



IDCC14 Workshop:

Costing Curation: are we on the right track?

Presented by: Neil Grindley, Katarina Haage, Joy Davidson, José Borbinha, Rachel Bruce

PROGRAMME

Timings	Duration	Activity
Section 1:		
09:00 - 09:30	30 mins	Introduction to the 4C project and the costs of curation (presentation)
09:30 – 10:00	30 mins	What is your organisation interested in? (Presentation and Q&A using the 4C Indirect Economic Determinants and the more broadly defined benefits of curation)
10:00 – 10:30	30 mins	How do different organisations count the cost of curation? (Exemplars & participants invited to briefly share experiences)
10:30 – 11:00	30 mins	<i>Break</i>

Project Summary


The Collaboration to Clarify the Costs of Curation (4C) project will help organisations across Europe (and beyond) to more effectively invest in digital curation and preservation.

Vision

The 4C vision is to create a better understanding of digital curation costs through collaboration.

Mission

Our mission is to provide useful, useable resources which support the process of cost management in digital curation.





Assessment

Tasks

- Assess cost models & strategies
- Examine good practice
- Analyse requirements
- Integrate components
- Produce guidance & briefing materials
- Setup costs exchange

Engagement

Tasks

- Engage stakeholders
- Raise awareness
- Organise meetings
- Promote Research & Innovation
- Build community network



Collaboration to
Clarify the
Costs of
Curation

Networking &
Coordination



Affiliate Partners
& Stakeholders

Events,
Workshops,
Meetings &
Reports

Enhancement

Tasks

- Examine and refine related concepts
- Value
- Risk
- Benefits
- Sustainability
- Economic Reference Model



Jisc

Project Coordination

Tasks

- Project meetings
- Project reporting
- EC liaison
- Budget oversight
- Outputs QA



Outputs



Reports for
General
Dissemination



Curation
Costs
Exchange



Reports for
European
Commission



Submission
of Roadmap
to the EC



Collaboration to Clarify the Costs of Curation



TIMELINE

May 2011

Nov 2011

Apr 2012

Feb 2013

Preparation



Project
Kickoff

Who we are
& what we do



iPRES2013

Summer 2013

Contact
Stakeholders

Participation,
debate, emerging
findings



Archiving



Summer 2014

Emerging
Resources

Dissemination,
legacy,
recommendations

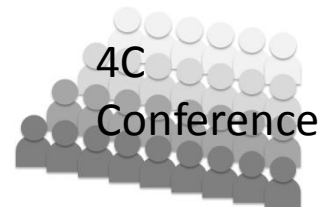
Economics of
Digital Curation
Roadmap



Jan 2015

Project Close

Oct 2014



4C
Conference

Curation Costs
Exchange



Why should we concern ourselves about the cost of curation? (What are the stakeholders saying ...)

- Understanding the cost of preservation may mean we can offer realistic and cost effective curation services to others.
- Understanding costs can support strategic planning.
- Understanding costs can support tactical decision-making.
- Understanding costs can provide evidence of cost-effectiveness and value.
- Clarifying and publishing the cost of digital curation can be used to enhance our organisation's credibility. But this must be done along with the context of how the costs were calculated
- Understanding economic drivers can help to strategically align an organisation

How is 4C going to help?

By carefully analysing all of the information we assemble and making sure that it is passed onto our stakeholders through ...

- A series of state of the art reports
 - Cost model evaluation and a needs & gap analysis
 - Trust and quality (in relation to the cost of curation)
 - Risk, benefit, impact and value (in relation to the cost of curation)
 - From cost models to business models
 - A roadmap for future economic considerations in relation to digital curation
- New frameworks and models to assist with designing new approaches and building future tools
 - Indirect economic determinants
 - An economic sustainability reference model
 - A gateway specification for future cost models
 - A cost concept model for digital curation
- A Curation Costs Exchange

- Engagement
 - Make sure all outputs are available for public
 - Get involved and build partnerships with individuals, groups and institutions that are active or interested in the topic of curation costs
 - Build a community network
 - Organise webinars, focus groups and other events to connect people
 - Provide a platform for exchange, interaction and discussion (CCEx)

What is your organisation interested in?

The 4C Indirect Economic Determinants (IED)

- 15 indirect economic determinants that are significant for digital curation
- Generic management tools to help ensure sustainable digital curation

Goals?

- > Indicators
- > Support
- > Feature within the ESRM

What the consultation showed:

- > Risk, trustworthiness and benefits are ranked with high importance
- > If grouped: 1) Risk/trustworthiness, 2) Sustainability, 3) Data protection issues

... and the more broadly defined benefits of curation

- direct
- indirect
- near term
- long-term
- private
- public

When it comes to digital curation: what are the three most basic needs the German National Library would choose?

- Access
- Integrity
- Authenticity

Access

- Infinite
- At any time
- Also: preserving the interpretability of the digital data
- Challenged by: constant changes in hard- and software

Integrity and authenticity

“The digital repository ensures the **integrity** of the digital objects during all processing stages.” (p.19f)

Integrity here refers firstly to the completeness of the digital objects and secondly to their intactness. The yardsticks for integrity are the characteristics of a digital object defined as worthy of preservation (cf. 9.2.).

“The digital repository ensures the **authenticity** of the digital objects during all stages of processing.” (p. 21f)

Authenticity here means that the object is genuine, i.e. that it represents, what it claims to represent. It also includes full documentation of all transformations to the objects carried out for the purpose of preservation.

http://files.d-nb.de/nestor/materialien/nestor_mat_08_eng.pdf



Exercise - Indirect Economic Determinants

To what extent would your organisation regard the following 15 outcomes as an investment priority?

IED	High	Medium	Low	Notes
Authenticity				
Benefit				
Efficiency				
Impact				
Innovation				
Interoperability				
Quality				
Reputation				
Risk				
Sensitivity				
Skills				
Sustainability				
Transparency				
Trustworthiness				
Value				

Neil ...

How do different organisations count the cost of curation?

(Exemplars & participants invited to briefly share experiences)

We need to think carefully about what exactly *is*
the problem that we are trying to solve ...

If they want to, organisations can work out how
much it costs them to manage their digital assets

4C Data Gathering Exercise Organisation A

Curation Categories

Pre-Ingest

Ingest

Preservation Planning

Data Management

Archival Storage

Access

Administration

Accounting Principles

Labour Direct	Labour Indirect	Capital Direct	Capital Indirect	Note

0	0.2	0.1	0.0	0.7	0.6	1.3			€ 5,000	€ 5,000
0	0.2	0.1	0.0	0.7	0.6	1.3			€ 5,000	€ 5,000
0	0	0.1	0.0	0.0	0.3	0.3			€ 1,000	€ 1,000
0	0	0.1	0.0	0.0	0.3	0.3			€ 1,000	€ 1,000

Time Period
2012

Total Cost
€252,000

4C Data Gathering Exercise Organisation B

Curation Categories

Ingest

Curation

Access

Cost Categories

Hardware	Software	Employment	Accommodation	External Services	Transfer
----------	----------	------------	---------------	-------------------	----------

Accounting Principles

Direct	Costs of Service	Absorbed Indirect Costs of Service	Unabsorbed Indirect Costs of Service
--------	------------------	------------------------------------	--------------------------------------

Time Period
2012

Total Cost

€645,683.26

4C Data Gathering Exercise Organisation C

Curation Categories

Ingest

Data Management

Archival Storage

Preservation Planning

Access

Administration

Common Services

Accounting Principles

Labour Costs	Capital Costs	Offset By Revenue
--------------	---------------	-------------------

Size of Collection
393 TB

Time Period
2012

Total Cost
€15,800,000

4C Data Gathering Exercise Organisation D

Curation Categories

Ingest

Archival Storage

Metadata Management

Access

Administration

Cost Categories Product/Service

Requirements	Customisation	Integration	Installation	Training	Support
--------------	---------------	-------------	--------------	----------	---------

Cost Categories Hardware

Time Period
?

Total Cost
€349,665

4C Data Gathering Exercise Organisation E

Curation Categories

Digital Archiving

Cost Categories

Hardware	Software Maintenance	Software Development	Staff Other

Time Period
2007

Total Cost
123,000

Time Period
2012

Total Cost
348,500



4C Data Gathering Exercise
Organisation F

Curation Categories
Long term Digital
Preservation

Size of Collection
2 TB

Assumption
€8k per TB
per year for
Storage Costs

Accounting Principles
Development & Improvement
Operation

Staff
Development, Technical Support
Training, Communications, Public Relations
Expenses
Software Design
Software Licenses
Support
External Development
Hardware Purchase
Hardware Operating costs
Graphic Design

Time Period
2007-2012

Total Cost
€205,000

4C Data Gathering Exercise Organisation G

Curation Categories

Digital Archiving

Content Management

Data Development

Hardware

Production

Systems Development

Delivery

User Support

Overhead

Operations

Management

Time Period
2012-13

Total Cost

€ 3,130,110

So ... What exactly are the problems we need to tackle?

- The random numbers problem - How can we meaningfully compare the numbers that we end up with? [*cost data*]
- Activity based costing versus financial accounting methods
- Describing what the organisation does [*cost metadata*]
- Describing the amount and type of data that is being looked after [*cost metadata*]
- Sensitivity around data – Many organisations are not particularly happy to broadcast what it costs them to manage their data. How can we effectively anonymise the sharing of data?
- Complexity - The detail builds up very quickly across different organisations and it doesn't map together easily
- And we somehow have to make sure that the benefits are presented alongside the costs

Discussion ...

PROGRAMME

Section 2:

11:00 – 11:30	30 mins	Introduction to the 4C draft Cost Concept Model (Presentation)
11:30 – 12:30	1 hour	How would you break down the cost of curation? (Exercise in small groups supported by 4C team member)
12:30 – 13:30	1 hour	<i>Lunch</i>

It starts with the



4C Stakeholder Consultation

Which informs the ...

D3.1—Evaluation of Cost Models and Needs & Gaps Analysis (MS12 Draft)

<i>Deliverable Lead:</i>	Det Kongelige Bibliotek (KBDK)
<i>Related Work package:</i>	WP3—Assessment
<i>Author(s):</i>	Ulla Bøgvad Kejser (KBDK) Kathrine Hougaard Edsen Johansen (DNA) Alex Thirifays (DNA) Anders Bo Nielsen (DNA) David Wang (SBA) Stephan Strodl (SBA) Tomasz Miksa (SBA) Joy Davidson (HATII-DCC) Patrick McCann (HATII-DCC) Jaan Krupp (NLE)
<i>Dissemination level:</i>	Public

URL ...

Which will be further elucidated, enhanced and expanded with the ...

The 4C Cost Concept Model

... & Gateway Requirement Specification

The goal of this task is not to create a single unified functional implementable cost modelling application; rather it is to design a common model based on common concepts and a generic specification (a gateway specification) that can be used in follow-up R&D projects. [4C Description of Work]

Digital Preservation Cost Modelling: Where did it all go wrong?

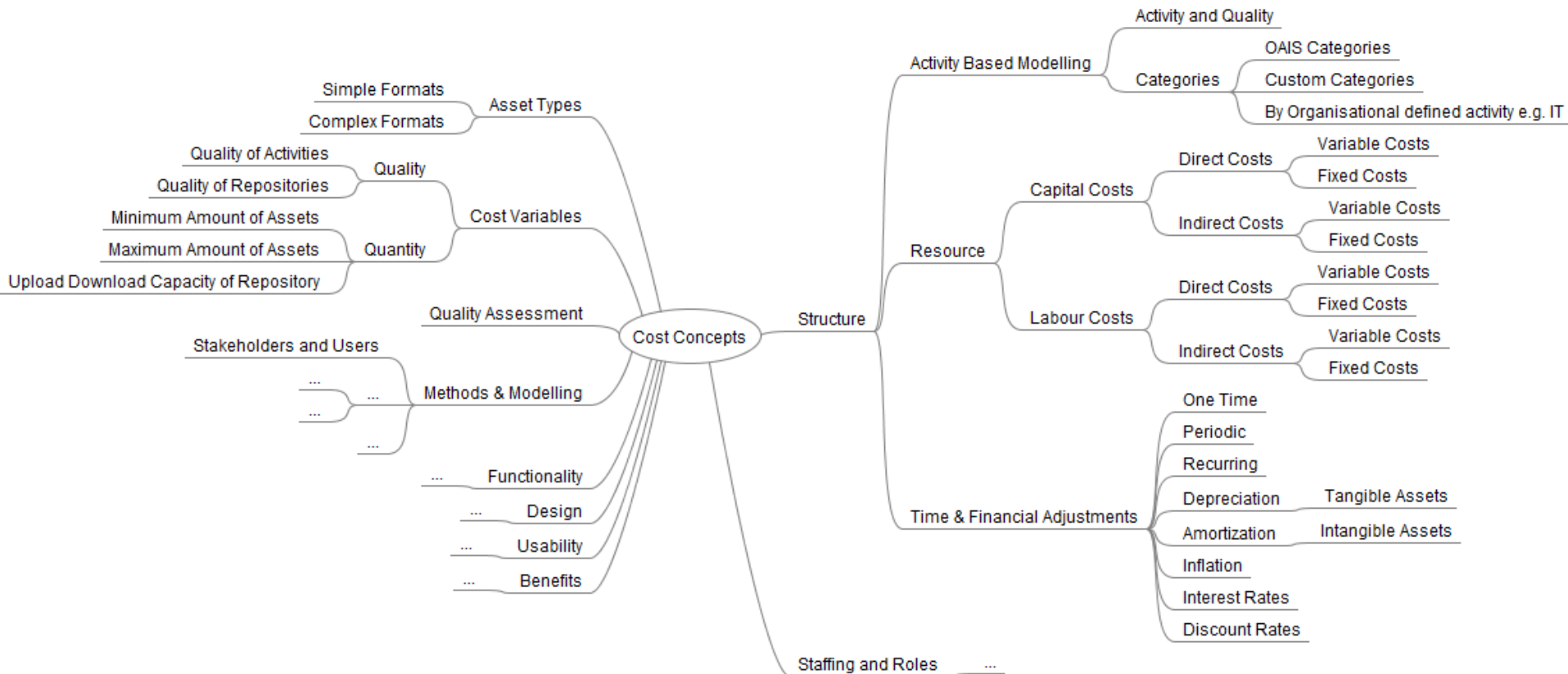
By paul on 29 June 2012 - 2:08pm



I recently spoke at a workshop on digital preservation costing organised by the lovely people at Knowledge Exchange and Nordbib. After briefly covering some of the work I was previously involved in as part of the LIFE Projects, I talked about why I think that

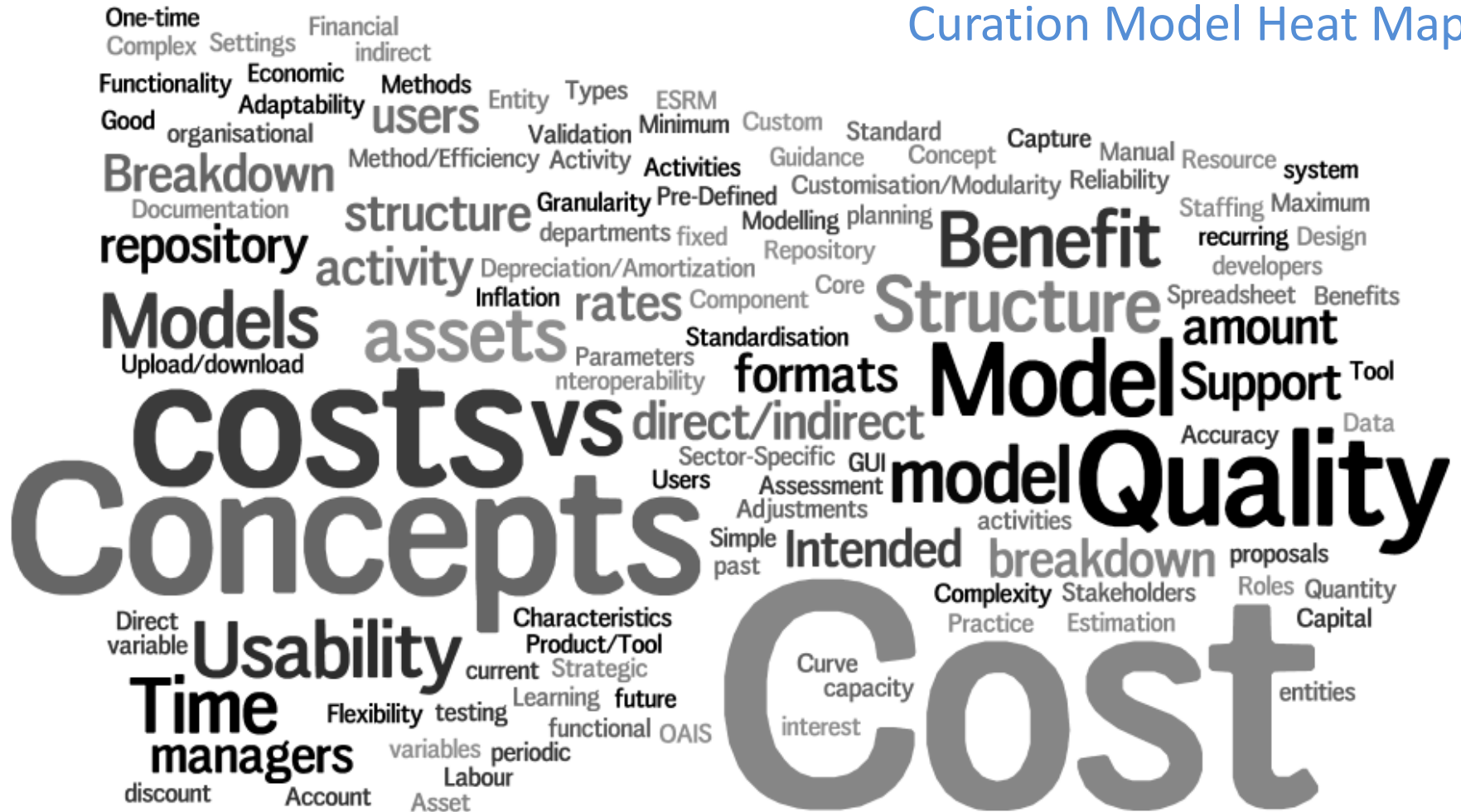
<http://openplanetsfoundation.org/blogs/2012-06-29-digital-preservation-cost-modelling-where-did-it-all-go-wrong>

We will create a Cost Concept Model (CCM) but we don't know what it looks like at the moment ... The Gateway Requirement Specification (GRS) will contain the detail – the CCM will need to be more schematic / accessible ...



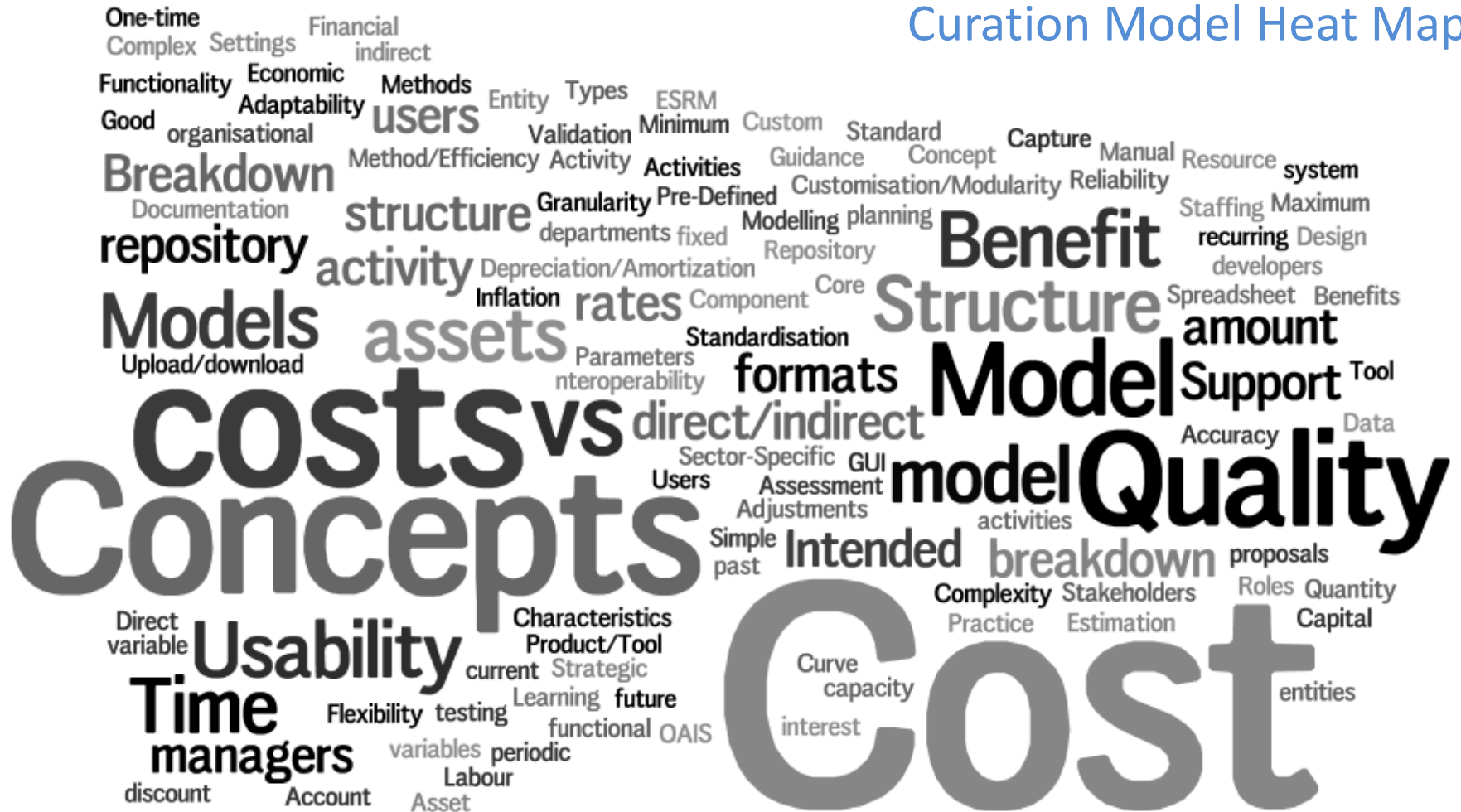
Model type	T-CMDP	NASA-CET	LIFE3	KRDS	CMDA	CMDP	DP4lib	PP-CMDS	CDL-TPC	EMLTS
-------------------	--------	----------	-------	------	------	------	--------	---------	---------	-------

Curation Model Heat Map



Model type	T-CMDP	NASA-CET	LIFE3	KRDS	CMDA	CMDP	DP4lib	PP-CMDS	CDL-TPC	EMLTS
-------------------	--------	----------	-------	------	------	------	--------	---------	---------	-------

Curation Model Heat Map



Only 15% of people in the stakeholder consultation indicated that they had tried to use a cost model

What can be done to ensure that models are developed in line with users' needs?

15 drivers for development are listed in the Needs & Gaps report along with 11 recommendations

Draft good practice proposals for model developers from GRS/CCM Framework:

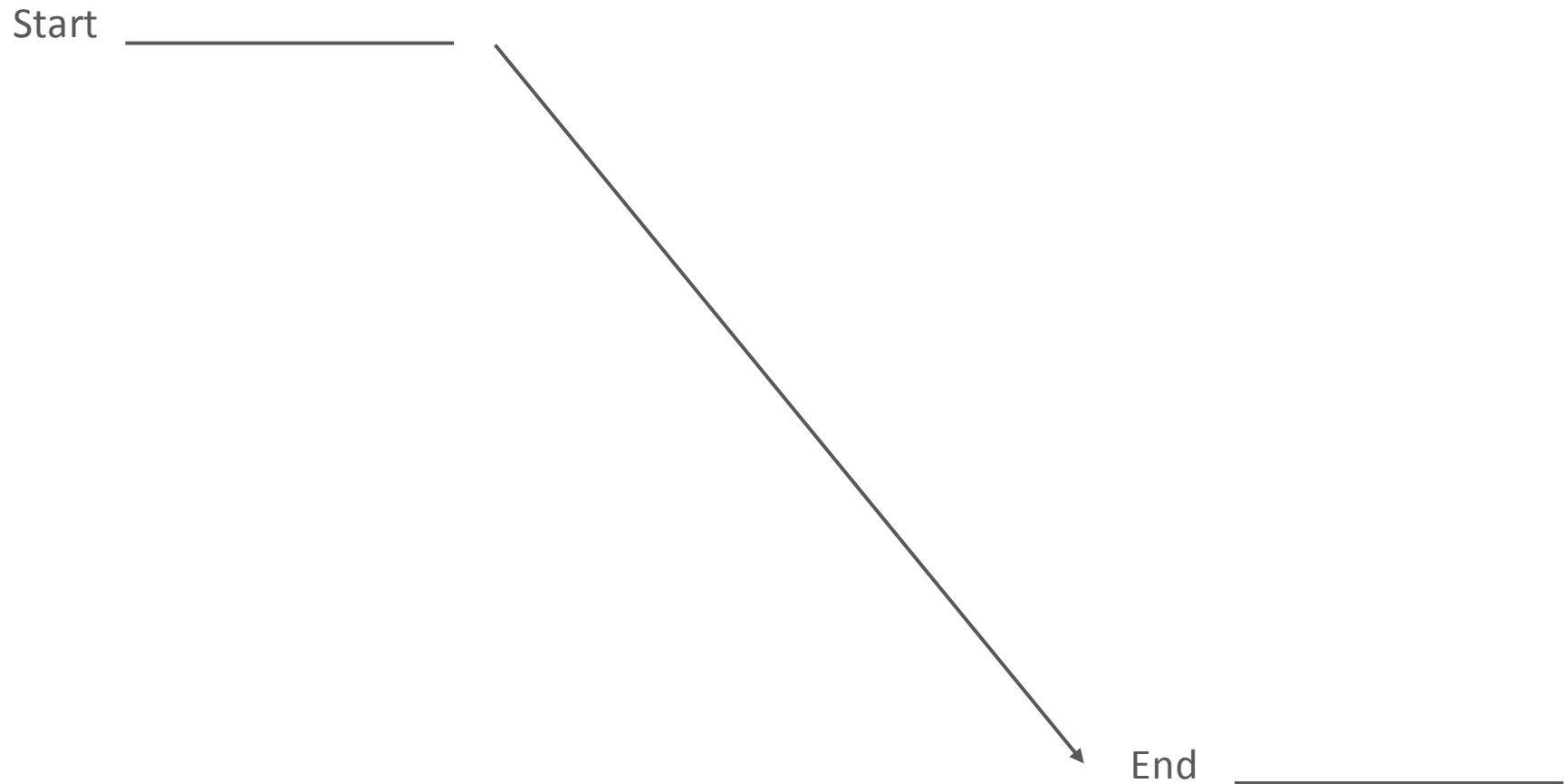
1. Use a standardised definition of digital curation
2. Limit the purpose of the model and define clearly the expected users of the model and its scaling capacity
3. Start out and continue to prioritise simplicity; be explicit about limitations on accuracy
4. Limit the time scope
5. Use simple formulae
6. Implement the model in a simple and widespread tool

So the CCM and GRS should help in various ways ...

- It builds on and raises awareness of the 4C stakeholder engagement and Needs & Gaps work
- It provides a detailed analytical foundation for the Curation Costs Exchange approach
- It will be a one-stop shop for thinking about the components of cost models
- It will enable comparisons of various existing cost models
- It will help people to design their own modelling approaches
- It will help to build consensus around definitions and terminology
- It will be an accommodating structure to showcase and raise awareness of future cost modelling work
- It will feed into the 4C Roadmap work which will recommend future activity in relation to further clarifying the economics of curation

Exercise

A journey that you are familiar with ...



Exercise

A journey that you are familiar with ...

Start _____

1 _____

2 _____

3 _____

4 _____

5 _____

End _____

What strikes you about
your journey?

Exercise 1

A journey that you are familiar with ...

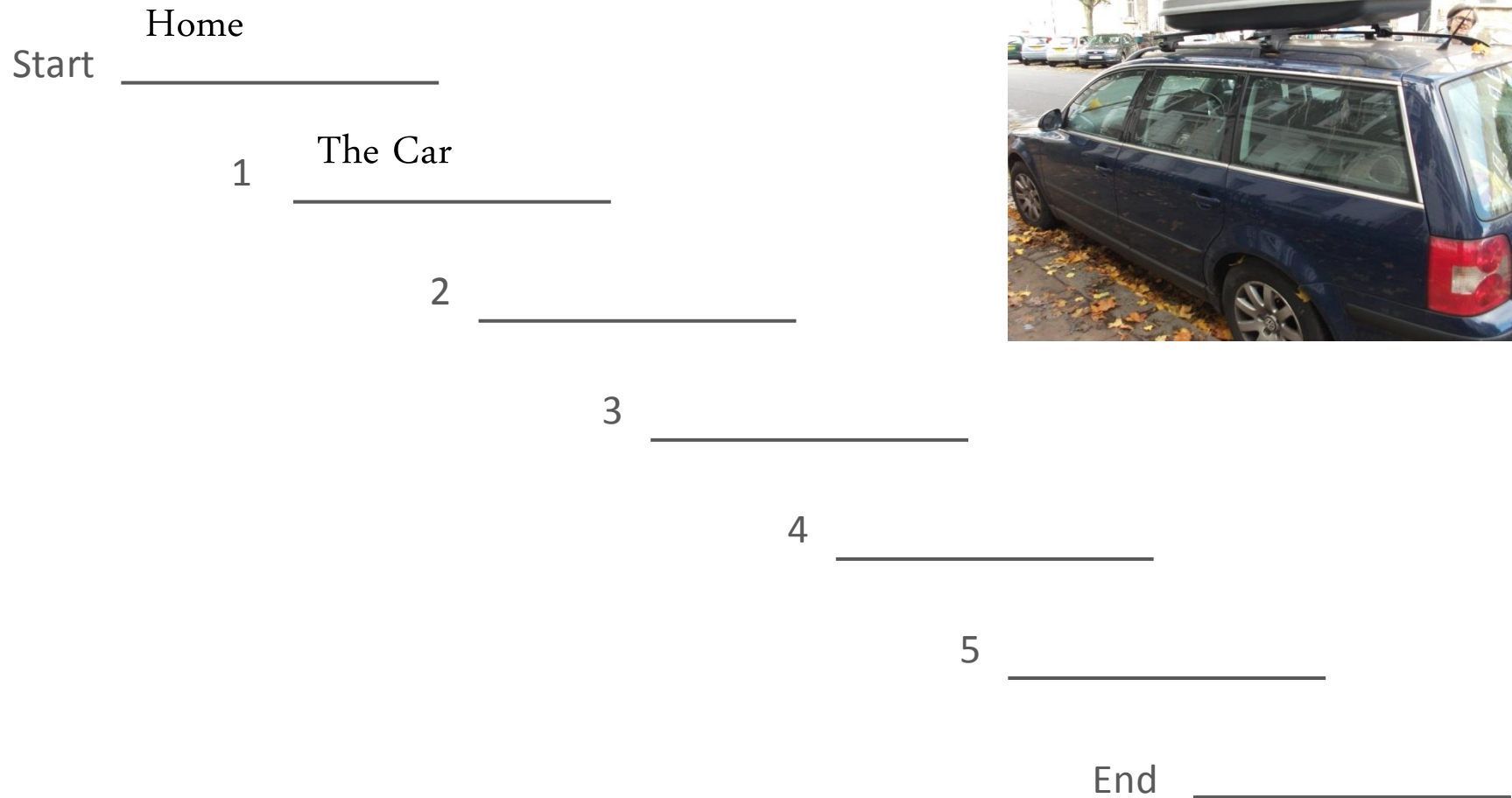
Start Home



End Gymnasium

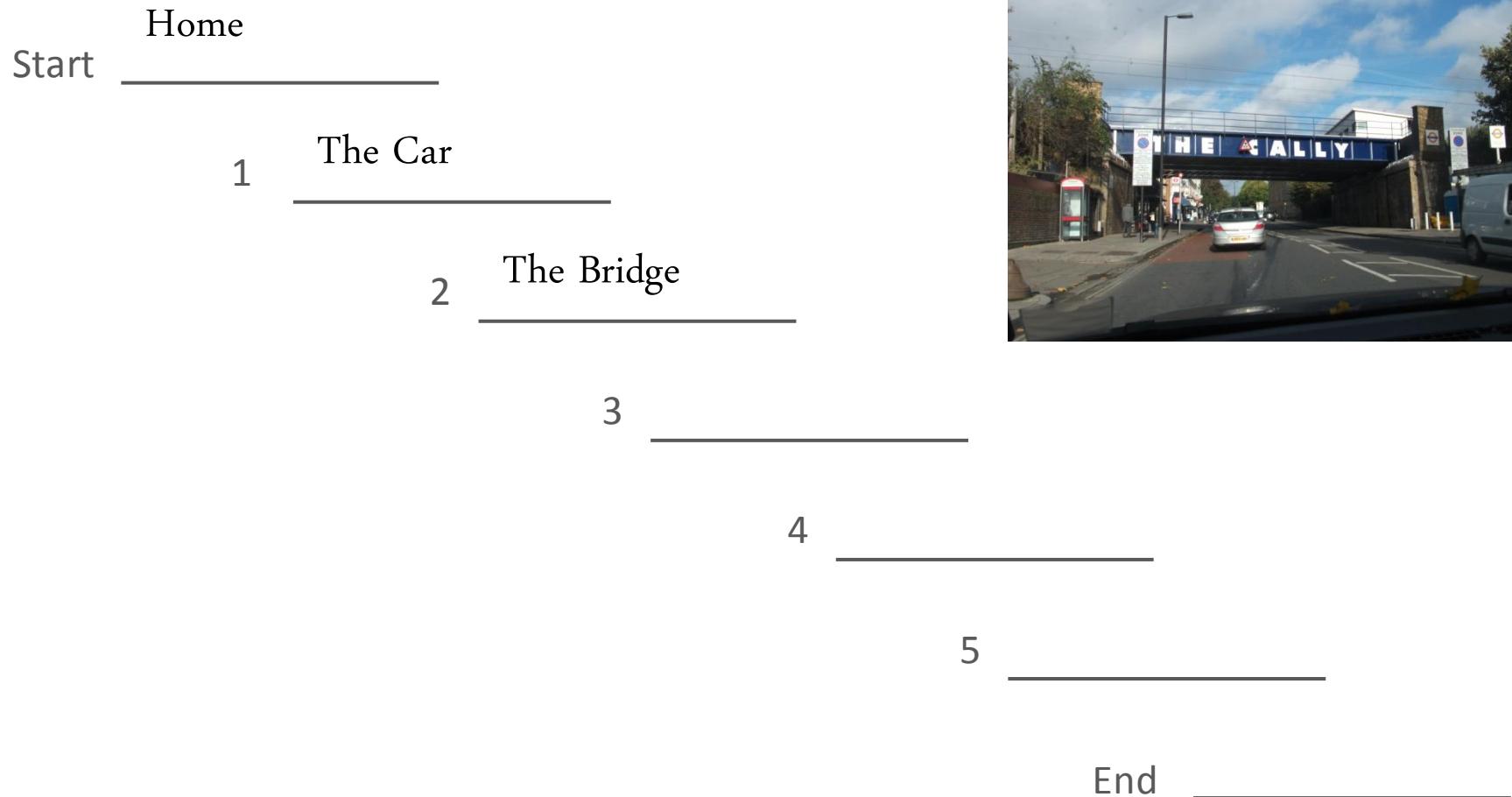
Exercise 1

A journey that you are familiar with ...



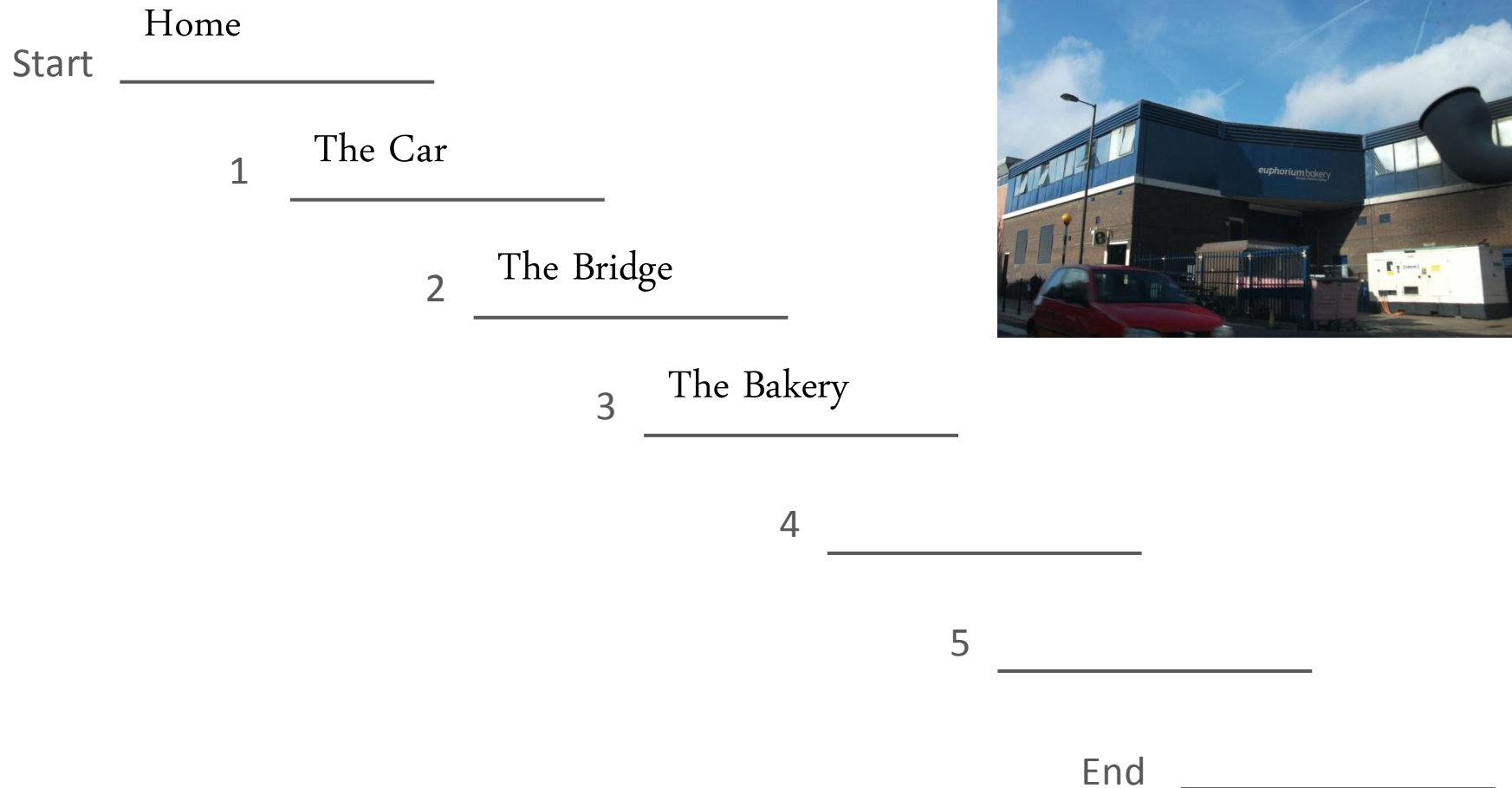
Exercise 1

A journey that you are familiar with ...



Exercise 1

A journey that you are familiar with ...



Exercise 1

A journey that you are familiar with ...

Start Home

1 The Car

2 The Bridge

3 The Bakery

4 The Turning

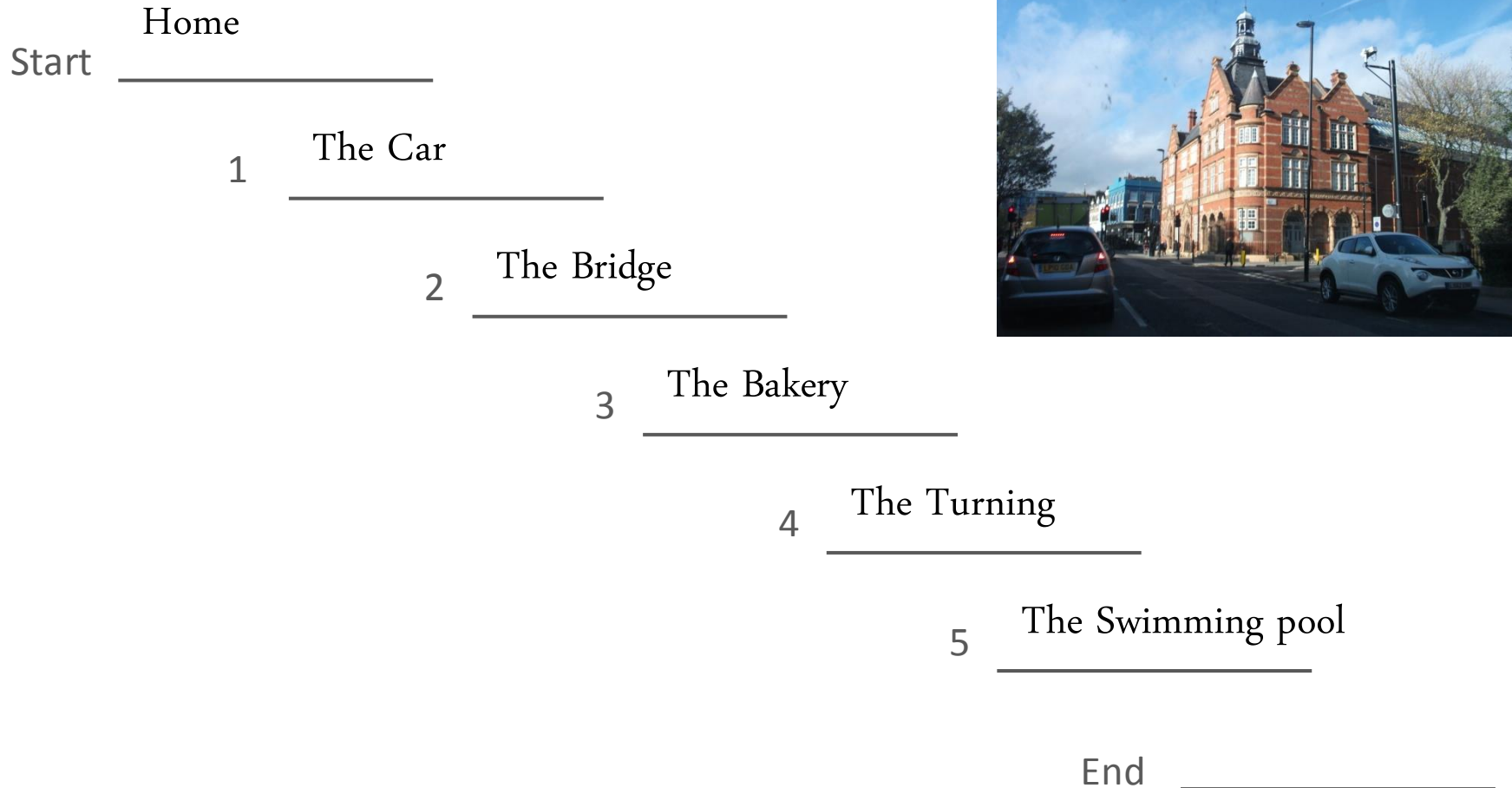
5 _____

End



Exercise 1

A journey that you are familiar with ...



Exercise 1

A journey that you are familiar with ...

Start Home

1 The Car

2 The Bridge

3 The Bakery

4 The Turning

5 The Swimming pool

End The Gymnasium

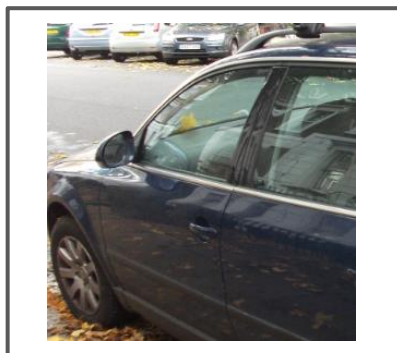


Exercise 2

In relation to that activity ... what are the associated cost categories?



The outfits



The car



Petrol



Parking



Gym Fees



Tea and cake



Competition Fees



Exercise 1

A journey that you are familiar with ...

Start _____

1 _____

2 _____

3 _____

4 _____


5 _____

End _____

What strikes you about
your journey?

Exercise 2

In relation to your activity ... what are the associated cost categories?



Exercise 3

How might you describe the sequencing of a digital asset in your organisation?

Start _____

1 _____

2 _____

3 _____


4 _____

5 _____

End _____

Exercise 4

In relation to your digital asset ... what are the associated cost categories?



PROGRAMME

Section 3:

13:30 – 14:15	45 mins	The CCEx and sharing costs (Presentation and Q&A)
14:15 – 15:00	45 mins	From costs to business models via risk (Presentation and Q&A)
15:00 – 15:15	15 mins	<i>Break</i>

Introduction to the Curation Costs Exchange (CCEx)



The 4C Project

- Our vision is to create a better understanding of digital curation costs through collaboration.
- Our mission is to provide useful, useable resources which support the process of cost management in digital curation.

The CCEx is one of these resources.

What is the CCEx?

- An online, virtual community platform for the exchange of curation cost information.
- Used to gather cost information from partner organisations and stakeholders
- Submission Form/Template to capture calculation processes, metrics, effort statistics, value calculations, from stakeholders in order to underpin future activity with empirical knowledge.

What is the CCEx?

CCEx will provide a means through which interested parties will be able to access information on the costs of curation, in exchange for a little information about the cost of their own digital curation activities.



How will it work?

1. Submission/Input

Submission Template will determine what information the CCEx will collect about our users:

- User type: lurker, registered user, member, partner, administrator
 - Which will enable access to different information
- Stakeholder type: funder, data collector, SME...
 - Based on various use cases
- Cost type: any/activity based/ financial reporting
- How the data is to be shared: all, nothing, anonymous



How does it work?

2. Output

- Cost information based on user profiles and use cases
 - Cost models and user guidance
 - User reviews
 - Related articles, blogs, fora
- Cost data
 - Normalised/aggregated cost data based on profiles and bands of similar size/type organisations and volumes of data
 - Confidential
 - Anonymous



What does CCEx mean to you?

As a ...

Please indicate your professional area (researchers, administrator, librarian, etc.)

I would expect to find ...

What information would you expect to see here?

I'd like to ...

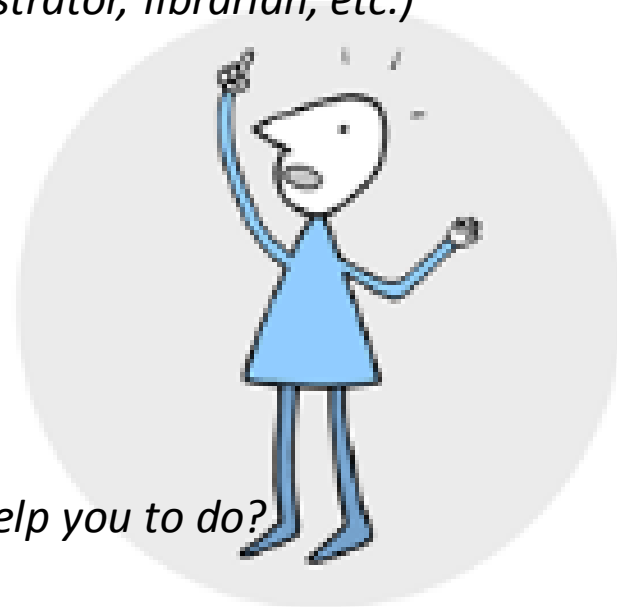
What functionality would you like to see in CCEx?

So I can ...

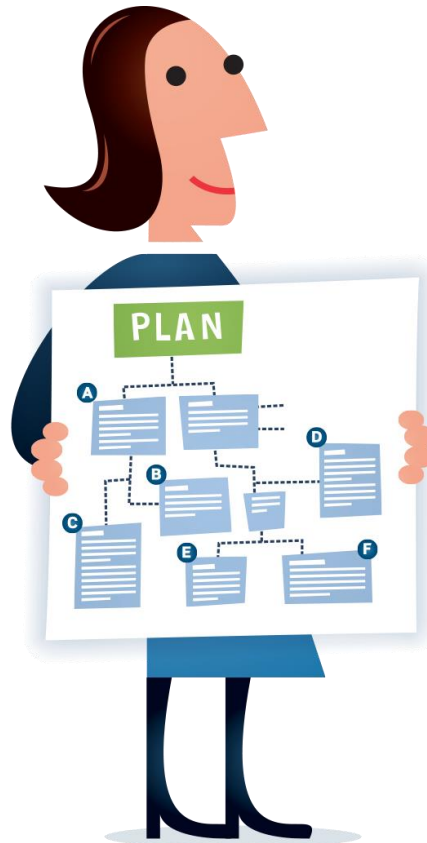
What could the information you'd aim to get out of CCEx help you to do?

I'd be prepared to share...

What data would you be willing to share with others via CCEx? Under what conditions (anonymity?)

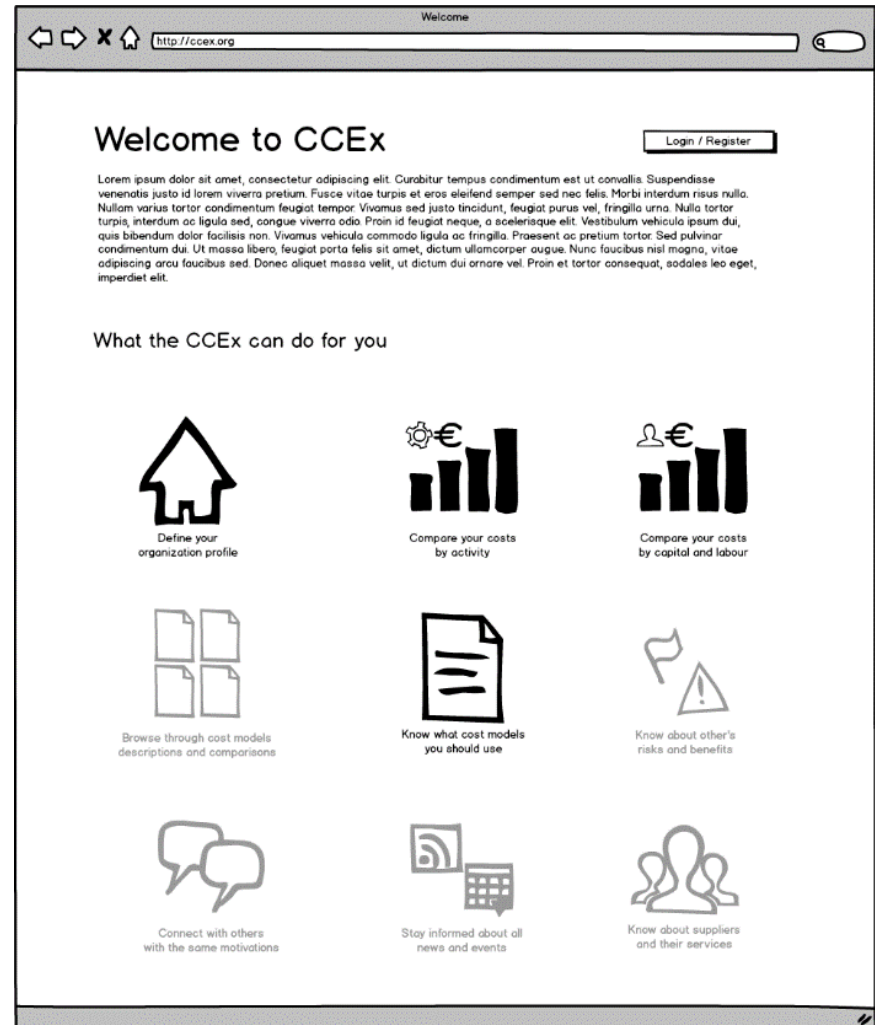


So what does it look like...at the moment?



Landing page options

- Define your organisation
- Compare costs by activity
- Compare costs by capital expenditure and labour
- Browse through cost model descriptions and comparisons
- Know what cost models to use
- Learn about others' risks and benefits
- Connect with others with the same motivations
- Stay informed about news and events
- Learn about suppliers and services



Organisation profile

- Name
- Description
- Cost determinants
 - Data volumes
 - Staff size
 - Duplicate policy
 - Country
 - Core business
 - Organisation type
- Privacy definitions
- Currency

The screenshot shows a web browser window with the address bar displaying 'http://focx.org/organisation'. The page title is 'Define your organization profile'. The form contains the following fields and options:

- Organization name:** A text input field.
- Organization description, purpose and mission:** A large text area.
- Cost determinants:**
 - Data volume in TB:** A text input field.
 - Organization staff size:** A dropdown menu with options: 'Less than 10 people', 'Less than 100 people', 'Less than 500 people', 'Less than 1000 people', 'Less than 10000 people', and 'More than 10000 people'.
 - Number of copies policy:** A dropdown menu with options: 'No replicas', 'One replica', 'Two replicas', 'Three replicas', and 'More than three replicas'.
 - Country (where most of staff lives):** A dropdown menu with options: 'Afghanistan', 'Albania', and 'Algeria'.
 - Core business:** A dropdown menu with options: 'Digital creation' and 'Other'. A 'Help' link is visible next to the dropdown.
 - Organization type:** A dropdown menu with options: 'Research funder' and 'Other'.
 - Other organization type:** A text input field.
- Other settings:**
 - Privacy definitions:** A dropdown menu with options: 'Don't share anything', 'List my organization but don't share costs', 'Share costs in anonymity', 'Share cost information only with other registered partners', and 'Share cost information with everyone'.
 - Currency definitions:** A dropdown menu with options: 'Euro €', 'British Pound Sterling £', 'US Dollar \$', and 'Other'. A 'Help' link is visible next to the dropdown.
 - Other currency:** A text input field.
 - Currency convert rate:** A text input field containing '1', followed by a '€' symbol and a link to 'External currency rate converter'.
- Save details:** A button at the bottom of the form.

Comparison of costs by activity


- Based on common terminology
- E.g. digitization, format migration, ingest, archival storage, access

Compare costs by activity


http://ccex.org/costs/activity/list

Compare costs by activity

Define your curation costs by activity and compare it with other users of the same type.



You must define [your organization profile](#) before being able to compare costs.
You still haven't? [Go there now!](#)



Did you know that you can [compare the costs by capital and labour](#) instead than by activity?
If you have your costs organized that way this will be a better option, but if you already have costs divided by activity then you are in the right place!

My costs by activity





Please add to the list below the costs of all activities that relate to digital curation within your organization. You can create the activities you want, as defined on your currently cost method, and then relate to our defined cost categories. After that you can compare your costs with other institutions that have done the same process.

Add

Group by

Filter by

Import/Export

Activity name	Cost	Normalized activities mapping	
Digitalization	14000 €	Ingest 90%, Archival Storage 10%	 
Format migration	2200 €	Archival Storage 50%, Access 50%	 
Sum 16200 €			

My normalized costs by activity

Below you can find your cost information normalized to a general set of activities that can be compared with others.

Normalized activity	Cost	% of total cost	Cost per TB
Ingest	12600 €	78 %	2520 €/TB
Archival storage	2500 €	15 %	500 €/TB
Access	1100 €	7 %	220 €/TB
Sum 16200 €			Sum 3240 €/TB

You [defined](#) that your organization has around 5 TB of data.

Compare my costs with others

Mapping of costs for comparison

My cost activity

http://ccex.org/costs/activity/new

Cost activity

Name*

Description

Cost* €

Mapping to normalized activities (for comparison with others)

Ingest 10% (1400 €)
Provides functions for Pre-Ingest and Ingest: Appraisal, submission agreement; Functions to accept digital assets SIPs from producers (validation, metadata enrichment), preparation of them for storage and management within the archive (creation of storage versions (AIPs))

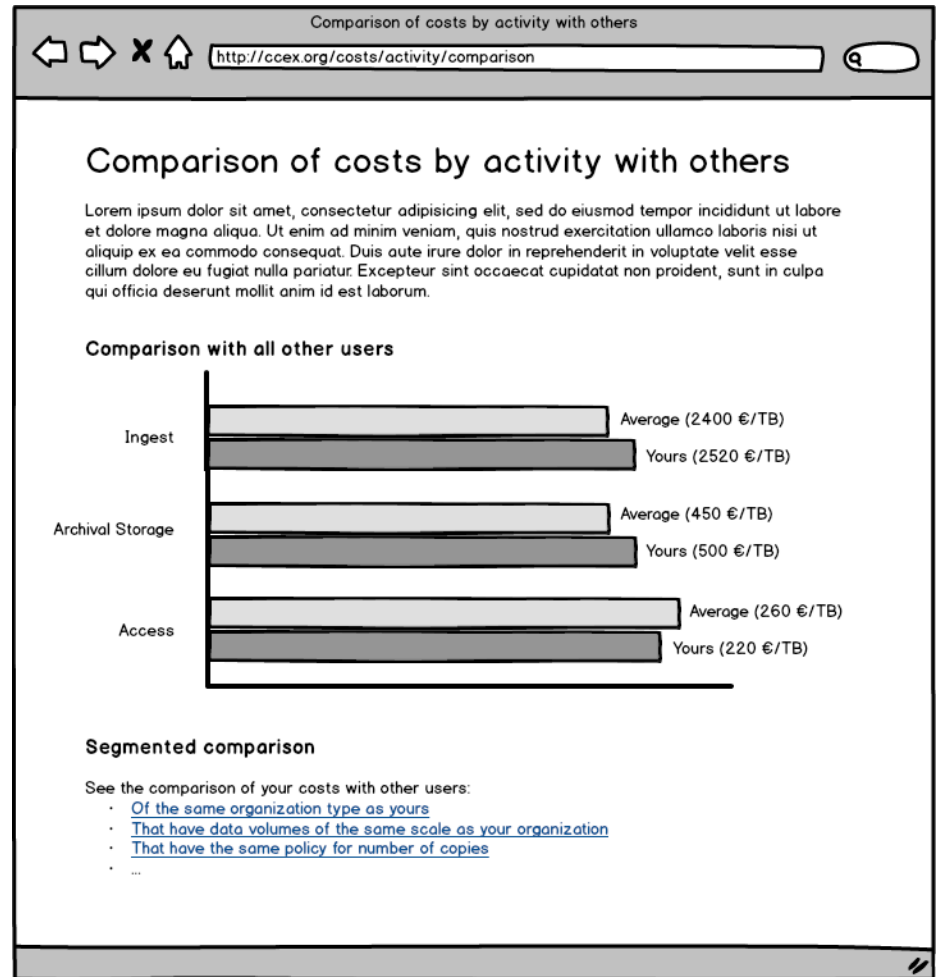
Archival Storage 10% (1400 €)
Functions for storage, maintenance and retrieval of the stored digital assets (AIPs). Includes media migration, storage hierarchy management, routine error checking and providing disaster recovery capabilities.

Access 10% (1400 €)
Functions that support Consumers in determining the existence, description, location and availability of digital assets. Allows consumers to request and receive the digital assets

Unmapped 70% (9800 €)
Beware that unmapped costs will not show on the cost comparison with others.

Save

An at a glance comparison of costs broken down by activity



Comparison of costs broken down by capital expenditure and labour

- E.g. hardware costs, software costs, development and building costs
- E.g. IT developers, operations/support staff, records management staff, other management

Compare costs by capital and labour

http://ccex.org/costs/financial/list

Compare costs by capital and labour

Define your curation costs by capital and labour and compare it with other users of the same type.

You must define your organization profile before being able to compare costs.
You still haven't? [Go there now!](#)

Did you know that you can compare the costs by activity instead than by capital and labour?
If you have your costs organized that way this will be a better option, but if you already have costs divided by capital and labour then you are in the right place!

My costs by capital and labour

Please add to the list below the all capital and labour costs that relate to digital curation within your organization. You can create the cost categories you want, as defined on your currently cost method, and then relate to our defined categories. After that you can compare your costs with other institutions that have done the same process.

[Add](#) [Group by](#) [Filter by](#) [Import/Export](#)

Name	Type	Cost	Cost categories	
Digitalization staff	Labour	10000 €	Support/operations 90%, Manager 10%	<input checked="" type="checkbox"/>
Digitalization machine	Capital	90000 €	Hardware 90%, Software 10%	<input checked="" type="checkbox"/>
		Sum 100000 €		

My normalized costs by financial cost category

Below you can find your cost information normalized to a general set of labour and capital cost categories that can be compared with others.

Labour

Labour category	Cost	% of total cost	Cost per TB
IT-developer	0 €	0 %	0 € / TB
Support/operations	9000 €	9 %	1800 € / TB
Records manager	0 €	0 %	0 € / TB
Manager	1000 €	1 %	200 € / TB
Sum 10000 €		Sum 10 %	Sum 2000 € / TB

Capital

Capital category	Cost	% of total cost	Cost per TB
Hardware	81000 €	91 %	16200 € / TB
Software	9000 €	9 %	1800 € / TB
External services	0 €	0 %	0 € / TB
Building costs	0 €	0 %	0 € / TB
Sum 90000 €		Sum 90 %	Sum 18000 € / TB

[Compare my costs with others](#)

You defined that your organization has around 5 TB of data.

You defined that your organization has around 5 TB of data.

Mapping of costs for comparison

My financial cost

http://ccex.org/costs/financial/new

Financial cost

Name*

Description

Cost* €

Type*

Mapping labour to standard staff roles (for comparison with others)

IT-developer 10% (1400 €)
Staff that develops software. Software engineers, programmers, it-developers, coders.

Support/operations 10% (1400 €)
Staff that executes technical tasks, for example testing digital material at ingest, operating the computers when migrations occur, burning optical disks, setting up robots, etc.

Records manager 10% (1400 €)
Staff that contributes to and executes the preservation planning of the managers, develop the organisation, appraise at pre-ingest, consultants at access, administrative tasks.

Manager 10% (1400 €)
Staff that organise and plan the work of digital curation in the organisations. Make tactical and strategic decisions, have staff responsibility, budget mandate and administrative tasks.

Unmapped 60% (8400 €)
Beware that unmapped costs will not show on the cost comparison with others.

My financial cost

http://ccex.org/costs/financial/new

Financial cost

Name*

Description

Cost* €

Type*

Mapping capital costs with standard acquisition types (for comparison with others)

Hardware 10% (1400 €)
Machines and media used throughout the whole digital asset lifecycle. They receive, store, validate, make copies, migrate and disseminate digital assets.

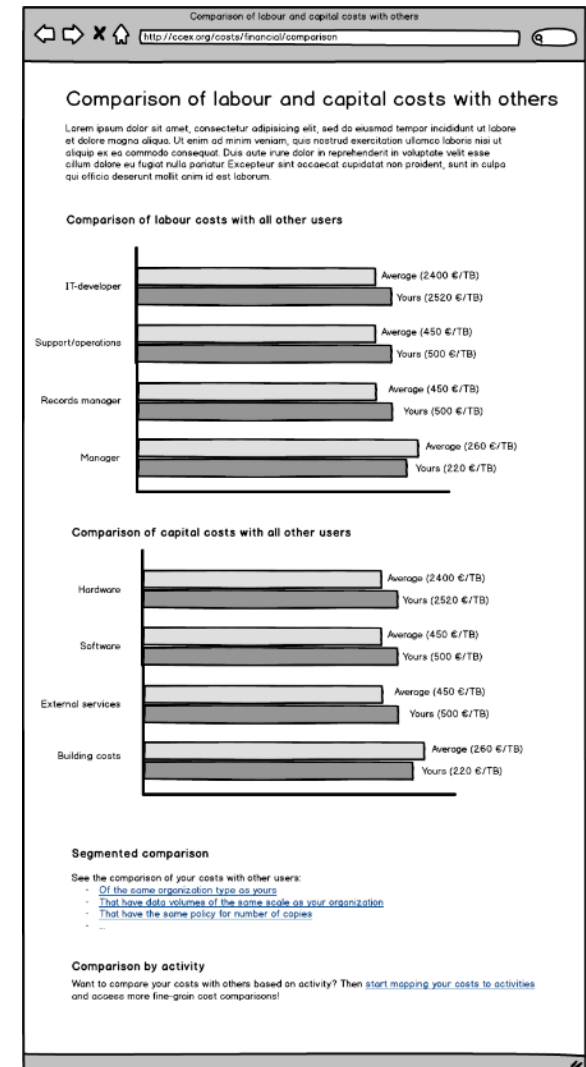
Software 10% (1400 €)
Programmes used throughout the whole digital asset lifecycle. They receive, process, validate, create copies, migrate and disseminate digital assets.

External services 10% (1400 €)
Costs spent to buy services from 3rd party providers.

Building costs 10% (1400 €)
All costs pertaining to building costs, electricity, water, overhead in general.

Unmapped 60% (8400 €)
Beware that unmapped costs will not show on the cost comparison with others.

An at a glance comparison of costs broken down by expenditure and labour



Cost model recommendations

- Accuracy preferences
- Motivation to determine costs
- Assets of interest
- Activity to be costed
- Preferred tool type

Know what cost models you should use

Please fill the following information so we can recommend you information about cost models, literature you might find of interest, and other resources as events, discussion forums, blogs and websites that might interest you.

Cost model accuracy preferences: Less accurate cost model that will finish in half a day, Average accuracy cost model that will finish under a week, Very accurate cost model that will take more than a week

Notes on cost model accuracy preferences

Motivation to determine costs: ☐ Risk, ☐ Efficiency, ☐ Confidentiality, ☐ Trustworthiness, ☐ Value, ☐ Interoperability, ☐ Benefits, ☐ Transparency, ☐ Flexibility, ☐ Sustainability, ☐ Reputation, ☐ Sensitivity, ☒ Other

Other motivations to determine costs

Assets of interest: ☐ Text documents, ☐ Websites, ☐ Geo data, ☐ Email, ☐ e-Journals, ☐ Images, ☐ Spreadsheets, ☐ Newspapers, ☐ Video, ☐ Databases, ☐ Audio, ☐ Digitized documents, ☐ Multidimensional datasets, ☐ Research data, ☐ Other, ☒ Others

Other assets of interest

Activities to be costed: ☐ Pre-ingest, ☐ Preservation Planning, ☐ Production, ☐ Ingest, ☐ Access, ☐ Management, ☐ Data Management, ☐ Administration, ☐ Other, ☐ Archival Storage, ☐ Common services, ☒ Others

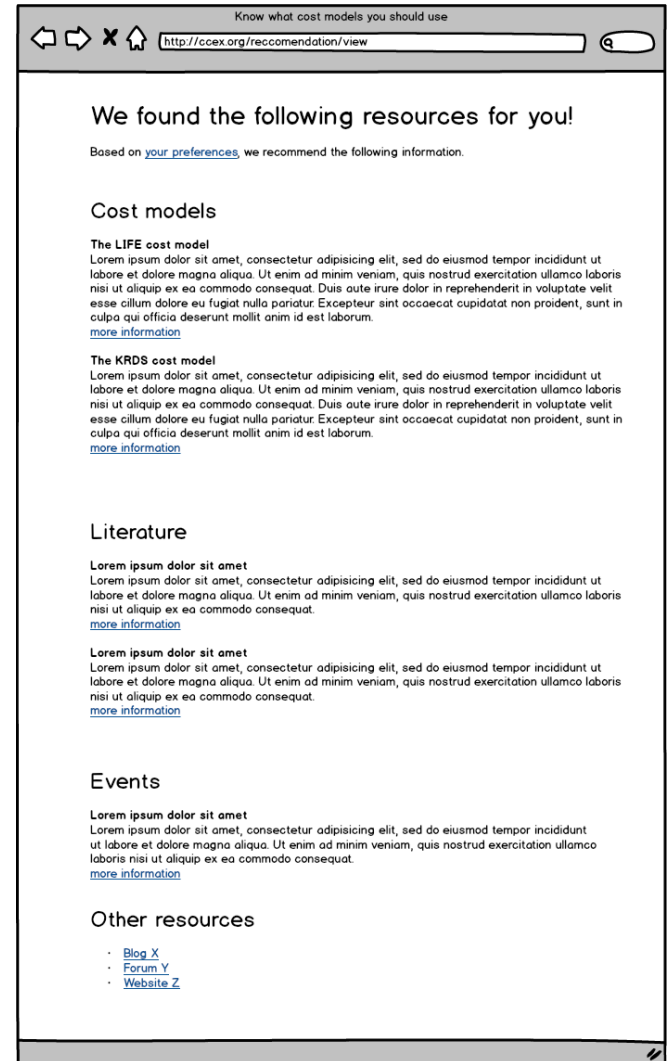
Other activities to be costed

Preferred tool type: Analysis, Estimation, Availability of spreadsheets with default data, Other

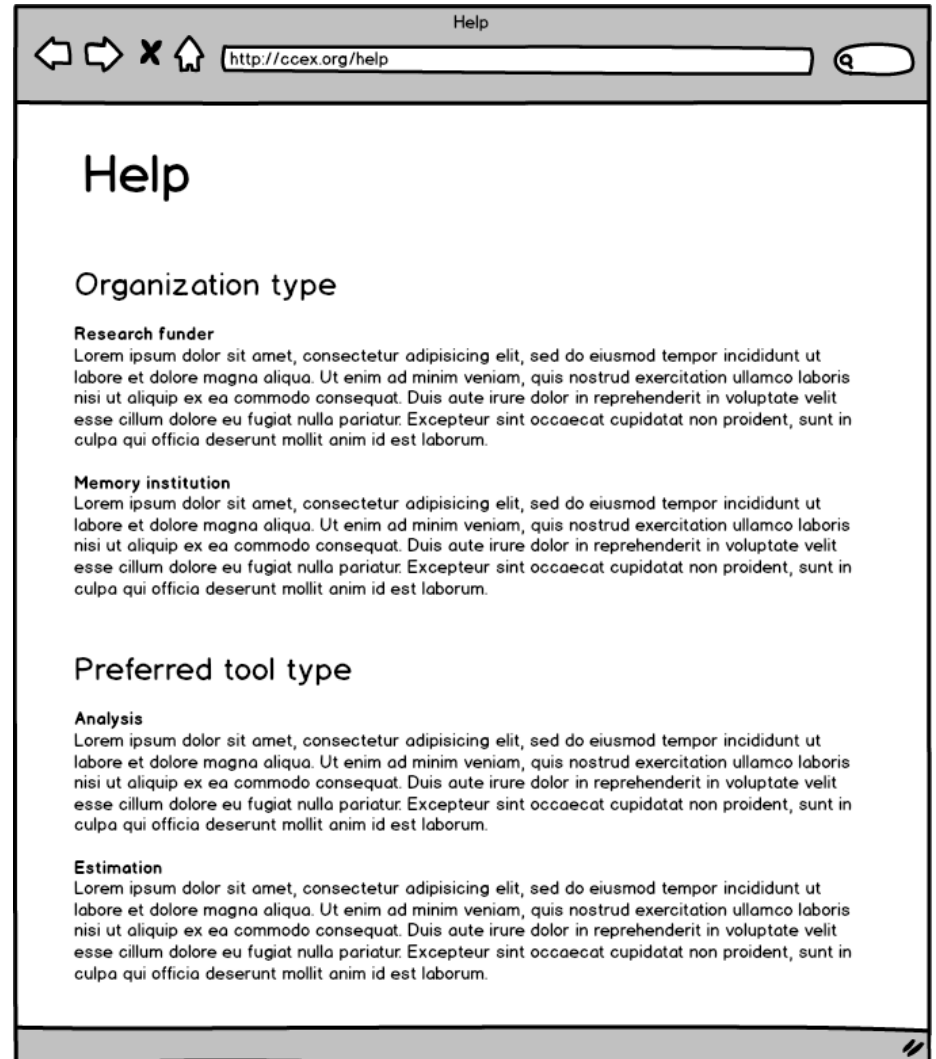
Other preferred tool types

Get recommendations

Cost model recommendations



More help and suggestions



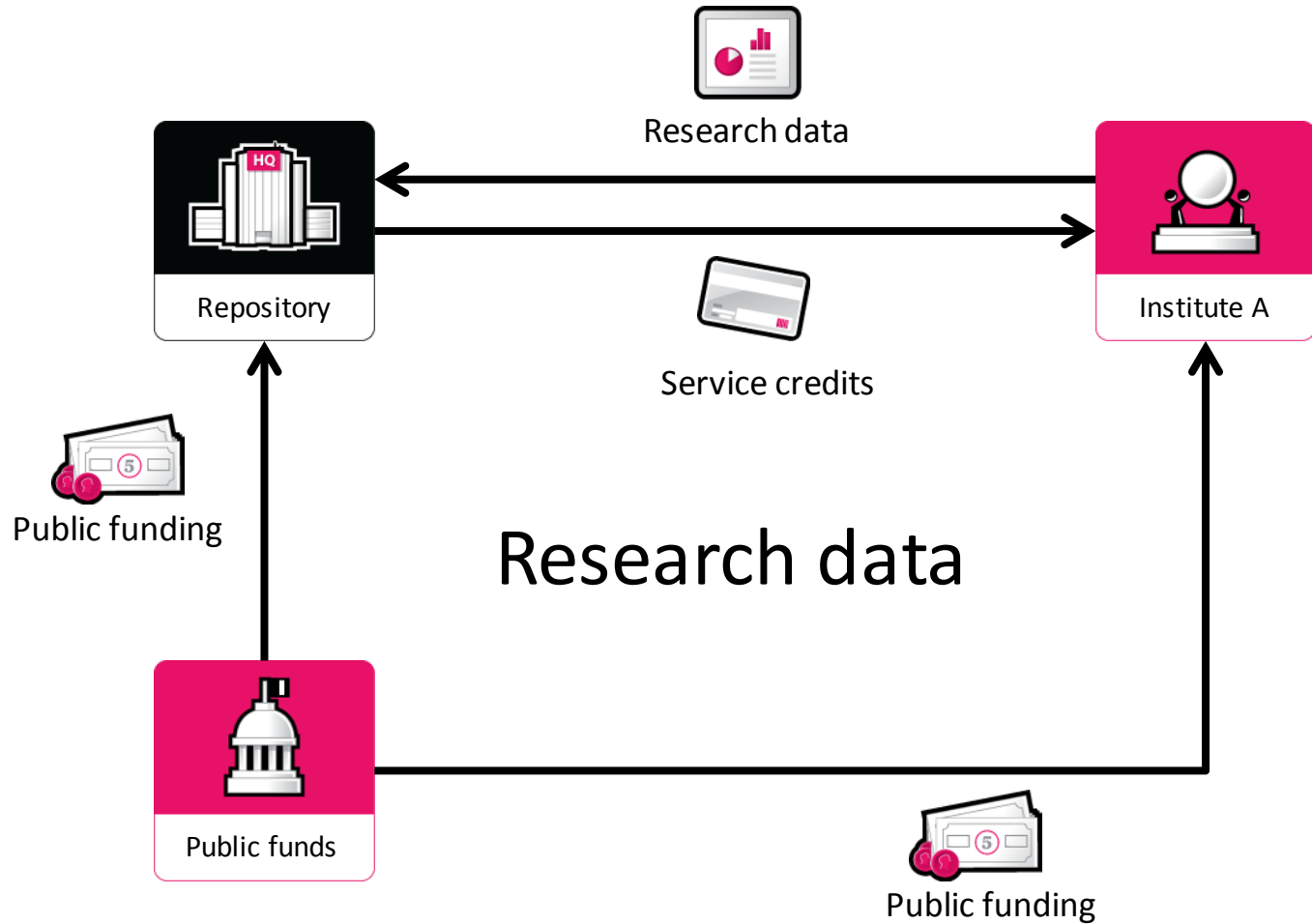
Your thoughts please...



From Costs to Business Models

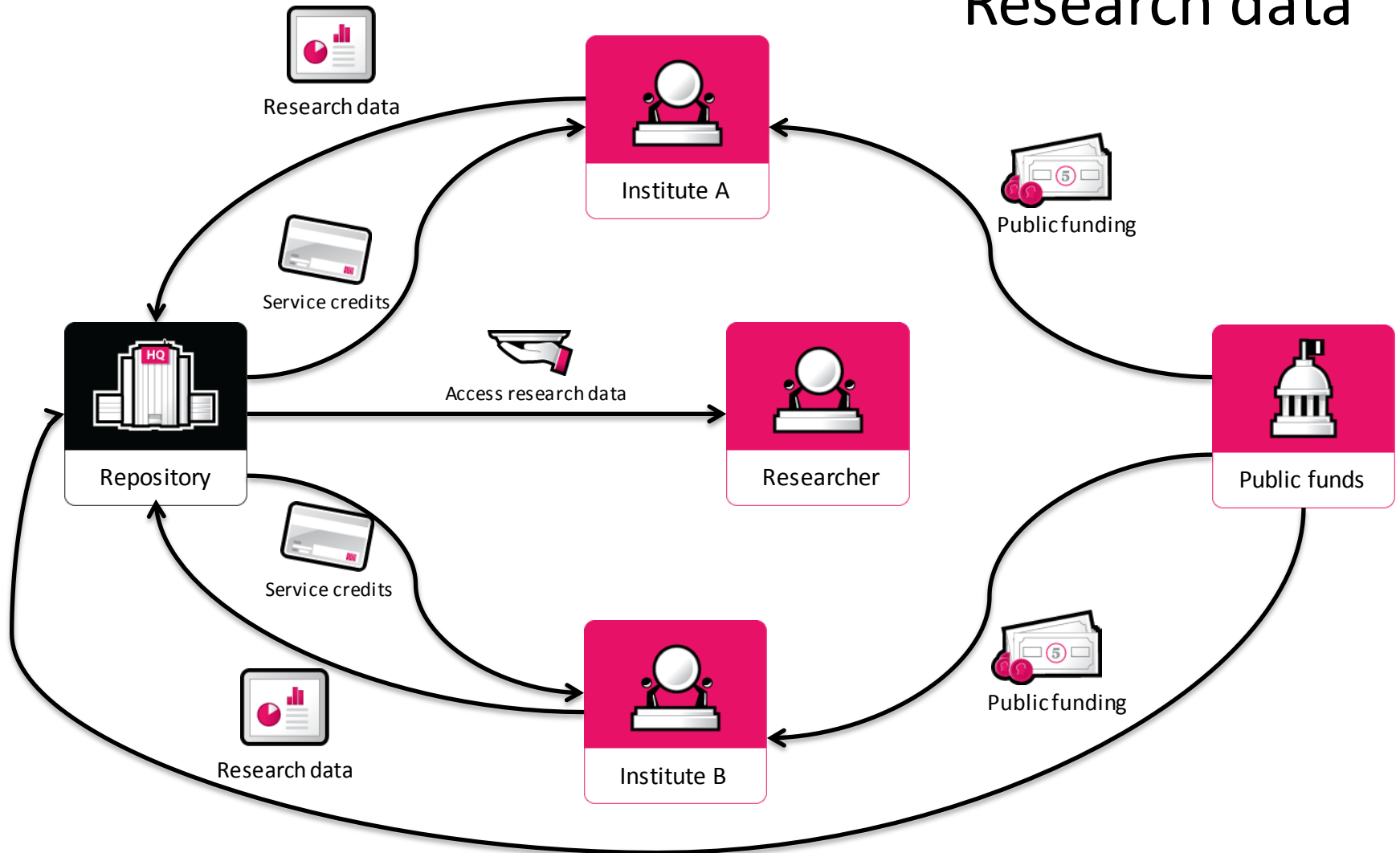
Many institutions would be interested in outsourcing their curation requirements if there was a viable, diverse and competitive market in which solution and service providers were operating. This task will take a look at potential business models and analyse the types of services needed, ways that these can be provided, and options for fee structures. *[4C Description of Work]*

Potential Business Models ...

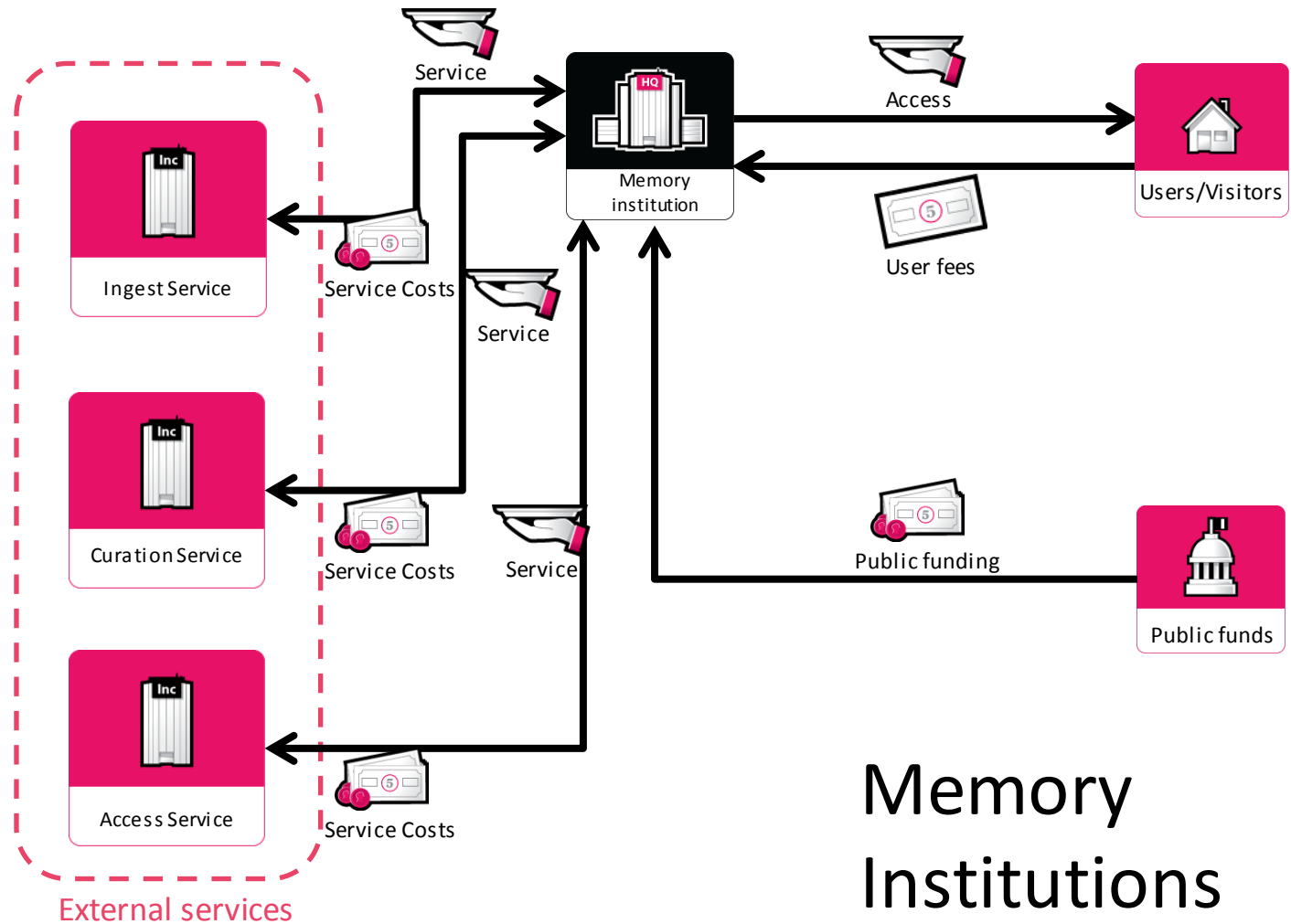


Potential Business Models ...

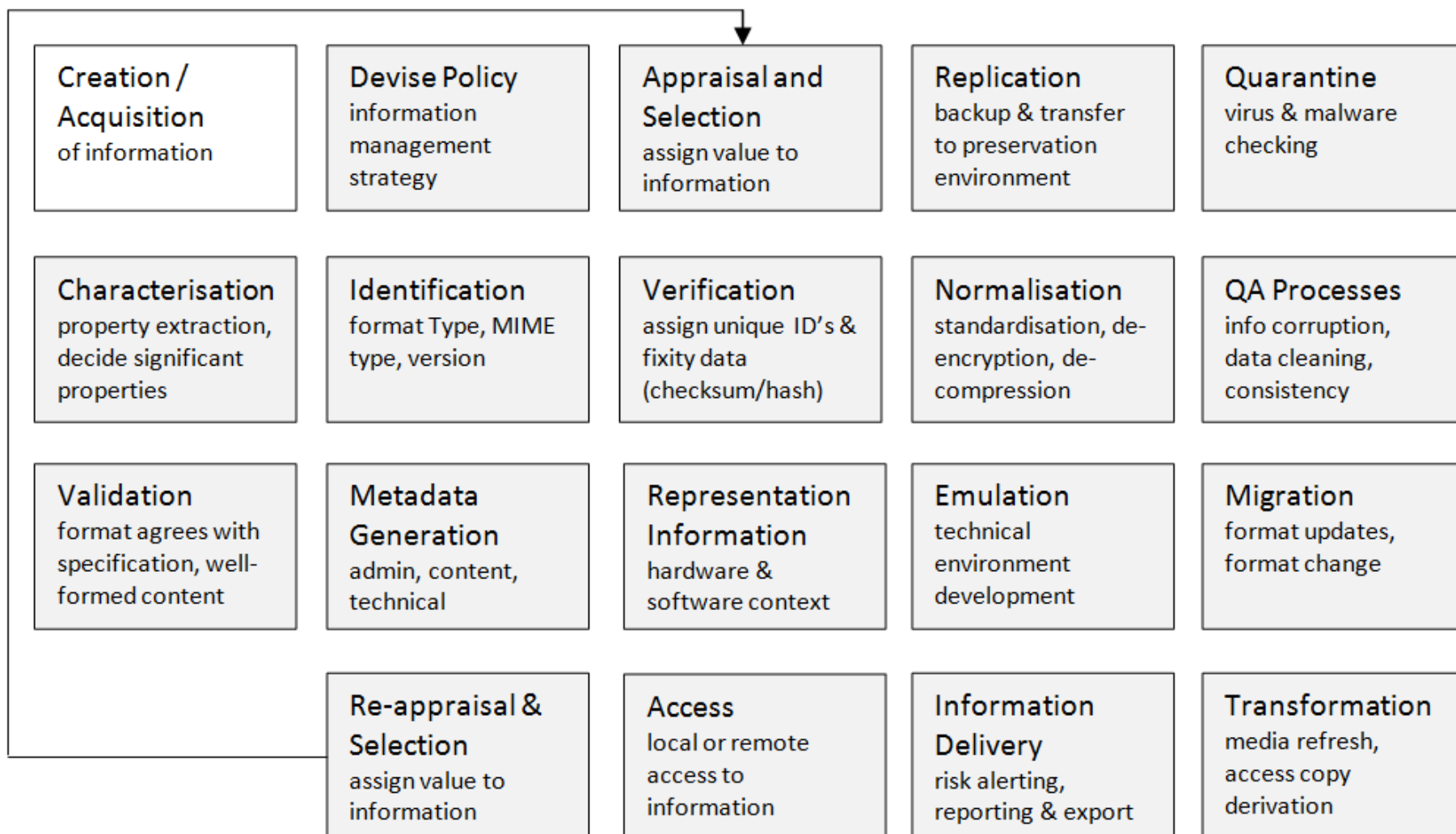
Research data



Potential Business Models ...



Types of Services Needed ...?



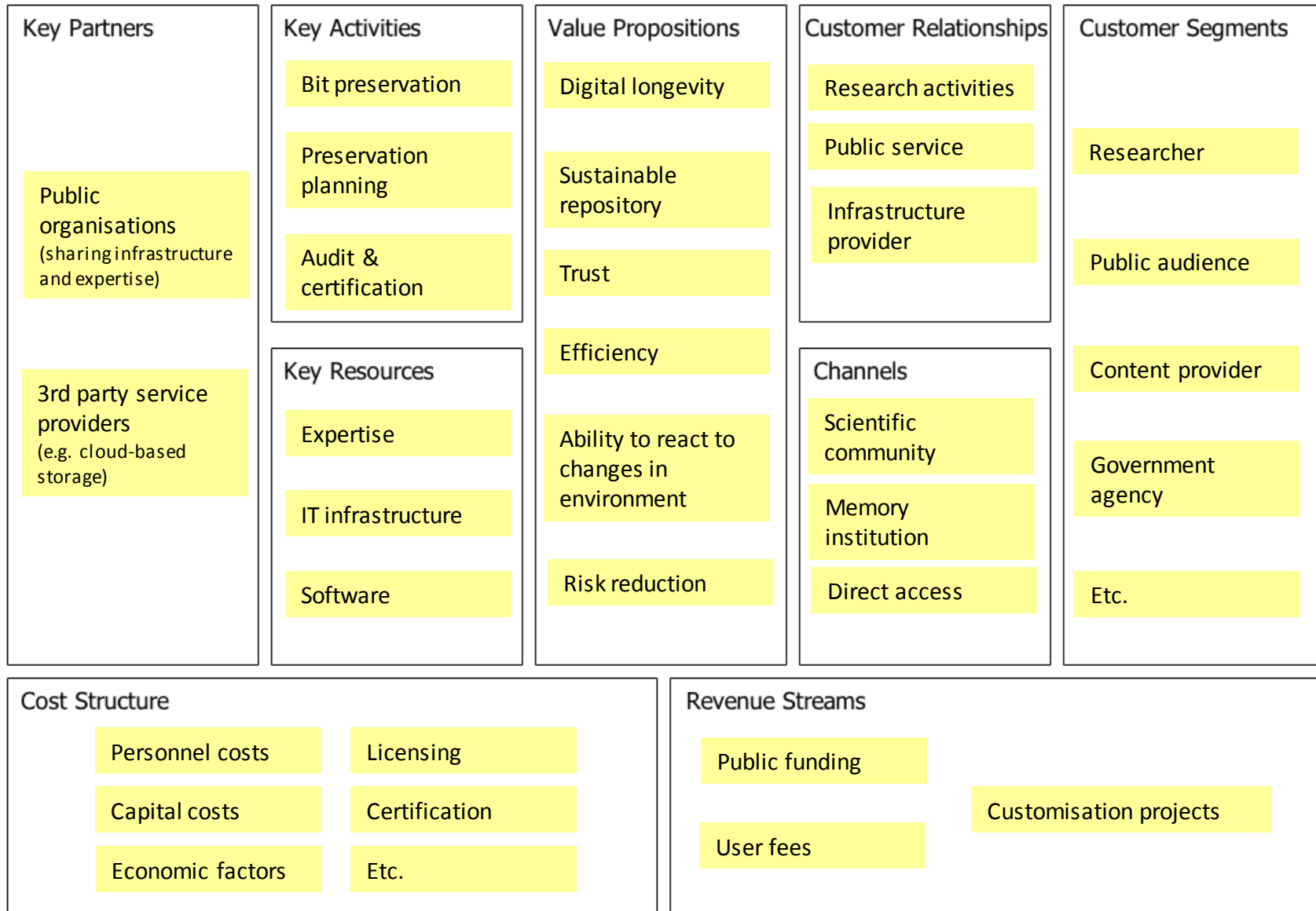
Types of Services Needed ...?

- Ingest
 - Bit level, format, content
 - Metadata extraction and conversion
- Curation
 - Bit storage (tape, disc) (off-line storage)
 - Monitoring and Q&A
 - Preservation Planning (and execution) on demand
- Access
 - Interfaces for access, search (on-line storage)



Ways that services can be provided ...?

The Business Model Canvas



Options for fee structures ...?

This work is in its early stages ...

We'd like to hear ideas about other relevant work and what
would be useful to do in this space ...

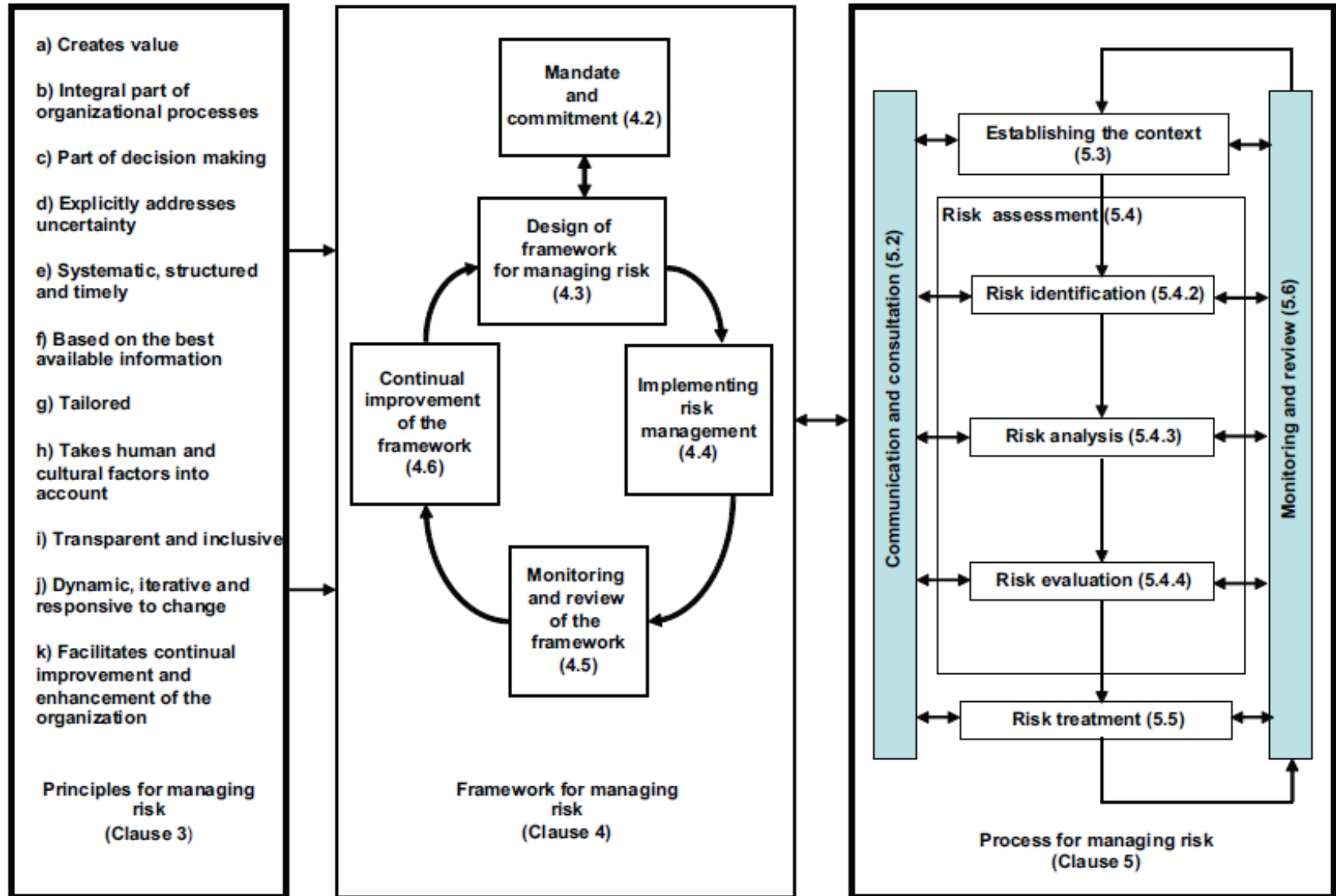
archivematica®

	Free	Basic	Premium
Installation	\$0	\$27,999	\$44,999
free software license (AGPL3)	✓	✓	✓
installation documentation	✓	✓	✓
community forum support	✓	✓	✓
installation technician	✗	on-site plus travel expenses	on-site plus travel expenses
dedicated telephone and email support	✗	3 months plus optional annual maintenance	6 months plus optional annual maintenance
secure remote support	✗	3 months plus optional annual maintenance	6 months plus optional annual maintenance
IT department liaison	✗	✓	✓
storage and network integration	✗	✓	✓
backup procedure review	✗	✓	✓
scalability testing & optimization	✗	✓	✓
administrator and end-user training	✗	✓	✓

On the role of risk, benefit, impact and value as
an economic determinant in digital curation...

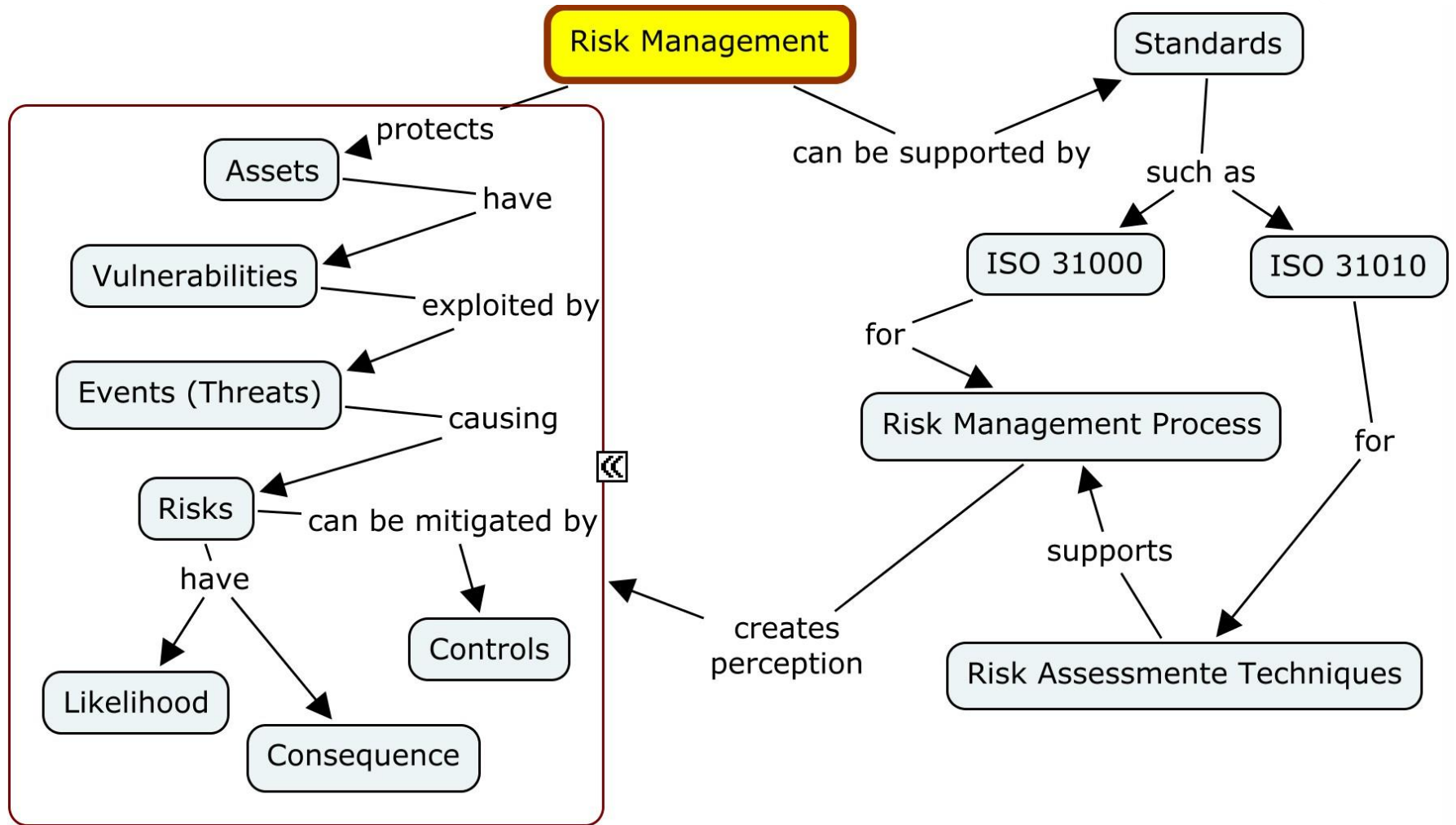
ISO/FDIS 31000:2009(E) - Risk management — Principles and guidelines

Figure 1 — Relationships between the risk management principles, framework and process



Some relevant concepts with concise definitions (ISO73 and ISO31000 provide more formal definitions):

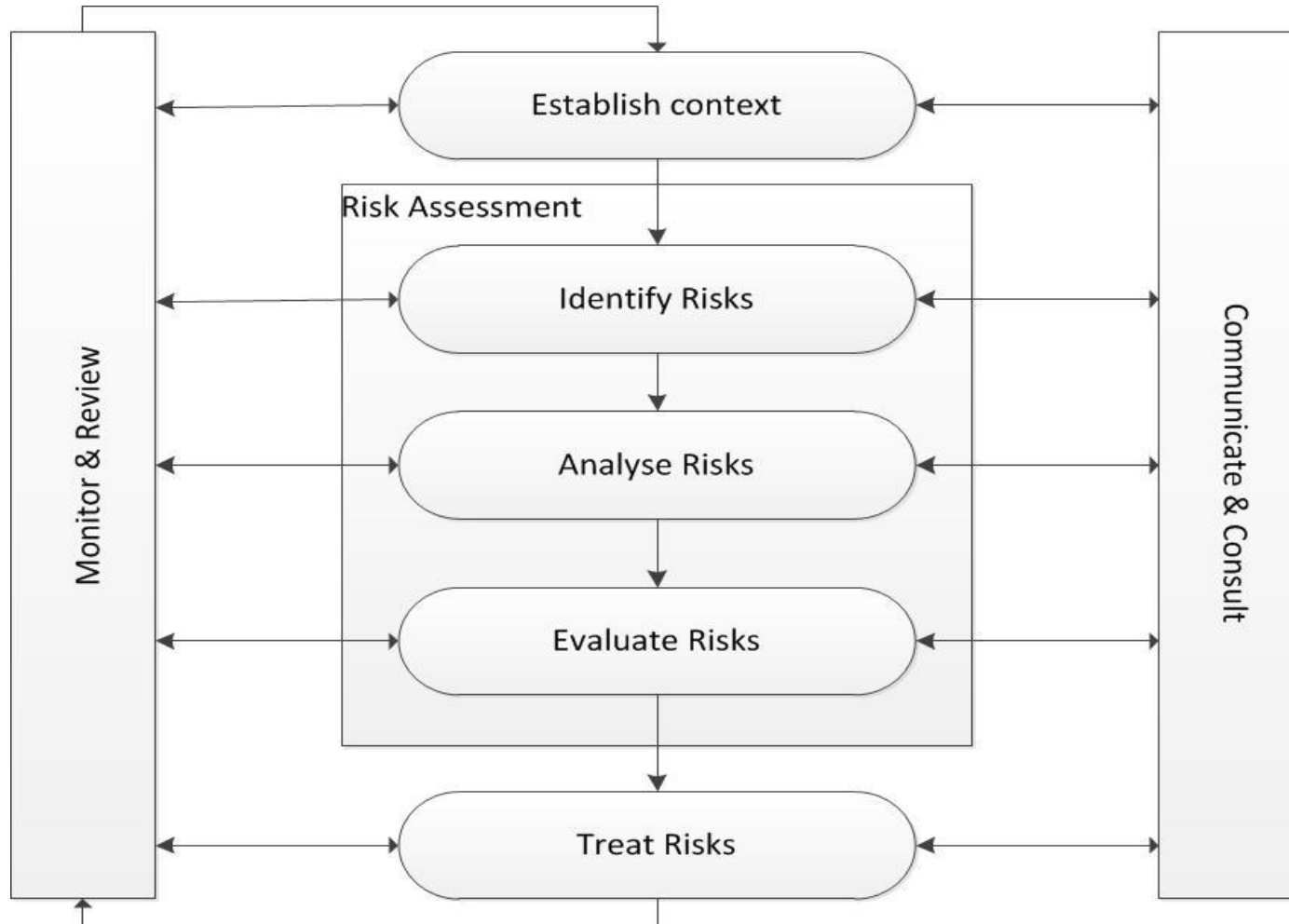
- **Event - occurrence (or not...) of a particular set of circumstances**
 - The event can be certain or **uncertain** (the probability can be estimated for a given period of time...)
- **Risk** – an objective exposed to **uncertainty**
- **Uncertainty** - lack of sufficient relevant knowledge for a particular objective
- **Vulnerability** – an **uncertainty** that could have a negative effect for a particular objective
- **Threat** – the exploitation of a vulnerability
- **Loss** – a negative effect on an objective from a threat
- **Opportunity** - an uncertainty that could have a positive effect for a particular objective
- **Gain** – a positive effect on an objective from the exploitation of an opportunity
- **Control** – measures that is modifying a risk
- **Risk Management** - The systematic process of identifying and analyzing risks and responding by **defining controls**
- **Risk register** - record of information about identified risks



Accordingly, we can propose that:

- **Control** – is a measure that we can put in practice to **minimize loss** (the main concern of digital preservation) or also to **maximize gain** (the maximum concern of digital curation)
- **Costs** – is what we have to give up for the **controls** !!!

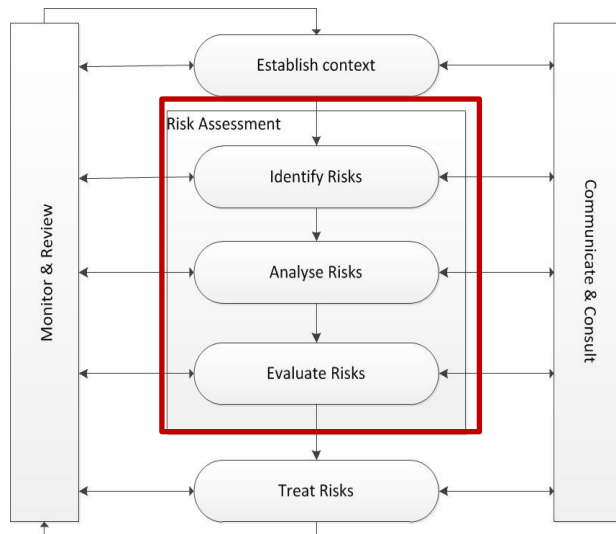
The generic Risk Management process (ISO31000)




IEC/FDIS 31010:2009(E) - Risk assessment techniques

Table A.1 – Applicability of tools used for risk assessment

- 1) SA - Strongly applicable.
- 2) NA - Applicable.
- 3) A - Applicable.



Tools and techniques	Risk assessment process					
	Risk Identification	Risk analysis				
		Consequence	Probability	Level of risk		
Brainstorming	SA ¹⁾	NA ²⁾	NA	NA		
Structured or semi-structured Interviews	SA	NA	NA	NA	NA	B 02
Delphi	SA	NA	NA	NA	NA	B 03
Check-lists	SA	NA	NA	NA	NA	B 04
Primary hazard analysis	SA	NA	NA	NA	NA	B 05
Hazard and operability studies (HAZOP)	SA	SA	A ³⁾	A	A	B 06
Hazard Analysis and Critical Control Points (HACCP)	SA	SA	NA	NA	SA	B 07
Environmental risk assessment	SA	SA	SA	SA	SA	B 08
Structure « What if? » (SWIFT)	SA	SA	SA	SA	SA	B 09
Scenario analysis	SA	SA	A	A	A	B 10
Business Impact analysis	A	SA	A	A	A	B 11
Root cause analysis	NA	SA	SA	SA	SA	B 12
Failure mode effect analysis	SA	SA	SA	SA	SA	B 13
Fault tree analysis	A	NA	SA	A	A	B 14
Event tree analysis	A	SA	A	A	NA	B 15
Cause and consequence analysis	A	SA	SA	A	A	B 16
Cause-and-effect analysis	SA	SA	NA	NA	NA	B 17
Layer protection analysis (LOPA)	A	SA	A	A	NA	B 18
Decision tree	NA	SA	SA	A	A	B 19
Human reliability analysis	SA	SA	SA	SA	A	B 20
Bow tie analysis	NA	A	SA	SA	A	B 21
Reliability centred maintenance	SA	SA	SA	SA	SA	B 22
Sneak circuit analysis	A	NA	NA	NA	NA	B 23
Markov analysis	A	SA	NA	NA	NA	B 24
Monte Carlo simulation	NA	NA	NA	NA	SA	B 25
Bayesian statistics and Bayes Nets	NA	SA	NA	NA	SA	B 26
FN curves	A	SA	SA	A	SA	B 27
Risk Indices	A	SA	SA	A	SA	B 28
Consequence/probability matrix	SA	SA	SA	SA	A	B 29
Cost/benefit analysis	A	SA	A	A	A	B 30
Multi-criteria decision analysis (MCDA)	A	SA	A	SA	A	B 31

IEC/FDIS 31010:2009(E) - Risk management — Risk assessment techniques

Likelihood rating	E	IV	III	II	I	I	I
	D	IV	III	III	II	I	I
	C	V	IV	III	II	II	I
	B	V	IV	III	III	II	I
	A	V	V	IV	III	II	II
		1	2	3	4	5	6
		Consequence rating					

Figure B.15 – Part example of a probability criteria matrix

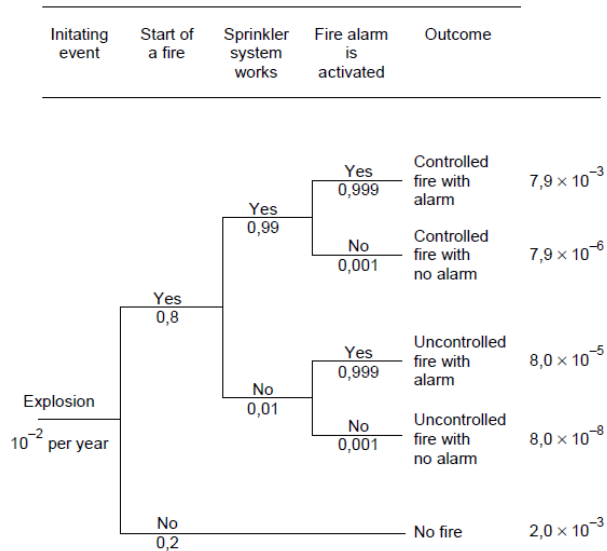


Figure B.3 – Example of an event tree

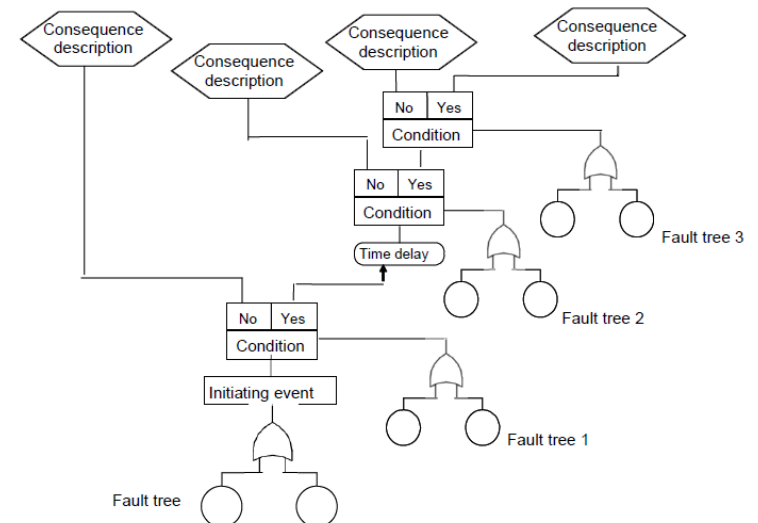
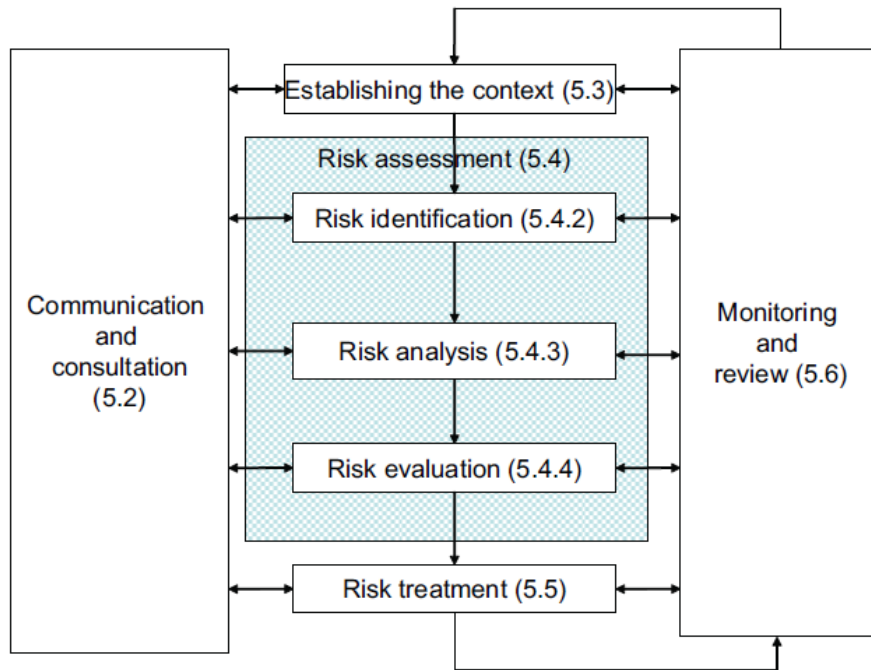


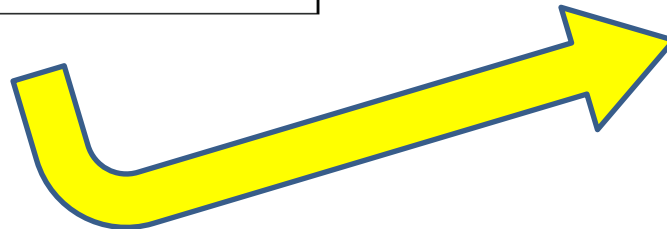
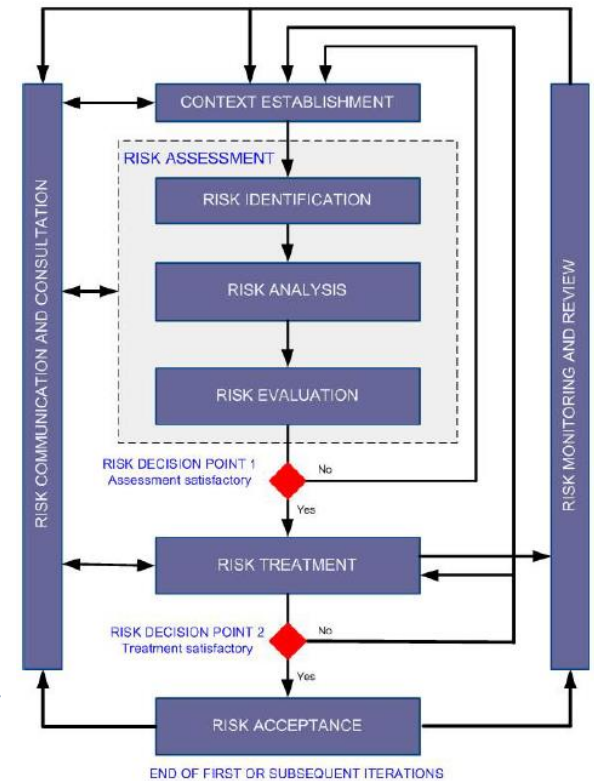
Figure B.4 – Example of cause-consequence analysis

ISO/FDIS 31000:2009(E) - Risk management —
Principles and guidelines

Figure 3 — Risk management process



ISO/IEC 27005:2011(E) - Information technology —
Security techniques — Information security risk
management



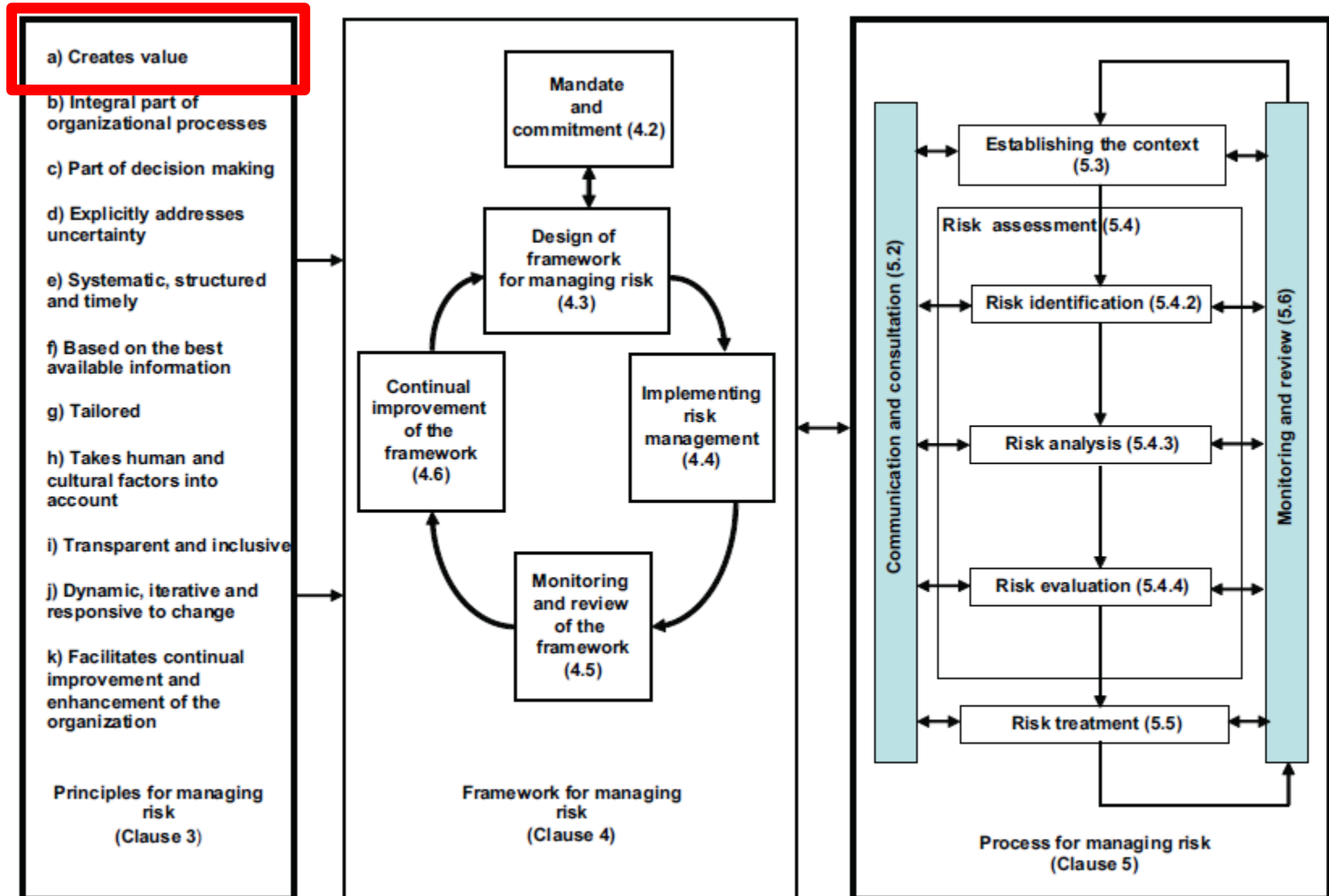
ISO/IEC 27005:2011(E) -
Information technology —
Security techniques —
Information security risk
management

Examples of vulnerabilities
(2 pages of them in the
standard...)

Types	Examples of vulnerabilities	Examples of threats
Hardware	Insufficient maintenance/faulty installation of storage media	Breach of information system maintainability
	Lack of periodic replacement schemes	Destruction of equipment or media
	Susceptibility to humidity, dust, soiling	Dust, corrosion, freezing
	Sensitivity to electromagnetic radiation	Electromagnetic radiation
	Lack of efficient configuration change control	Error in use
	Susceptibility to voltage variations	Loss of power supply
	Susceptibility to temperature variations	Meteorological phenomenon
	Unprotected storage	Theft of media or documents
	Lack of care at disposal	Theft of media or documents
	Uncontrolled copying	Theft of media or documents
Software	No or insufficient software testing	Abuse of rights
	Well-known flaws in the software	Abuse of rights
	No 'logout' when leaving the workstation	Abuse of rights
	Disposal or reuse of storage media without proper erasure	Abuse of rights
	Lack of audit trail	Abuse of rights
	Wrong allocation of access rights	Abuse of rights
	Widely-distributed software	Corruption of data
	Applying application programs to the wrong data in terms of time	Corruption of data
	Complicated user interface	Error in use
	Lack of documentation	Error in use
	Incorrect parameter set up	Error in use
	Incorrect dates	Error in use

ISO/FDIS 31000:2009(E) - Risk management — Principles and guidelines

Figure 1 — Relationships between the risk management principles, framework and process



On “Value”:

Considering:

- **Costs** – is what we have to give up for the **controls**
- **Control** – is a measure that we can put in practice to **minimize loss** (the main concern of digital preservation) or also to **maximize gain** (the maximum concern of digital curation)

We can propose these concepts:

- **Relative value** – ...ranking or weight of a set of assets for what we do not have their absolute value (this can be often the case of assets in the cultural heritage sector).
- **Added value** – ... the value that results from a control that can be applied to explore an opportunity. If that value can be measured, than we can claim we can quantify an added value for the related asset as a result of that action (this is **curation!!!**).

- **Step 1 – Definition of the Method (...)**
 - **1.1 – Principles and Guidelines:** Based on ISO 31000 ..
 - **1.2 – Assessment techniques:** Based on ISO 31010, on previous work (such as DRAMBORA) ...
 - **1.3 - Risk Register:** Propose the structure of a risk register for the domain, and demonstrate it populated with the results of existing sources.
- **Step 2 – Validation of the Method** (Iterate to Step 1 if necessary)
 - **2.1 – Internal Conceptual Review (...)**
 - **2.2 – Internal Validation (...end March...)**
 - **2.3 – External Closed Validation:** Ask external stakeholders as experts (...end April...)
 - **2.4 – External Open Validation:** Disseminate the method and ask for volunteers for external validation (...end May...)
- **Step 3 – Revise and publish (...end June!!!)**



Collaboration to Clarify the Costs of Curation



Discussion...?

PROGRAMME

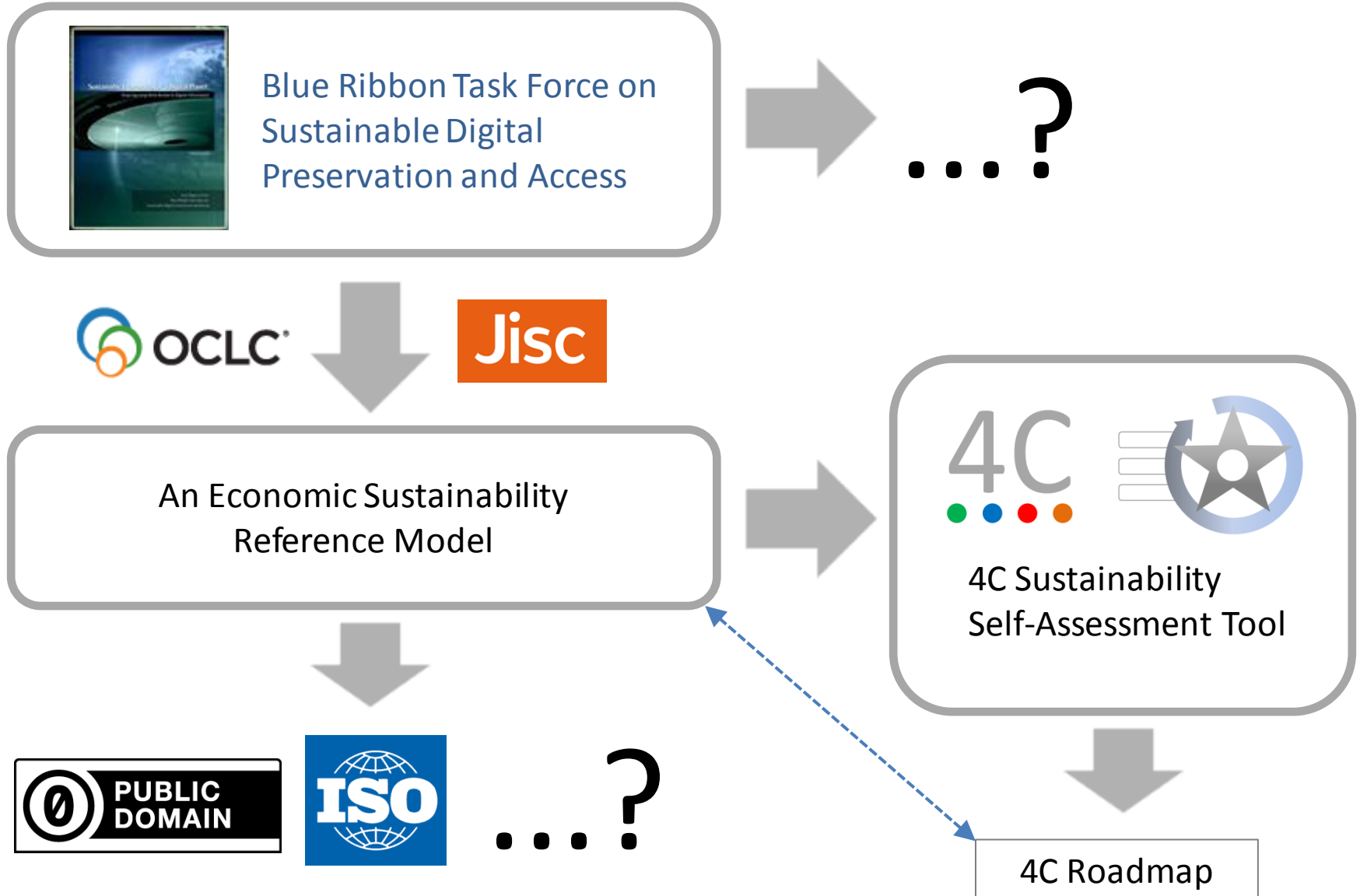
Section 4:

15:15 – 16:15	1 hour	Sustaining solutions and services using the ESRM (Exercise)
16:15 – 16:30	15 mins	Recap, summing up and feedback
16.30		Finish

Introduction to the Economic Sustainability Reference Model (ESRM)

The ESRM maps out the key elements of the problem space planners face when designing a sustainability strategy for their digital curation activities.

It focuses on the general concept of a sustainability strategy, breaks it down into its key components, and draws planners' attention to the properties of those components most relevant for economic sustainability.



Some questions to consider as we go through the session ...

Does it serve a useful purpose?

Who would use it?

Why would they use it?

When would they use it?

How would they use it?

What's missing?

How could it be more effective?



Operational Staff



Operational Managers



Senior Managers



Funders & Investors

Some points to bear in mind as we assess the ESRM + Tool ...

They are works in progress ...

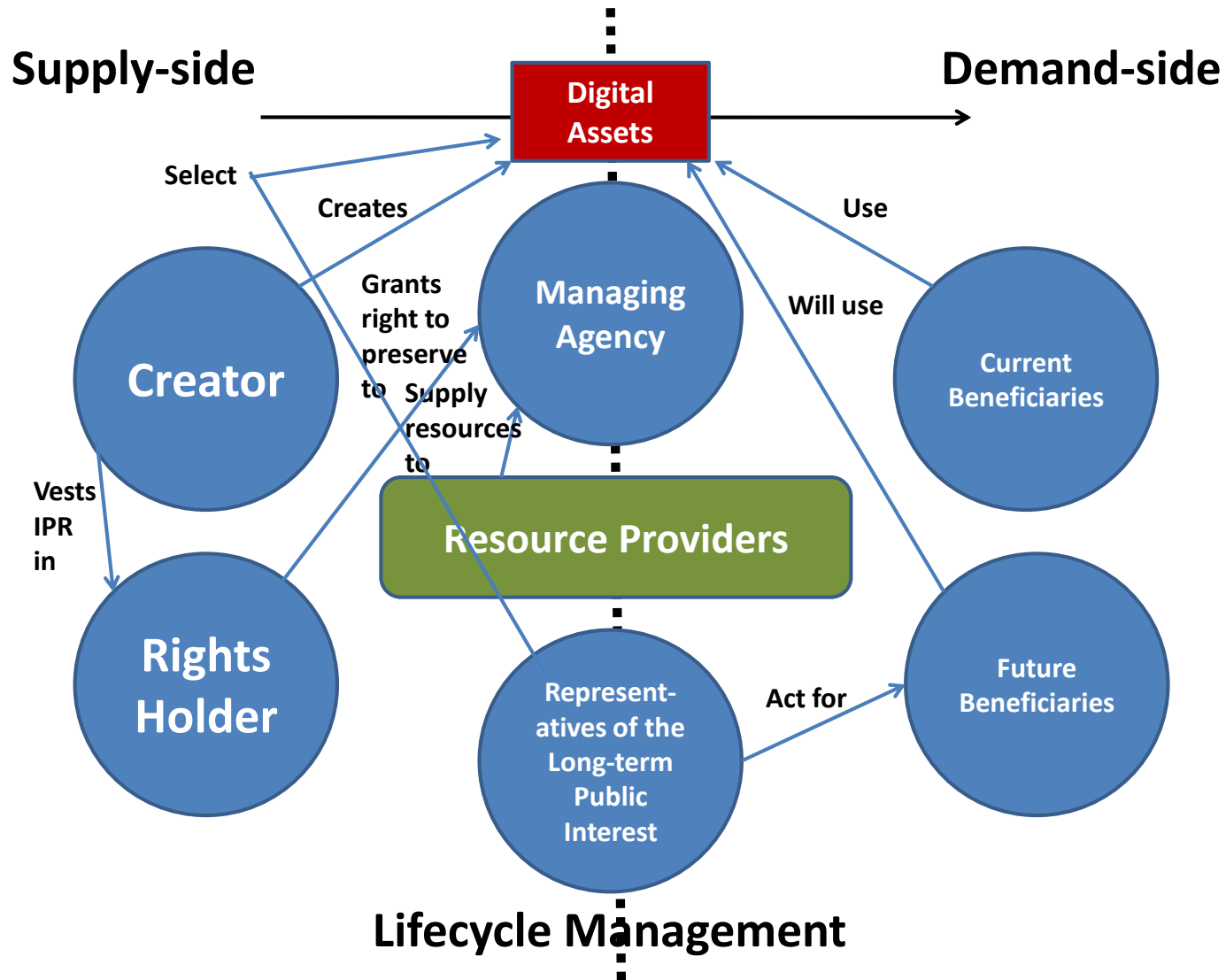
They need quite a lot more work done on (both) of them to streamline them and make them efficient and usable

Both resources are an attempt to generalise and/or simplify.

It's tempting (and possible) to jump straight away to the Self-Assessment Tool and try to just fill it in with no reference to the ESRM text. But the concepts in the tool are taken from the text and the text gives more context and meaning to the questions in the tool.

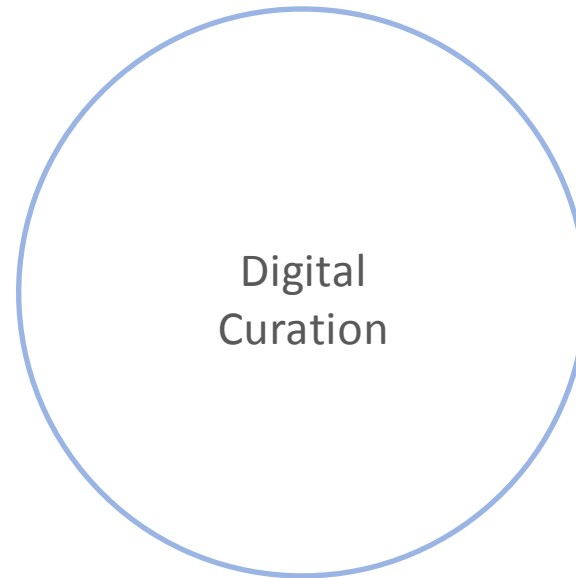
For instance ...

Q6 – Has the stakeholder ecosystem been surveyed or mapped?

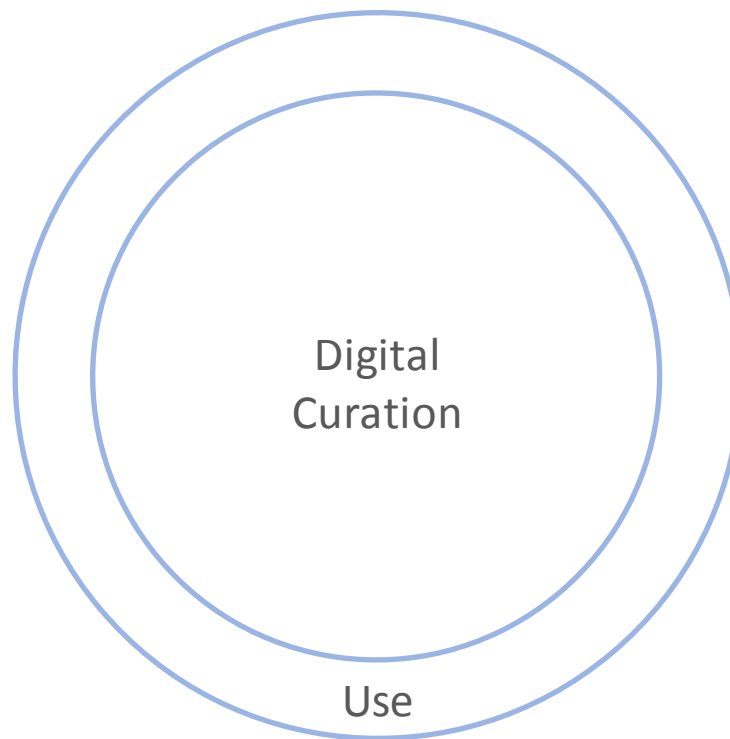


The ESRM proposes that a sustainability strategy requires consideration of four categories of issues:

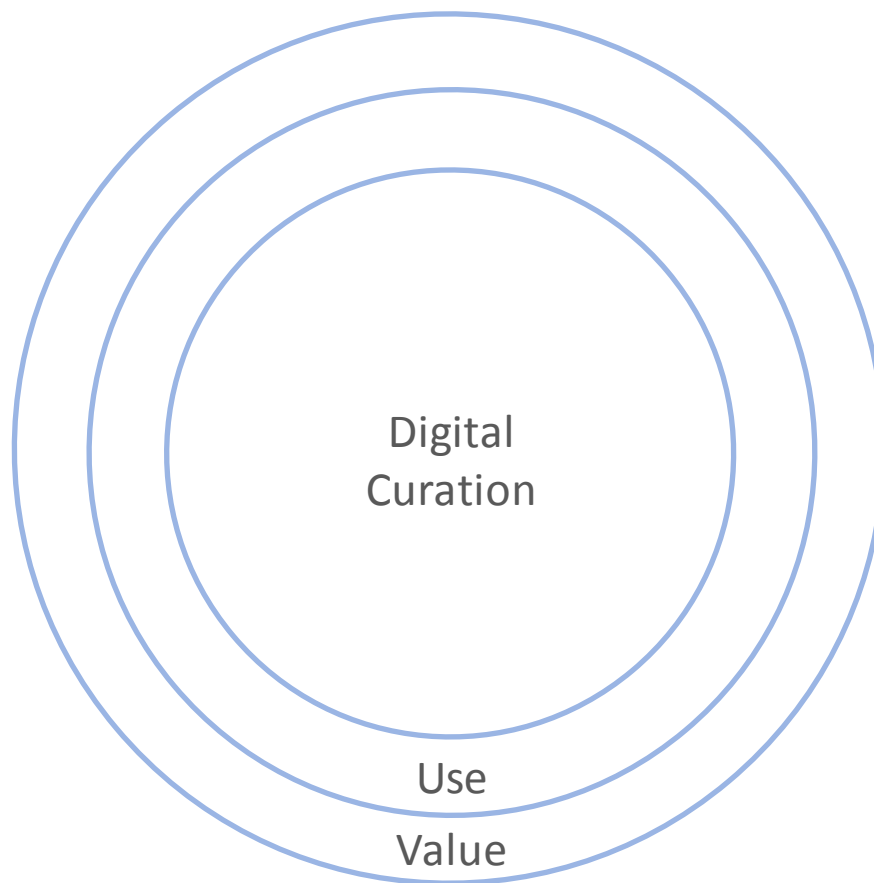
- The Economic Lifecycle
- Sustainability Conditions
- Key Entities
- Economic Uncertainties



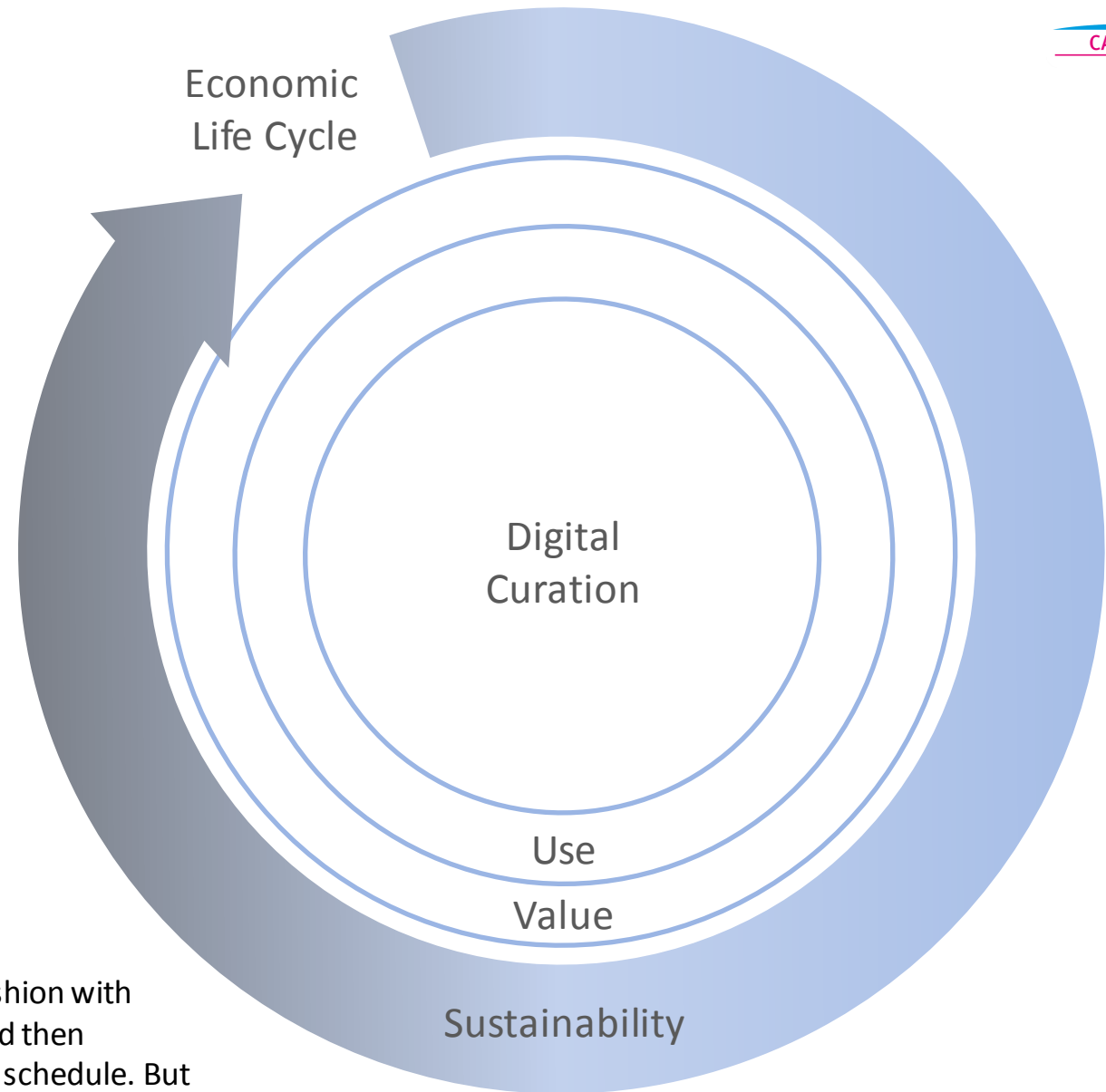
The activity of digital curation is assumed to be the central active component and the engine that will ensure the sustainability of digital assets



Investment into curation will in turn facilitate use (or the potential for use)

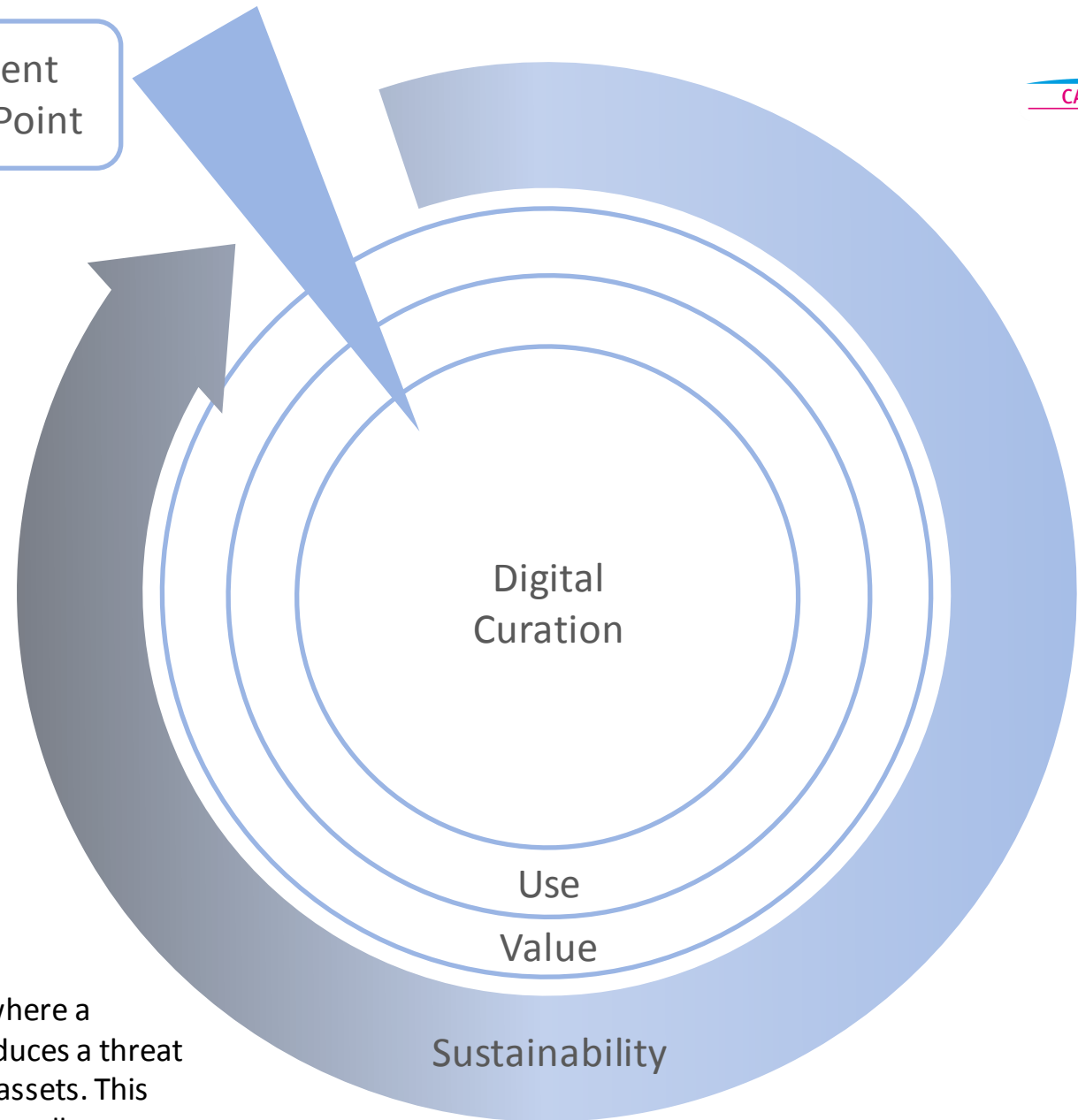


And use (or the potential for use) will realise value, thereby delivering a return on the investment



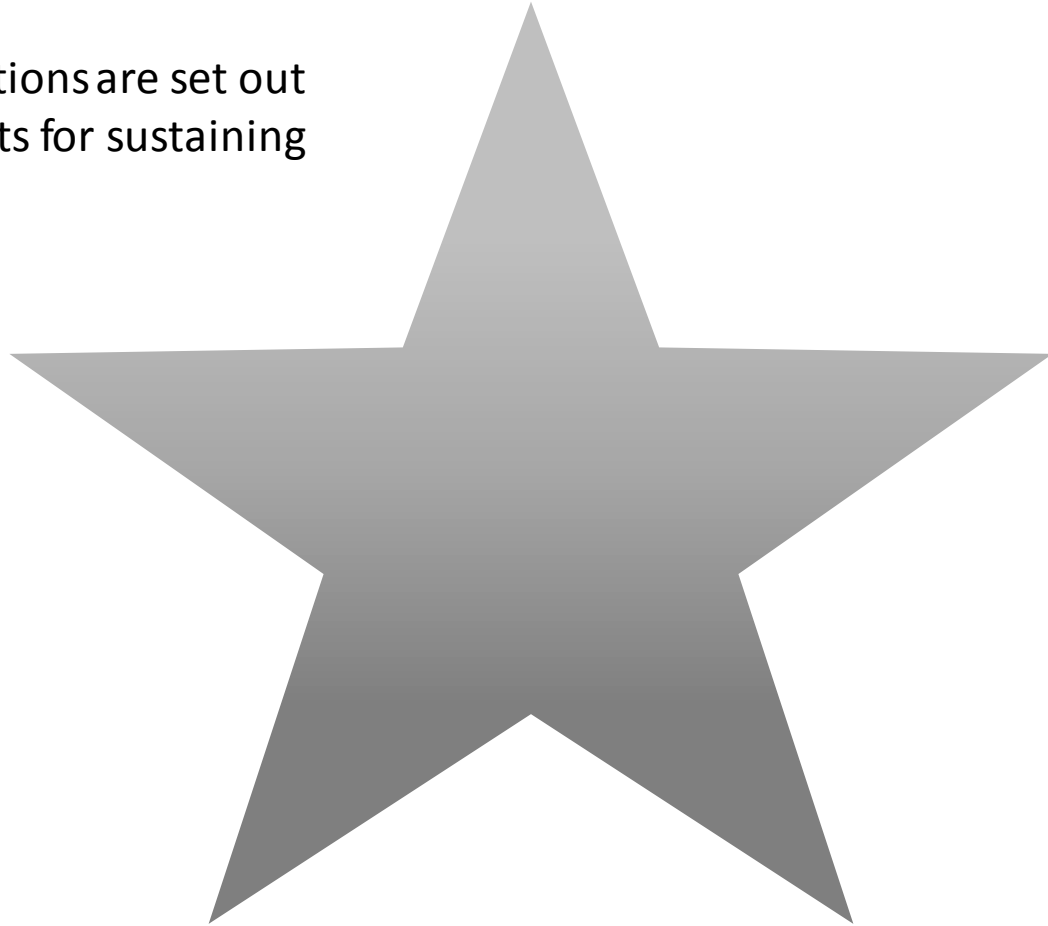
This could play out in a linear fashion with assets being created, curated and then deleted according to a retention schedule. But in the context of sustainability, it is more likely to be a cyclical process

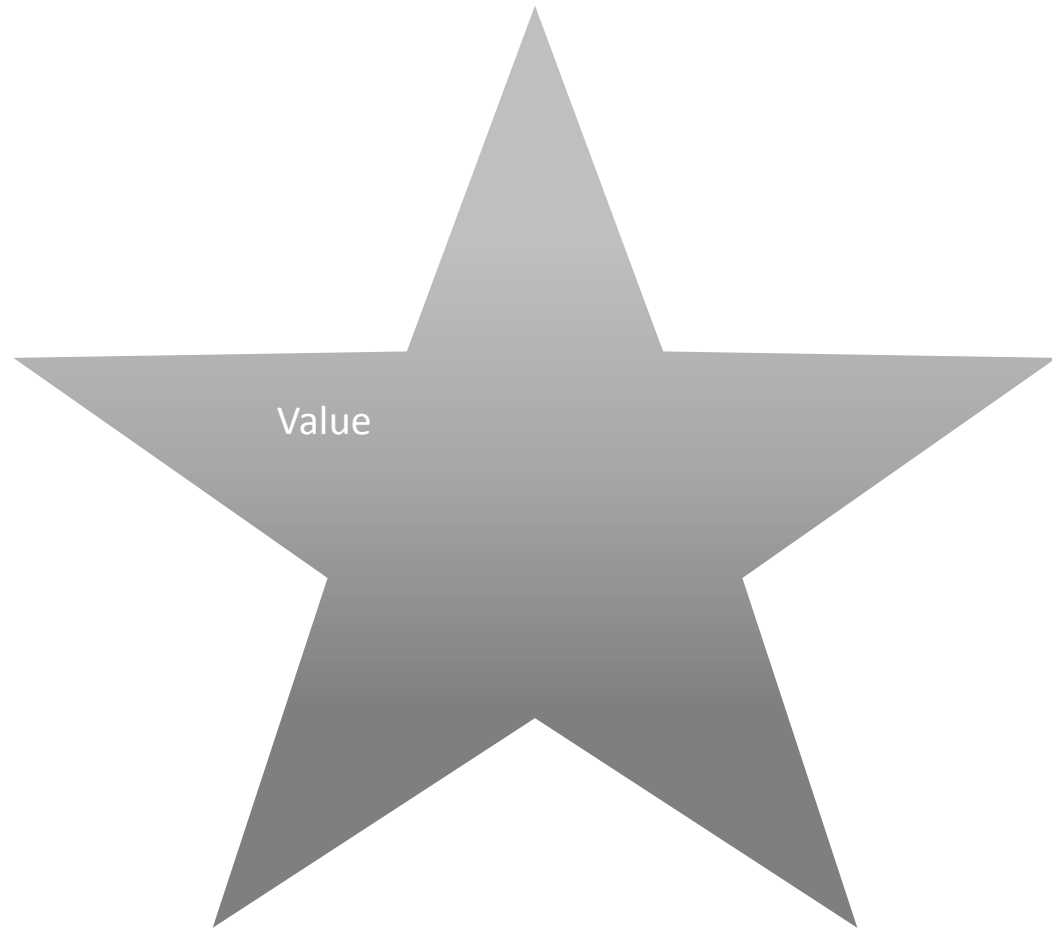
Investment
Decision Point



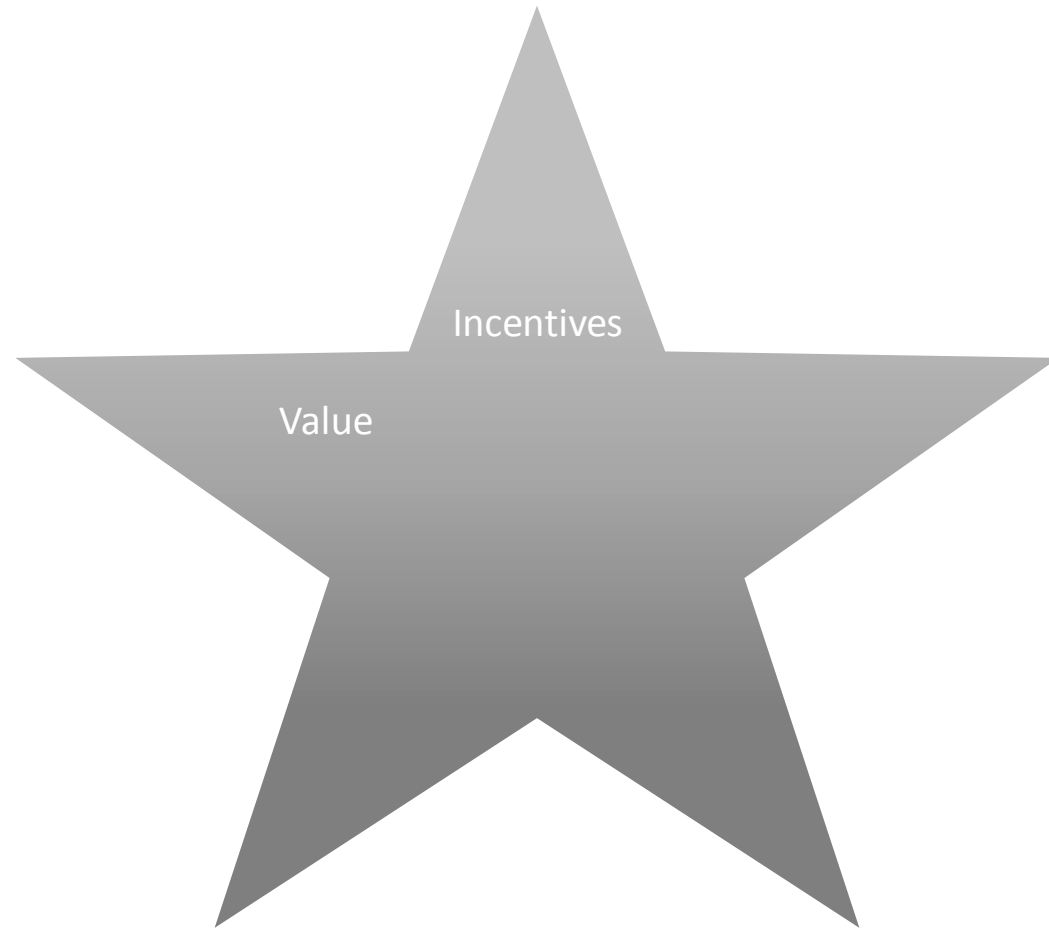
There will be a gap in the cycle where a technical or business issue introduces a threat to the continued viability of the assets. This becomes a decision point ... Do we allocate more resources to tackling the problem?

Five Sustainability Conditions are set out to maximise the prospects for sustaining assets

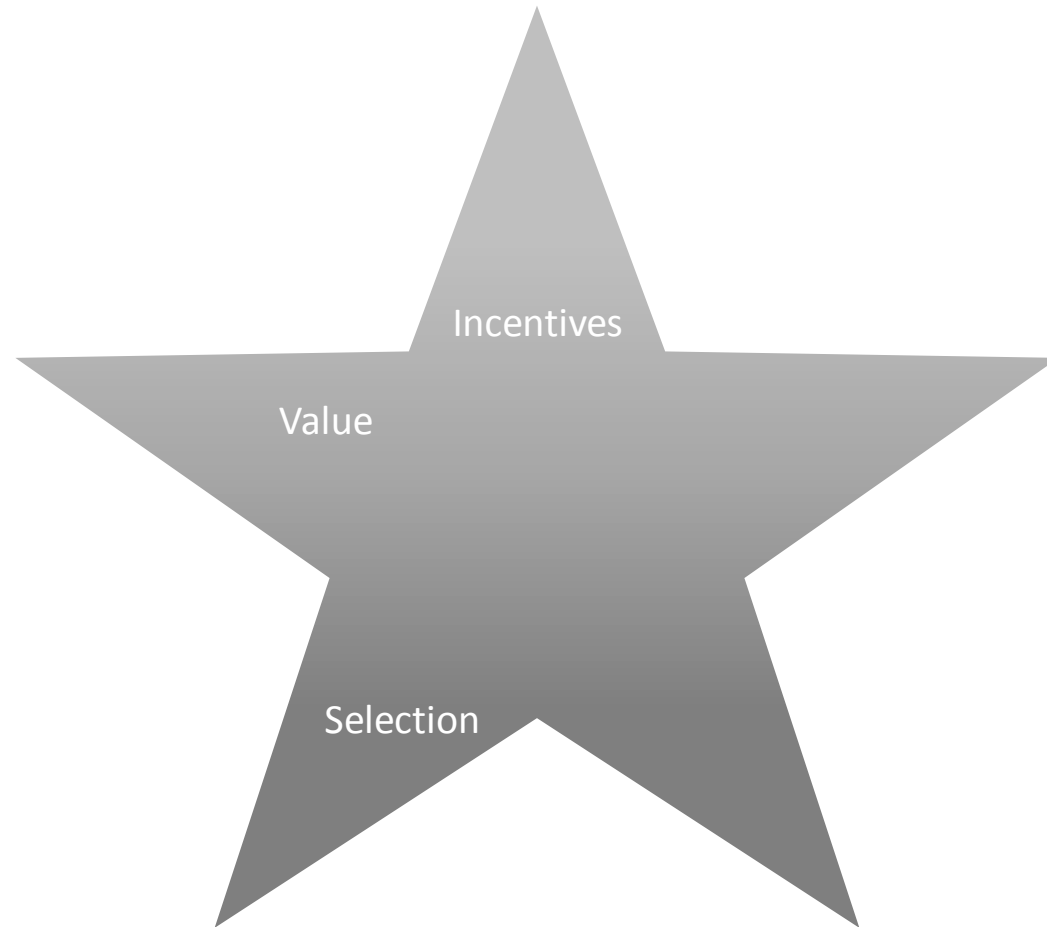




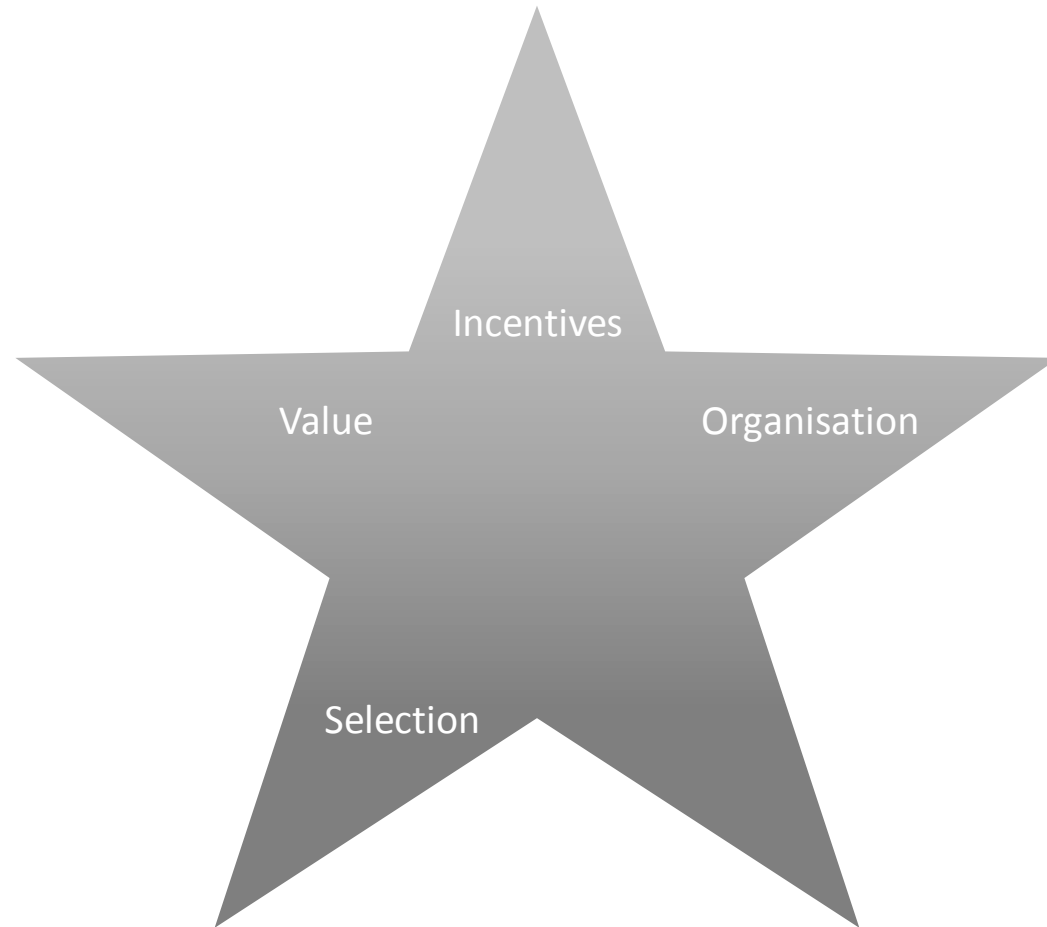
The assets must be understood (or perceived) to have tangible or intangible value



Relevant stakeholders must be sufficiently motivated to support curation



Where resources are scarce then discretion must be used to prioritise curation of the most valuable assets



The organisation should have an appropriate mandate; a supportive governance structure; and be optimally configured to sustain the assets



There must be a sufficient flow of ongoing resources (including financial and human capital) to achieve long-term goals



Key Entities



Three Key Entities are set out which are found in all digital curation contexts. Sustainability requires the nature of these entities to be understood

Three Key Entities are set out which are found in all digital curation contexts. Sustainability requires the nature of these entities to be understood

ASSETS

Every type of digital asset exhibits various attributes or properties that to a greater or lesser extent may affect the how they are curated

STAKEHOLDERS

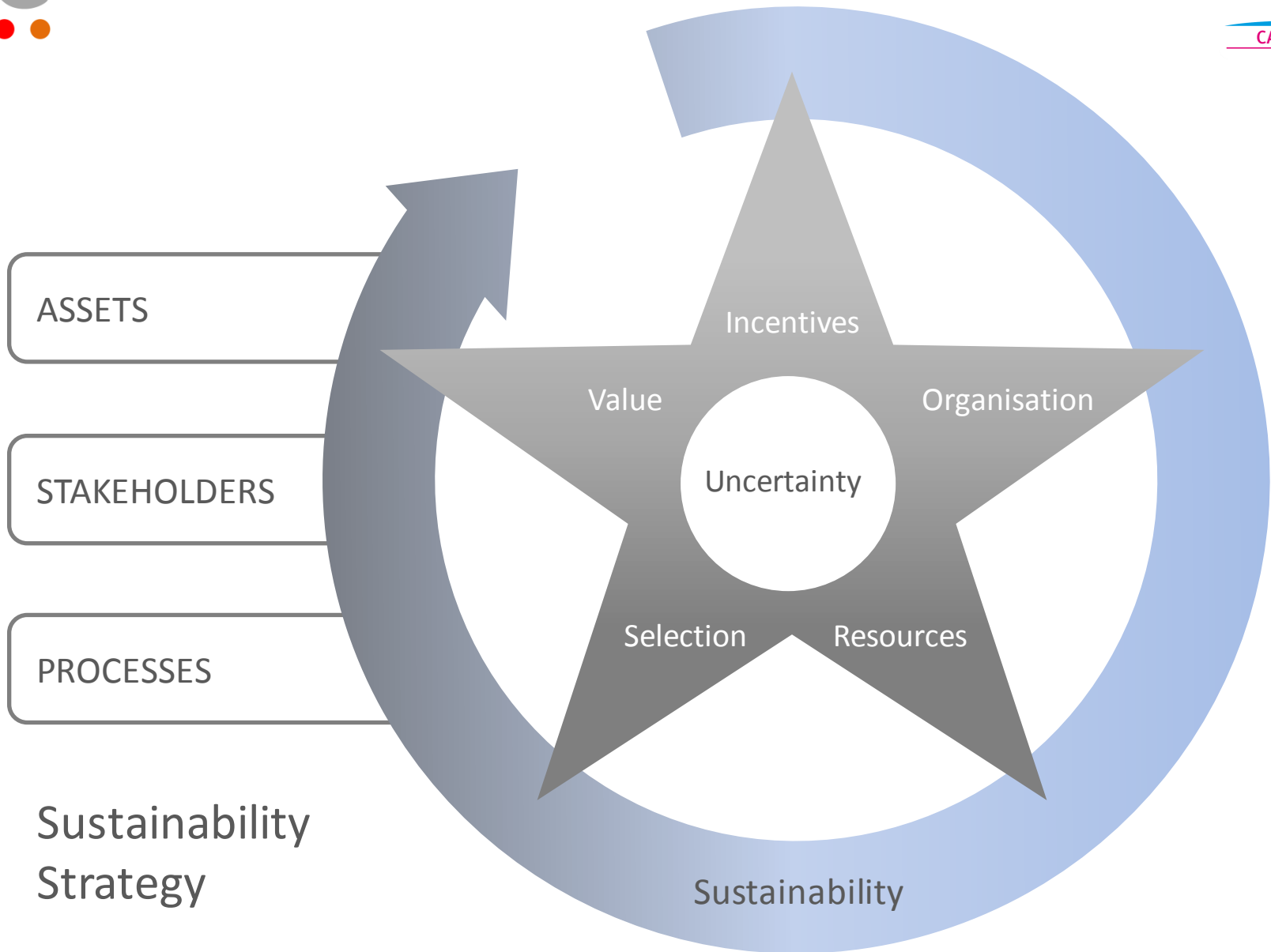
The stakeholder ecosystem for digital assets can be complex and the supply side and demand side should be understood in relation to who is undertaking the curation for the benefit of whom

PROCESSES

The processes involved must be capable of (and optimised for) efficiently enhancing the value of the assets

The inclusion of Economic Uncertainties is an acknowledgement that even the best sustainability strategy cannot accurately predict the future and that some expectation or mitigation of uncertainty (both threats and opportunities) should be built into the strategy where possible





THANK YOU
VERY VERY
MUCH!