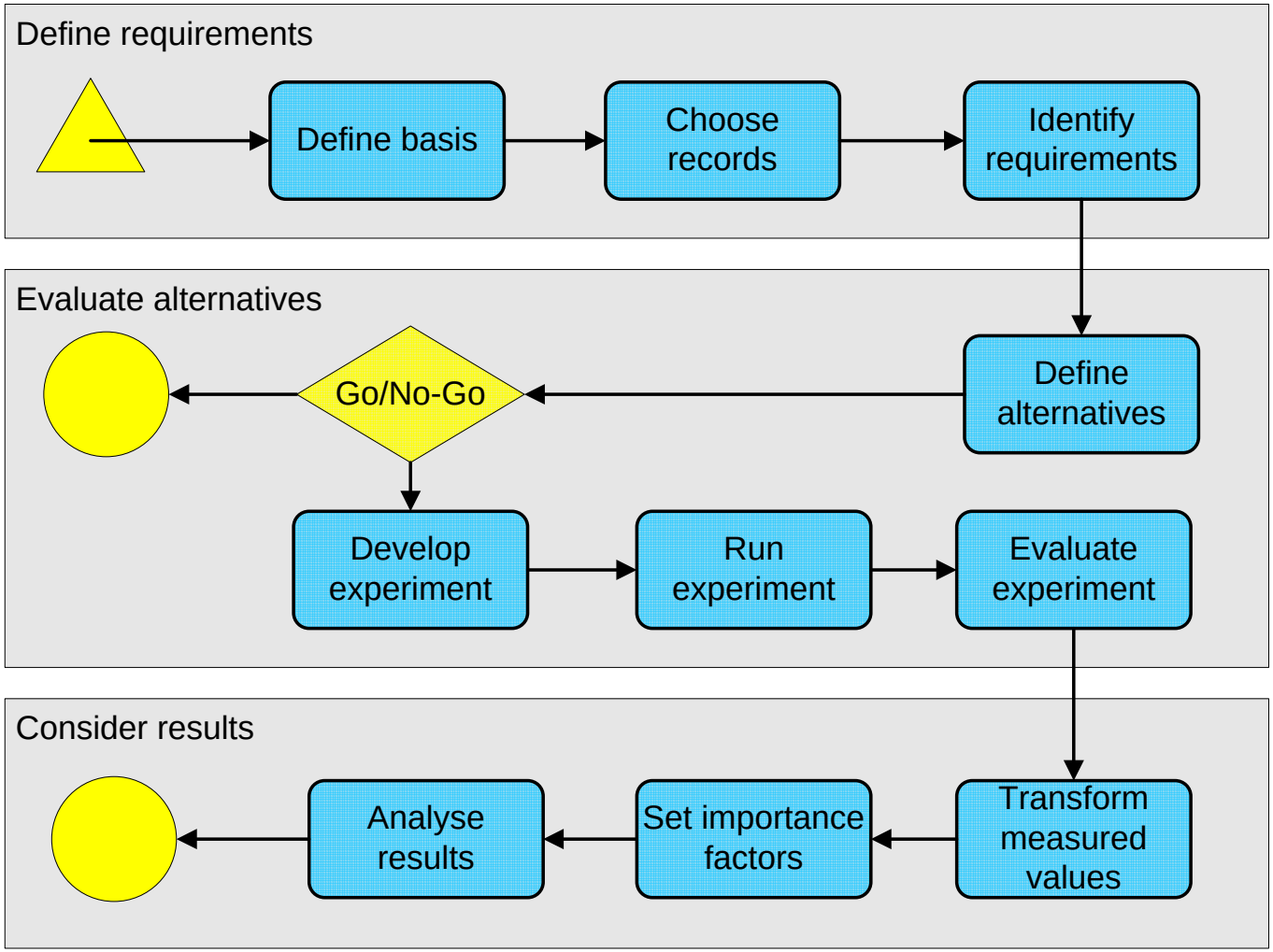

state of technology
usage information
define policy
regulatory context
results from experiments
results from tests/ experiments
Define plan
Policies, rules, standards
Collection profile
preservation plans
execute plan
Monitor collection
Collection profile
digital objects
digital objects/ +profiles
monitor + analyse performance
Evaluation information

Decision support for preservation planning

- Systematic procedure for evaluating preservation actions/strategies
  - By conducting experiments on sample content
  - Based on the Dutch Testbed and subsequently applied in DELOS
- Case studies
  - Electronic documents, interactive art, web archives...
  - Identify essential characteristics and requirements for preservation strategies
  - Validate methodology and workflow
- Development of software tool
  - Plato – Planning Tool
  - Web application supporting the workflow

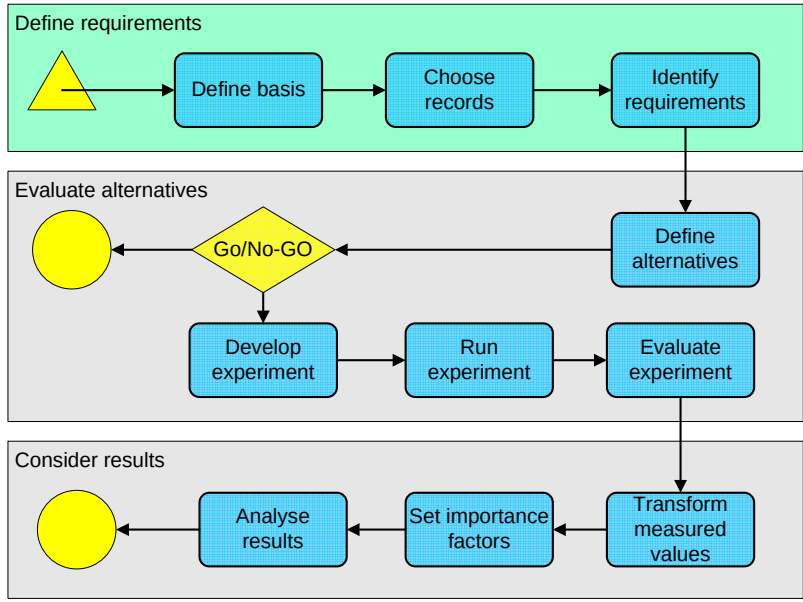
planets nationaal archief

Workflow

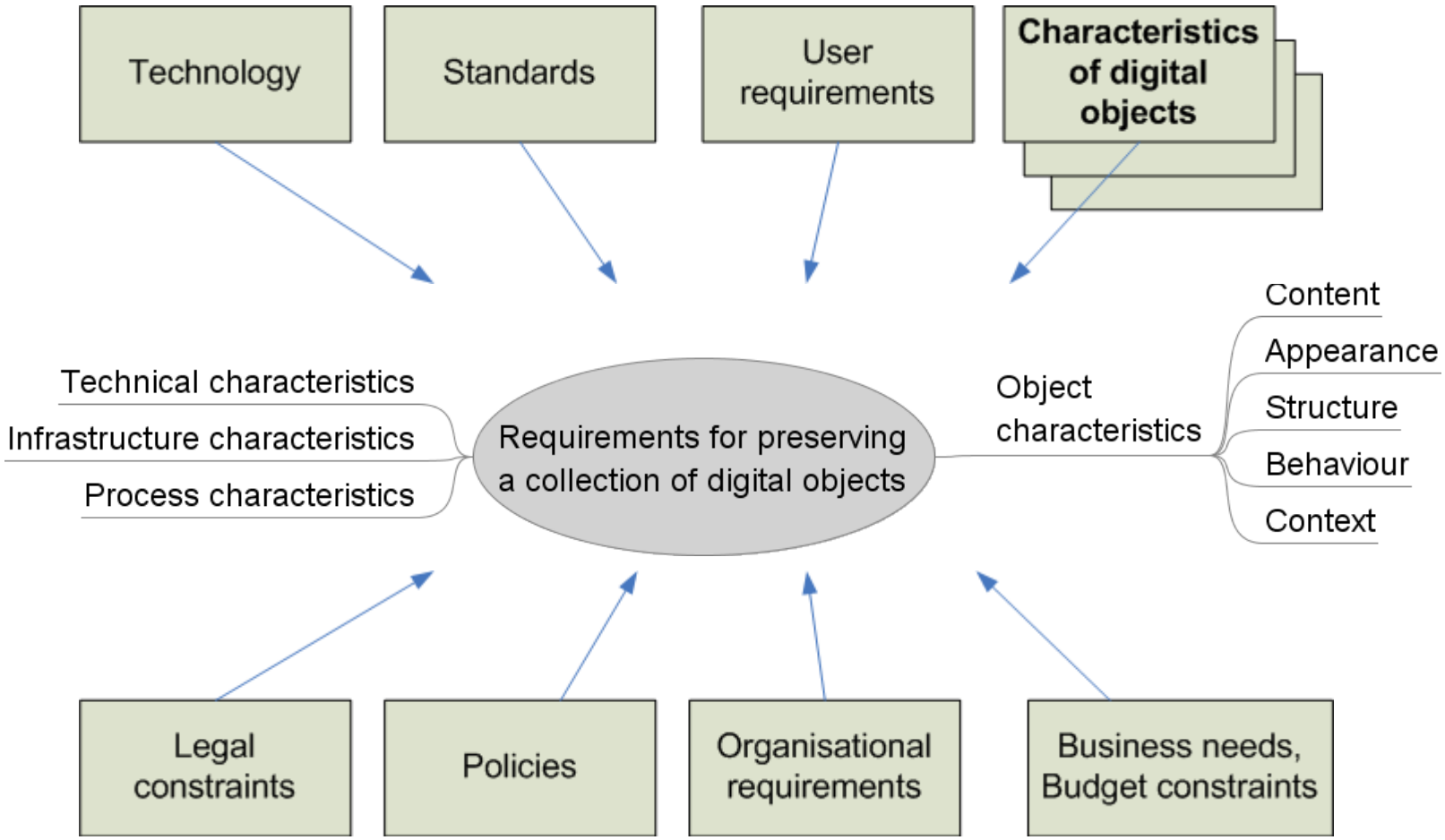

Define requirements
Define basis
Choose records
Identify requirements
Evaluate alternatives
Define alternatives
Go/No-Go
Develop experiment
Run experiment
Evaluate experiment
Consider results
Transform measured values
Set importance factors
Analyse results

Phase 1: Define requirements

1. Define basis
   - Describe Collection (profile)
   - Institutional settings
2. Choose sample objects/records
   - Representative for the type of objects that requires action
   - Right choice of samples is essential
3. Define requirements
   - “Objective tree”

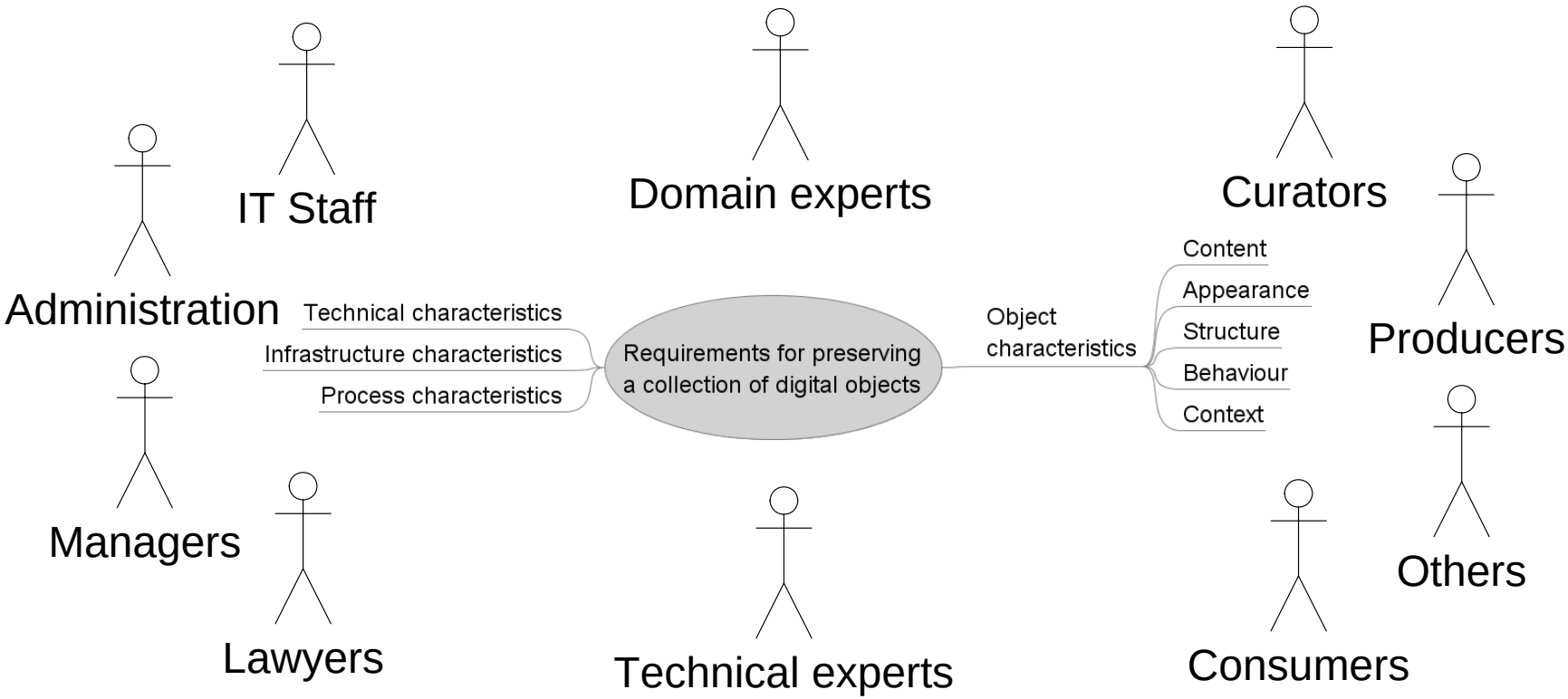


Influence Factors

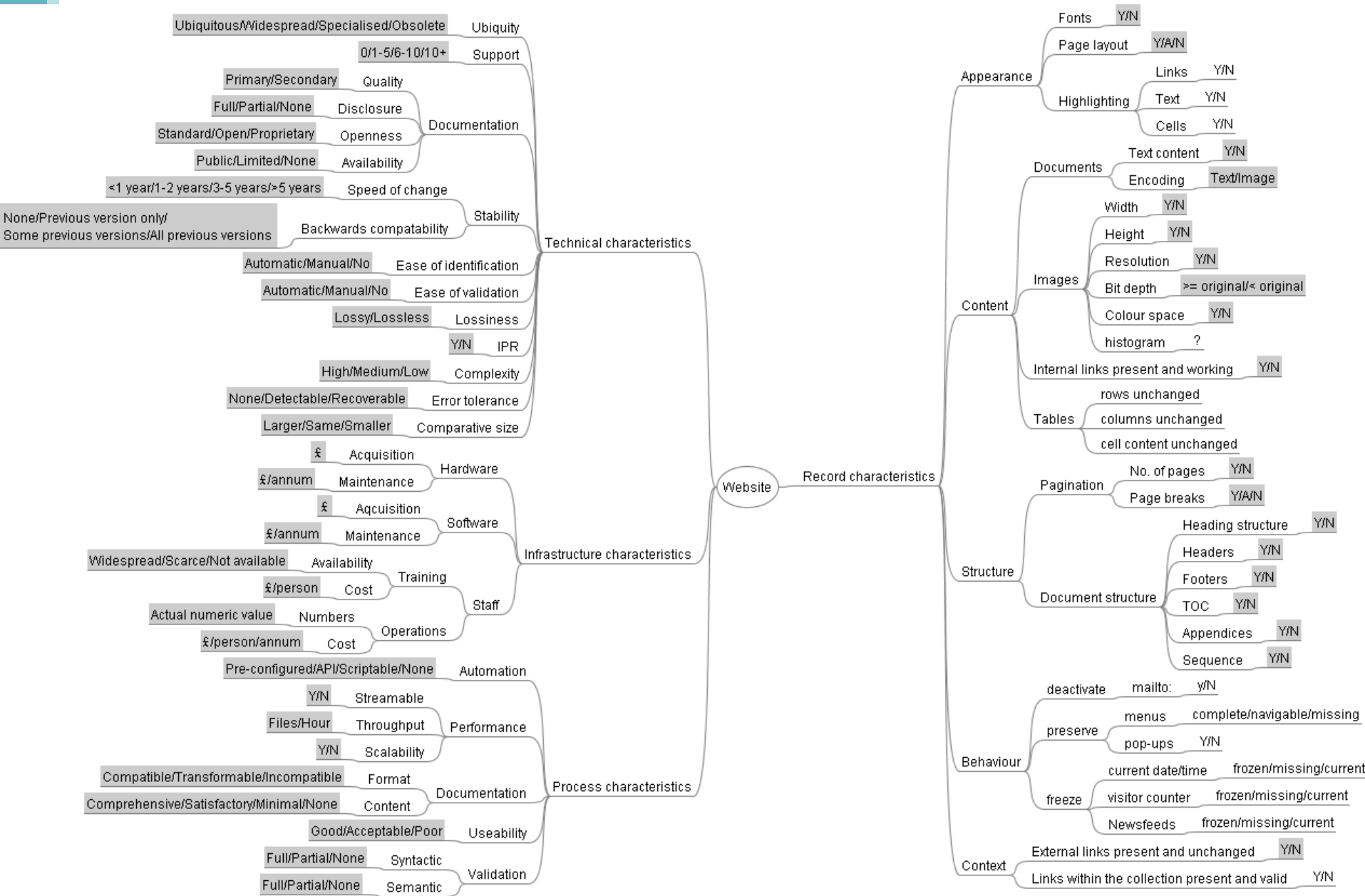


Stakeholders

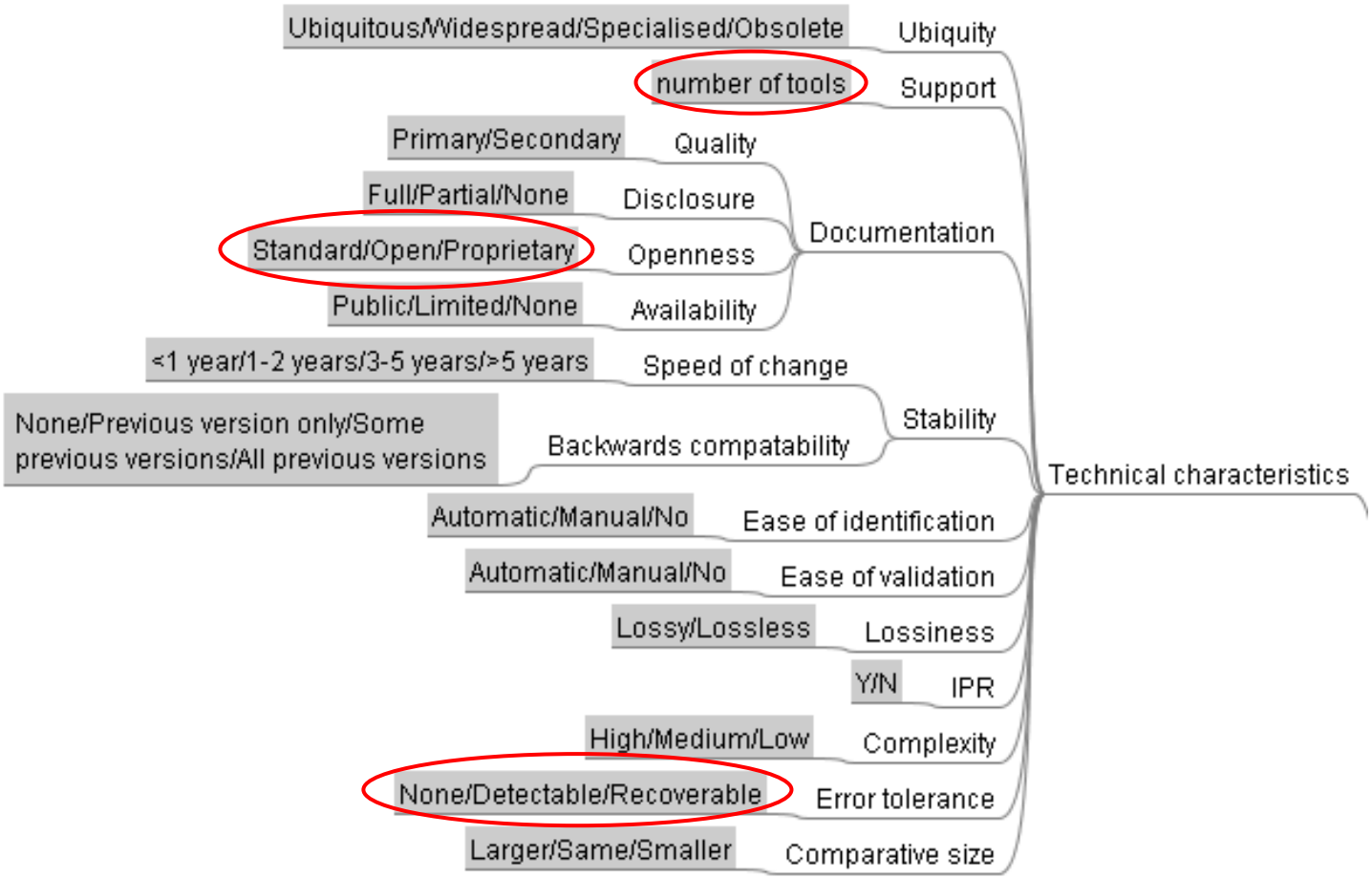
- Input from a wide range of persons, depending on the institutional context and the objects



The Objective Tree

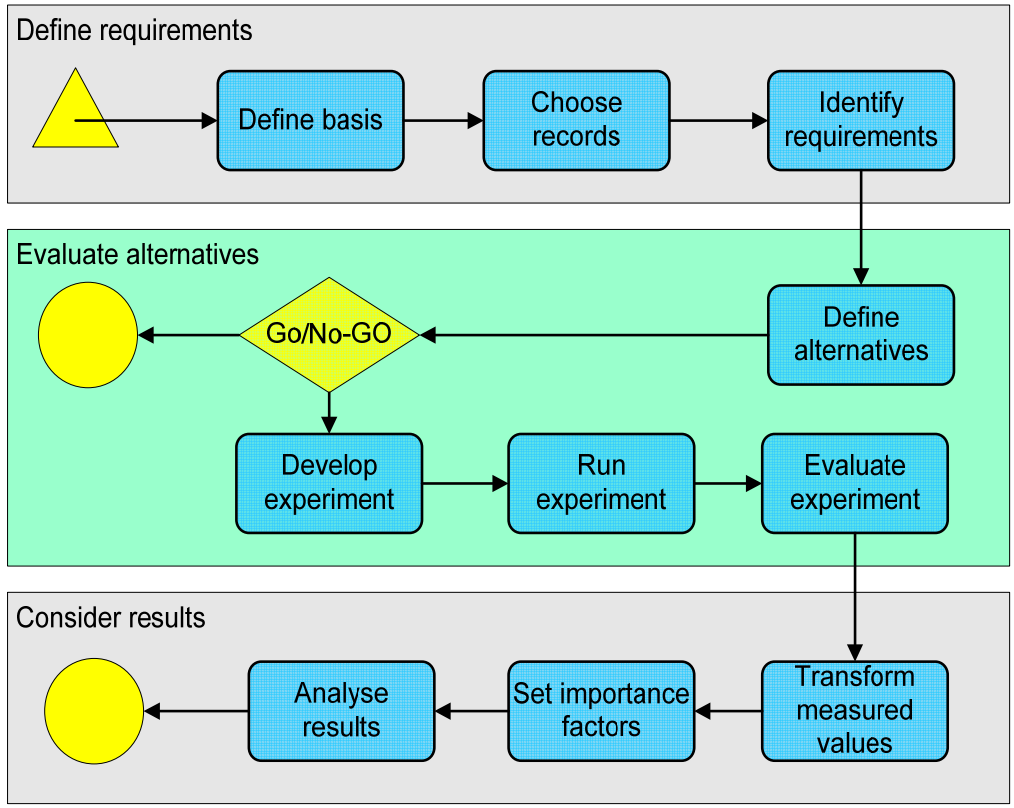


Assigning Scales



Phase 2: Evaluate Alternatives

4. Define Alternatives
5. Go/No-Go decision
6. Develop experiment
7. Run experiment
8. Evaluate experiment



Evaluate Experiment

Expand All | Collapse All

Website

Focus	Node
	▼Website
X	▼Record characteristics
X	▶Appearance
X	▶Content
X	▶Structure
X	▼Behaviour
X	▶deactivate
X	▶preserve
X	▶freeze
X	▶Context
X	▶Technical characteristics
X	▶Infrastructure characteristics
X	▶Process characteristics

Process characteristics > Automation

Alternative	Single result
solutionA	Pre-configured
solutionB	API

Performance > Streamable

Alternative	Single result
solutionA	Yes
solutionB	Yes

Performance > Throughput

Alternative	first	Unit	second	Unit
solutionA	0.0	files per hour	0.0	files per hour
solutionB	0.0	files per hour	0.0	files per hour

Performance > Scalability

Alternative	Single result
solutionA	No
solutionB	Yes

Documentation > Format

Alternative	Single result
solutionA	Compatible
solutionB	Incompatible

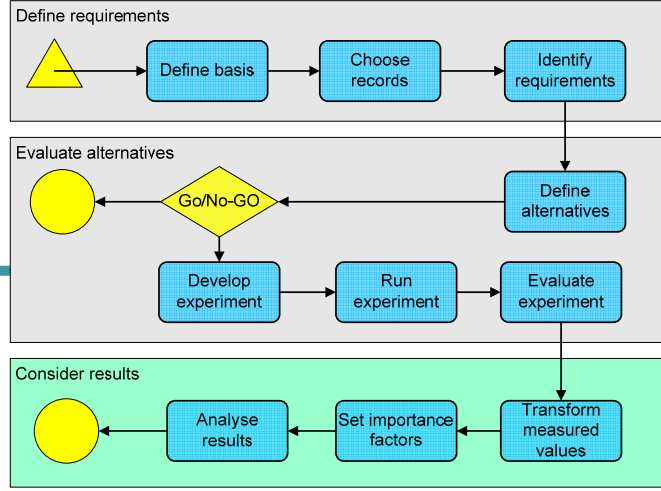
Documentation > Content

Alternative	first	second
solutionA	Comprehensive	Comprehensive
solutionB	Satisfactory	Satisfactory

Process characteristics > Useability

Alternative	Single result

Phase 3:
Consider Results

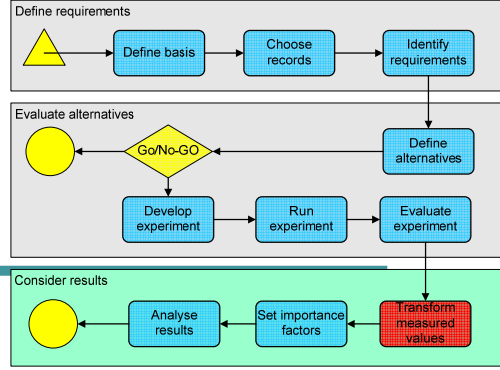


9. Transform measured values to a unified scale to make them comparable

10. Set importance factors to model the relative importance of siblings in each branch

11. Analyse results

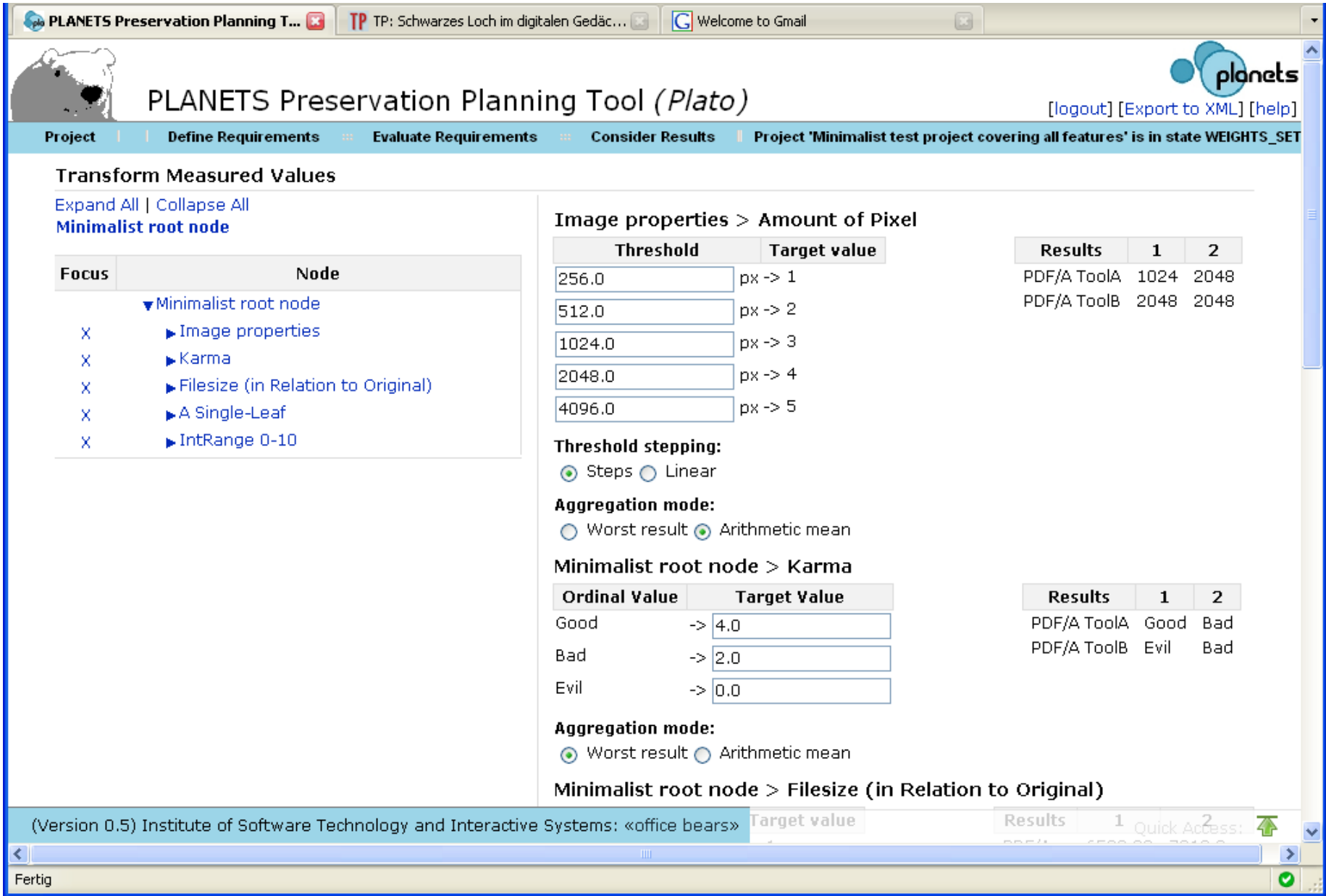
Transform measured values



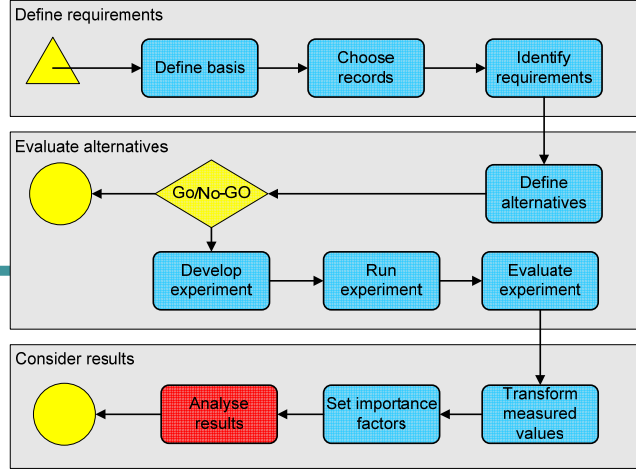
- Measures come in seconds, euro, bits, goodness values,…
- Need to make them comparable
- Transform measured values to uniform scale
- Transformation tables for each leaf criterion
- Scale 0-5 (0 is *unacceptable*)

Threshold	Target value
256.0	px -> 1
512.0	px -> 2
1024.0	px -> 3
2048.0	px -> 4
4096.0	px -> 5

Transformation



Analyse Results



- Aggregate values
  - Multiply the transformed measured values in the leaf nodes with the leaf weights
  - Sum up the transformed weighted values over all branches of the tree
- Rank alternatives according to overall performance value at root
- Performance of each alternative
  - overall
  - for each sub-criterion (branch)
- Comparison of different alternatives

PLANETS Preservation Planning Tool *(Plato)*

Analyse Results

Sum

☑	PDF/A (Tool A)
☑	PDF/A (Tool B)

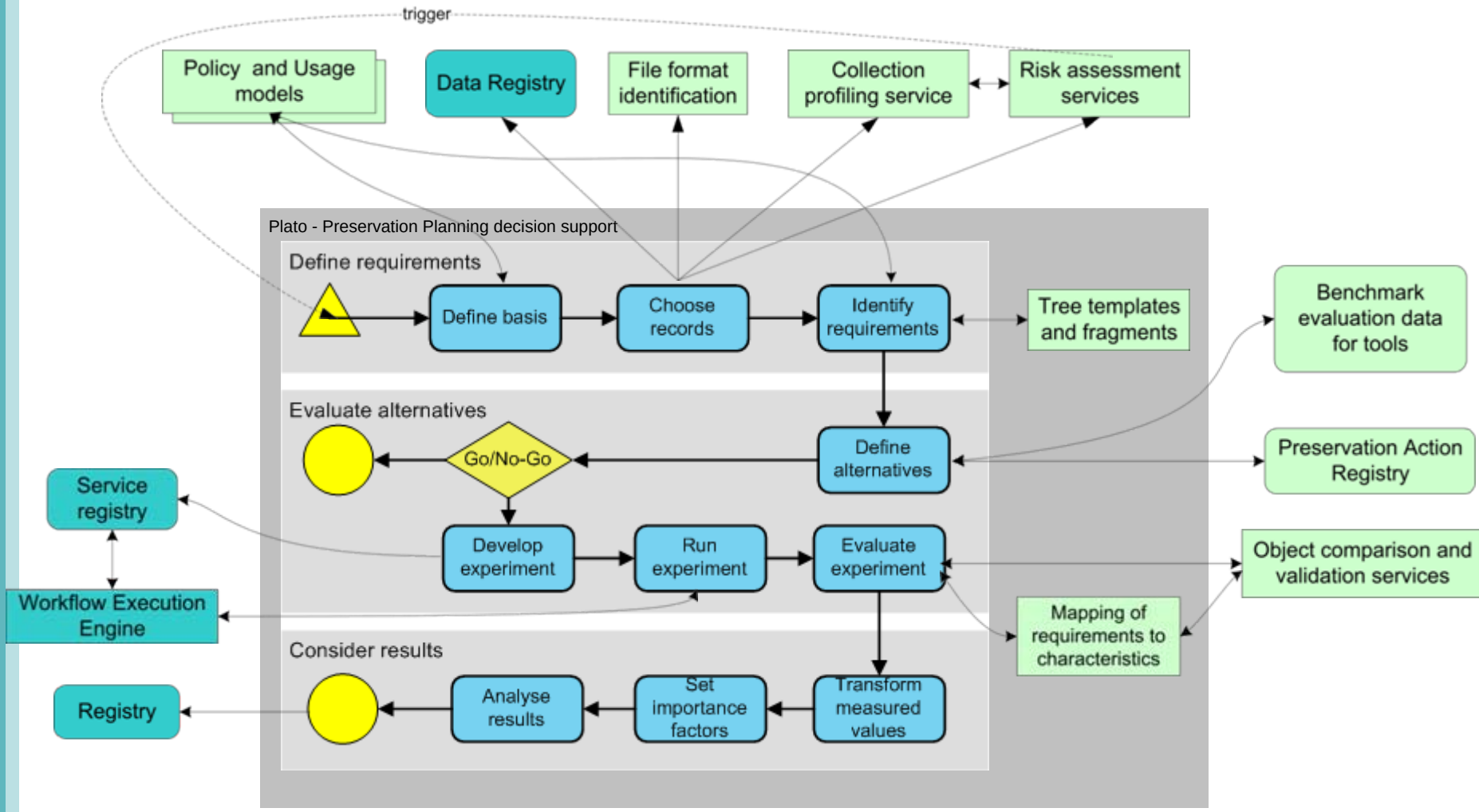
Show

Expand All | Collapse All

Minimalist root node

Focus	Name	Result
	▼Minimalist root node	PDF/A (Tool A): 2,98 PDF/A (Tool B): 3,19
X	▶Image properties	PDF/A (Tool A): 0,70 PDF/A (Tool B): 0,80
X	▼Karma	PDF/A (Tool A): 0,40 PDF/A (Tool B): 0,00
X	▼Filesize (in Relation to Original)	PDF/A (Tool A): 0,78 PDF/A (Tool B): 0,99
X	▼A Single-Leaf	PDF/A (Tool A): 0,40 PDF/A (Tool B): 0,80
X	▼IntRange 0-10	PDF/A (Tool A): 0,70 PDF/A (Tool B): 0,60

Integrating Planets concepts and services



Summary

- Systematic approach for identifying all criteria that will influence preservation planning

- Workflow for evaluating and choosing preservation actions

- Tool support: Plato
  - 1st version end of November 2007 (project internal)
  - 2nd version publicly available, second half of 2008

- Planets: developing one integrated environment for preservation planning

Thank you very much for your attention.

hans.hofman@nationaalarchief.nl
www.planets-project.eu