Developing impact assessment indicators: Making a proposal for the UK Web Archive

Julie Fukuyama
National Diet Library
ju-fukuy@ndl.go.jp

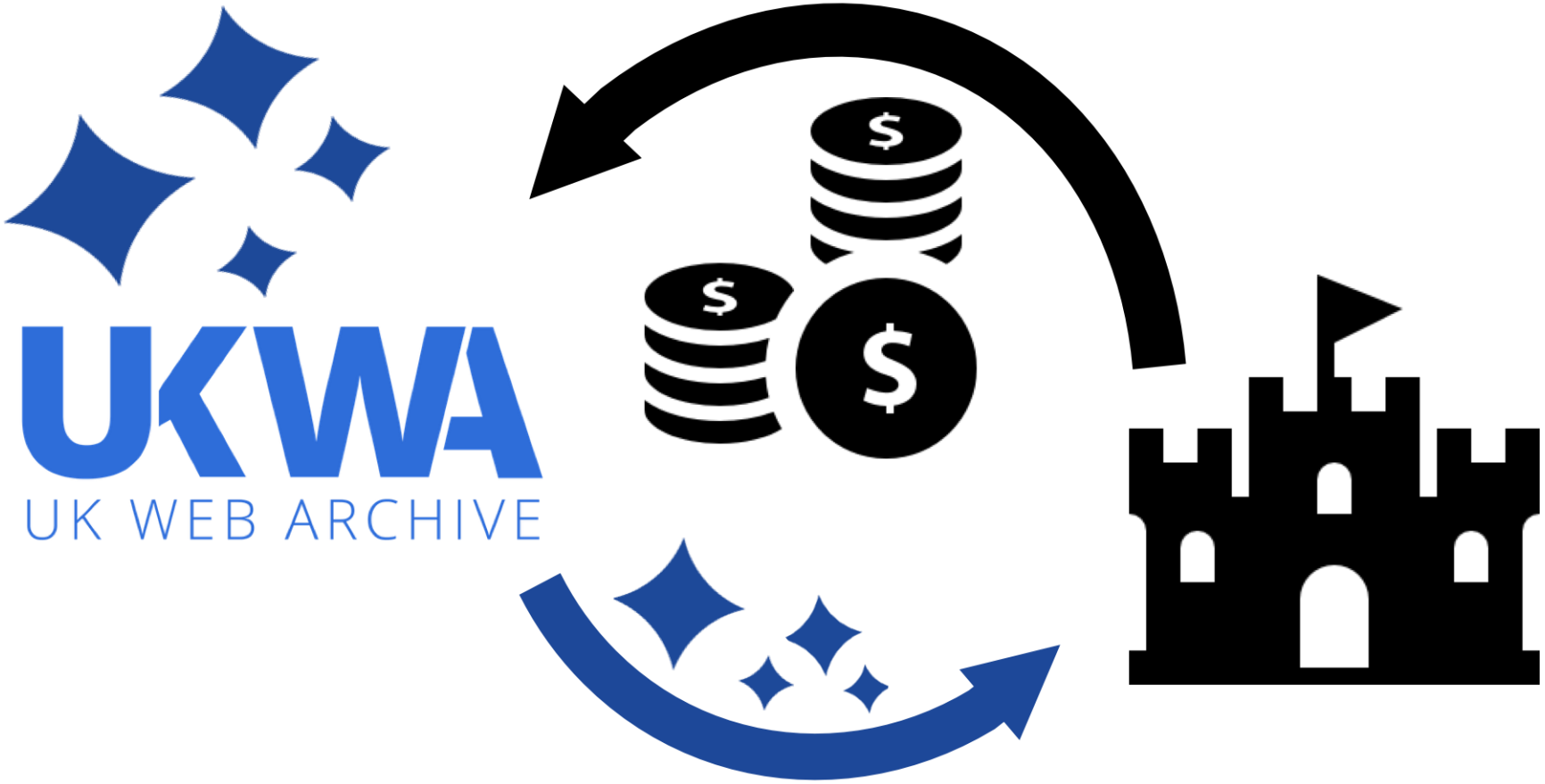
&

Simon Tanner
King's College London
simon.tanner@kcl.ac.uk
@SimonTanner

IIPC Web Archiving Conference 2019, Zagreb, Croatia, 7 June 2019

国立国会図書館 National Diet Library, Japan
KING'S College LONDON

Background



Research Question

What would the indicators of impact assessment for the UKWA be?

UKWA
UK WEB ARCHIVE

Approach to develop the UKWA IA indicators

Stage 1 Set the Context

Digital ecosystem | Objectives | Stakeholder analysis | Situation analysis

Focus areas ← Stakeholders

Strategic Perspectives
Economic / Social / Innovation / Operational

Value Lenses
Utility / Existence / Education / Community / Inheritance

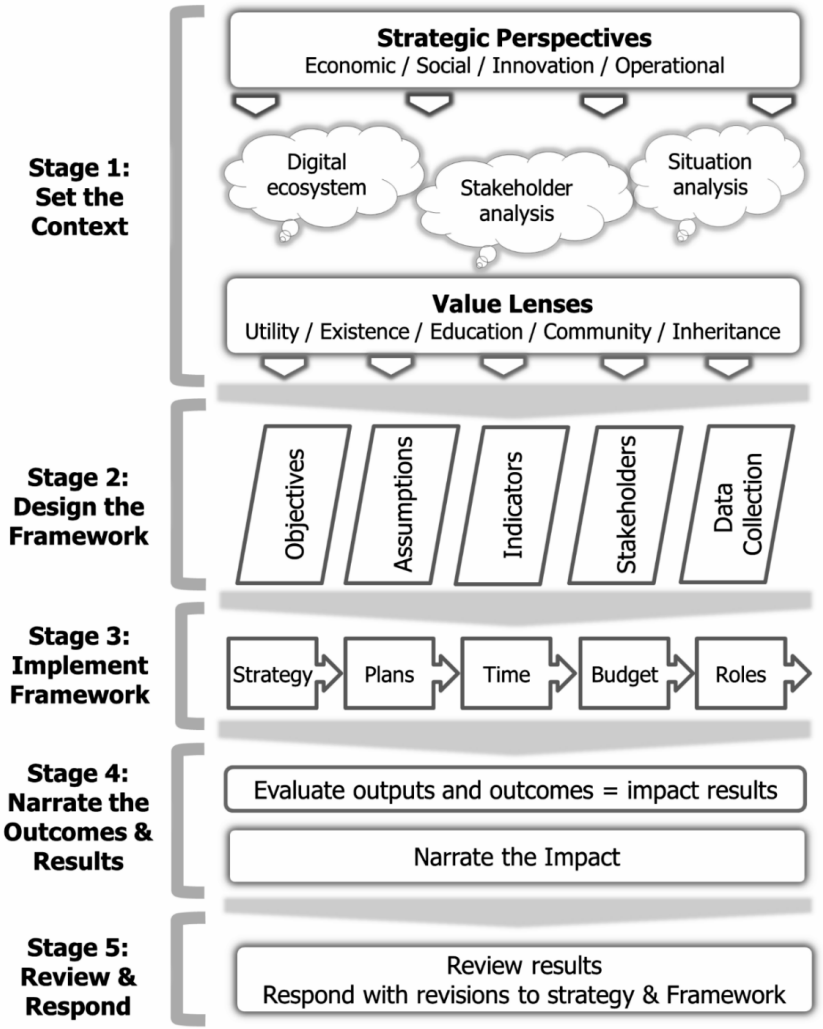
Stage 2 Develop Indicators

Success criteria → Questions → Indicators + Evidence

Stage 3 Evaluate Indicators

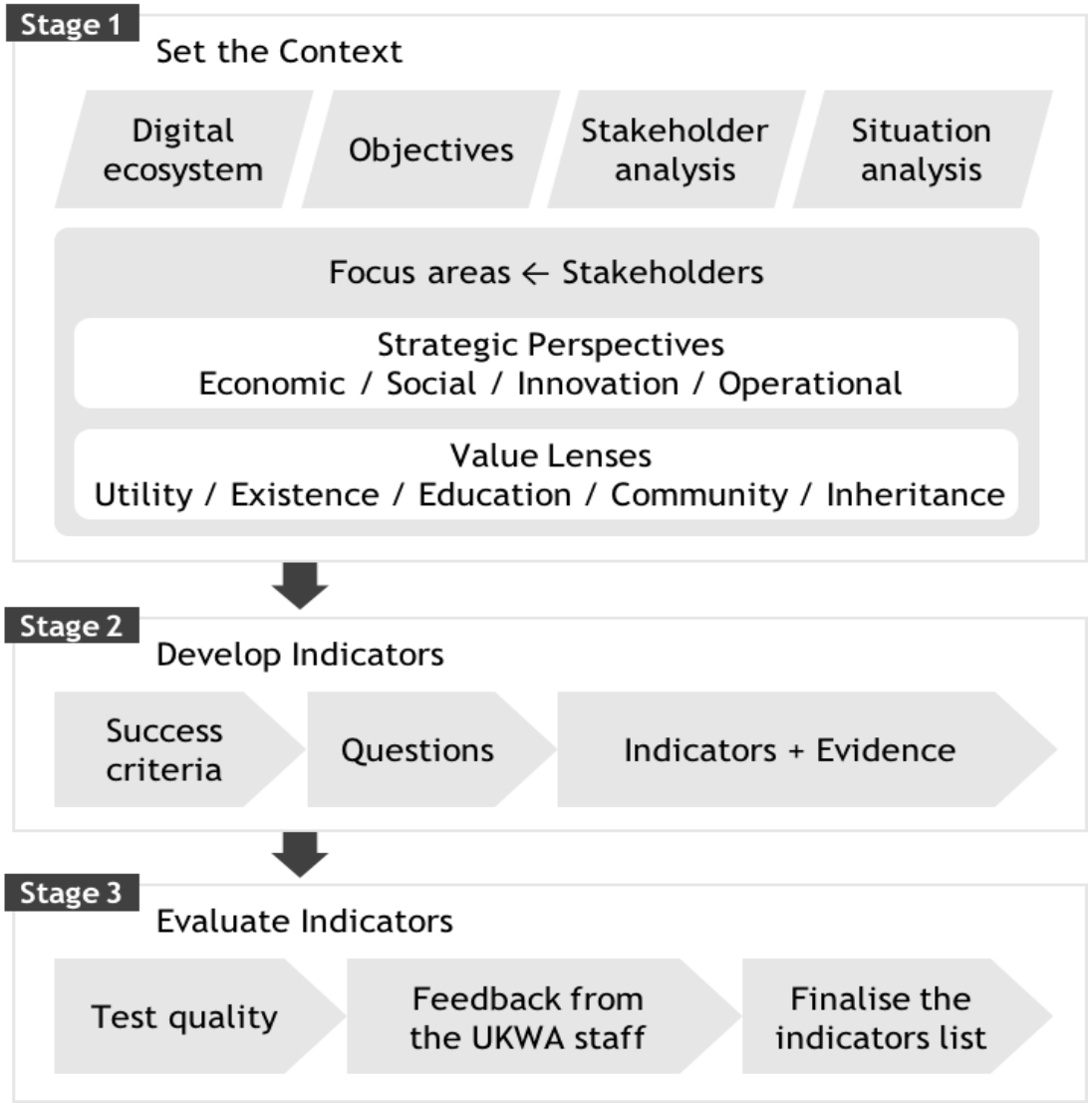
Test quality → Feedback from the UKWA staff → Finalise the indicators list

The Balanced Value Impact Model



http://simon-tanner.blogspot.com/2017/10/BVI-Model-V2.html

Approach to develop the UKWA IA indicators



国立国会図書館 National Diet Library, Japan
KING'S College LONDON

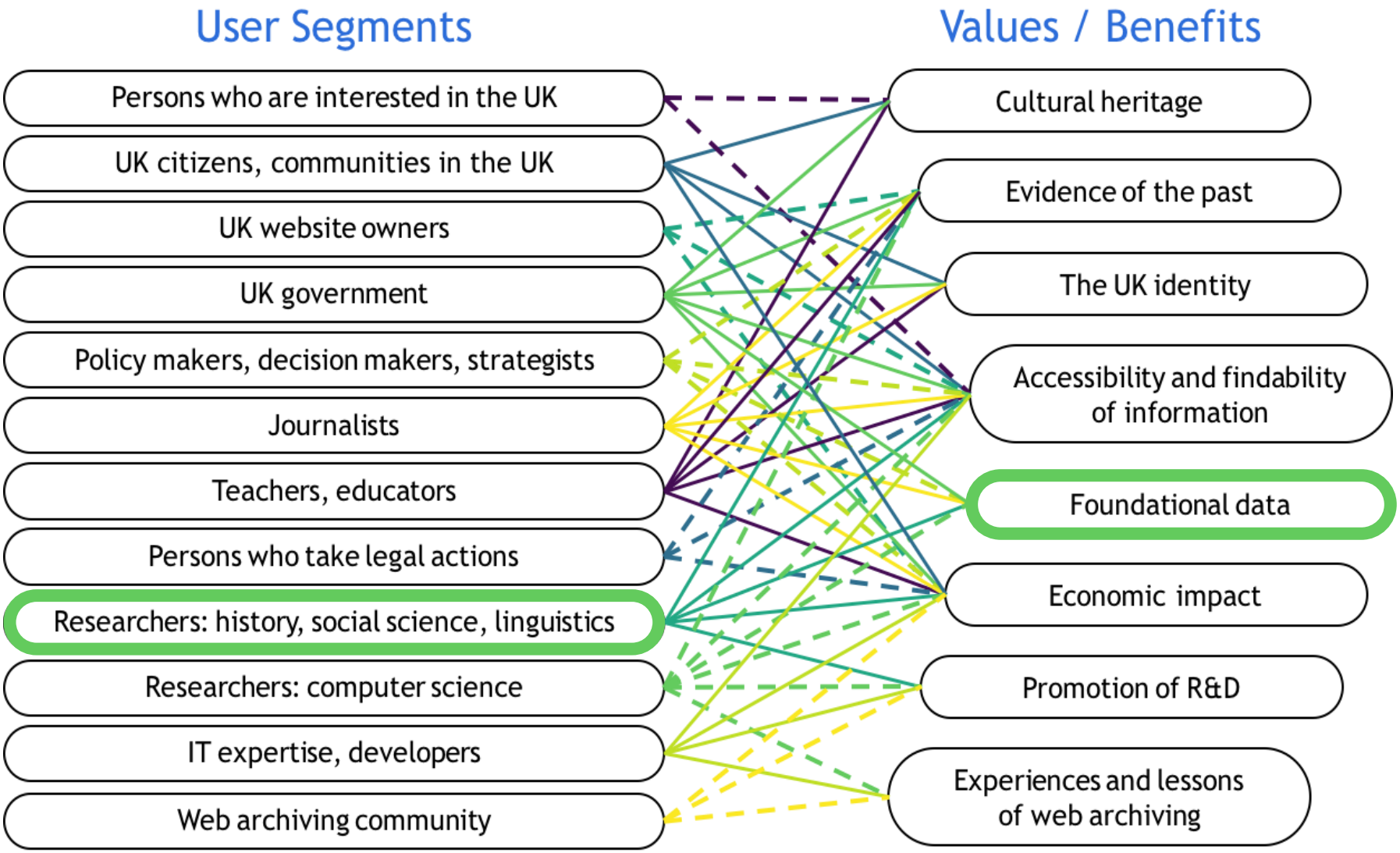
Sorted out mission of the UKWA

Activity	Objects	How	For whom
- Collect - Preserve - Give permanent access	- All of the UK Web Space as national collection of intellectual record	- Comprehensively	- Current and future generations and researchers

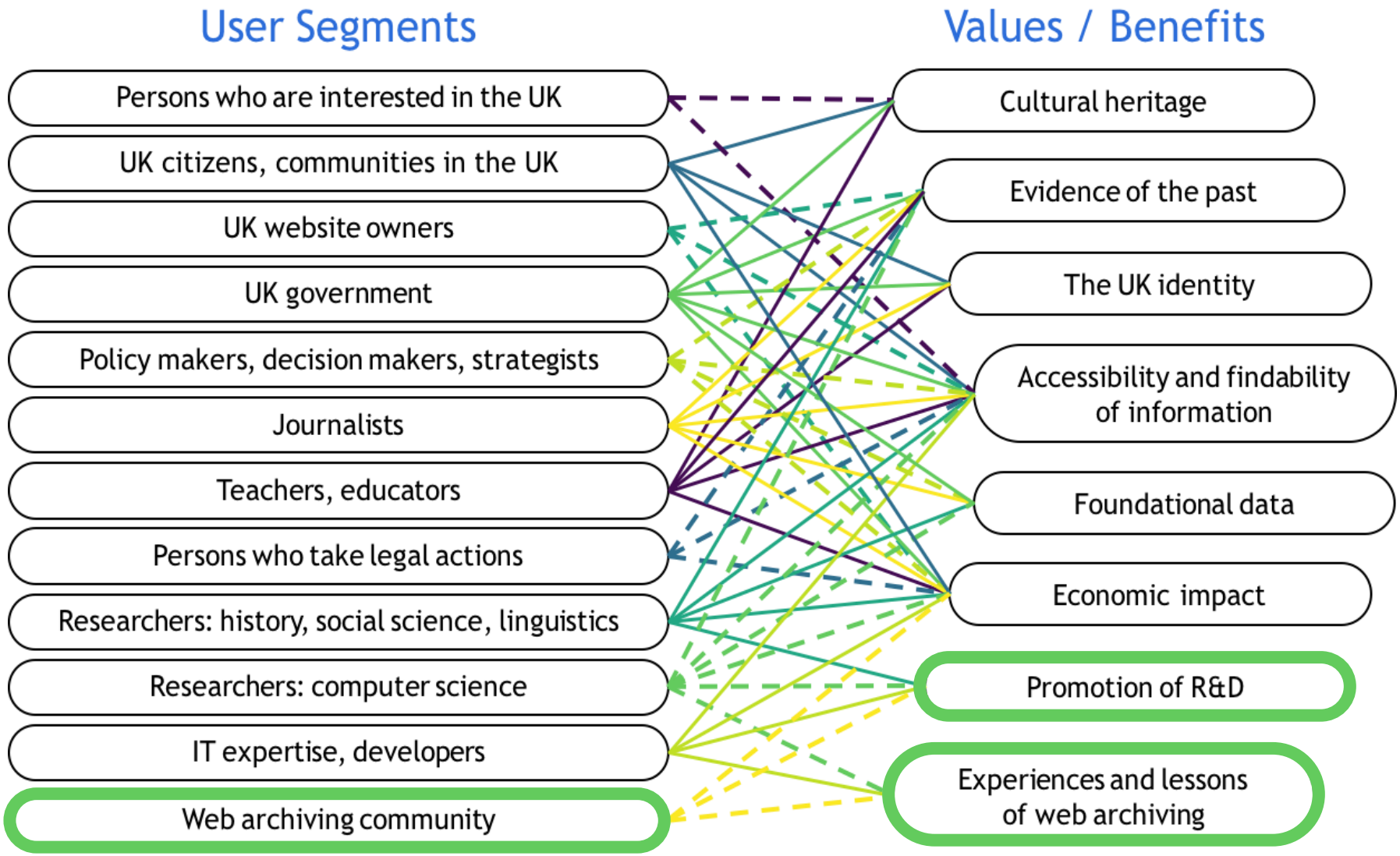
The UKWA's mission gathered from the website, the government's guidance on the regulations, and the Agency for the Legal Deposit Libraries (ALDL)

	Activity	Objects	How	For whom
UKWA	- Collect - Capture - Save - Preserve - Give permanent access	- All of the UK Web Space - Key UK websites - Assets: html pages, images, pdf's, video's - Information that is only available on the Web - Significant witness to our times - Highly ephemeral: altered, disappear	- Annually - Comprehensively	- Future generations - Current & future researchers (academic or otherwise)
Guidance	- Archive - Preserve	- National collection of non-print publications	- Efficient	- Future generations
Agency	- Collect - Preserve - Make it available	- Nation's published output - Nation's intellectual record - Future published heritage	- Systematically	- Future generations - Readers within the deposit libraries

User segments and their benefits



User segments and their benefits



Balancing perspectives and value lenses for the UKWA

Perspectives	Value lenses	Activities	Values
Social impact	Utility	Archiving, GUI, R&D	Satisfying access needs to resources, accessibility, findability
	Existence/ Prestige	Archiving, GUI, Engagement	National/unique collection, preservation of cultural heritage, UK identity, high reputation/acknowledgement
	Education	Archiving, GUI, R&D, Engagement	Satisfying access needs to resources, improvement on knowledge, national/unique collection, preservation of cultural heritage, accessibility, findability, support for creativity/inspiration, learning/skills development
	Community	Archiving, GUI, Engagement	National/unique collection, preservation of cultural heritage, UK identity, high reputation and acknowledgement, encouragement of social/civic engagement
	Inheritance /Bequest	Archiving, GUI, Engagement	National/unique collection, preservation of cultural heritage, UK identity, high reputation/acknowledgement
Economic impact	Utility	Archiving, Datasets, R&D, Engagement	UK identity, contribution to employment, foundational data of research, providing opportunities for R&D, sharing high-level experiences/lessons
Innovation impact	Utility	Archiving, Datasets, R&D, Engagement	Trustworthiness, evidence of the past, foundational data of research, providing opportunities for R&D, sharing high-level experiences/lessons
Operational impact	Utility	R&D, Engagement	Providing opportunities for R&D, sharing high-level experiences/lessons

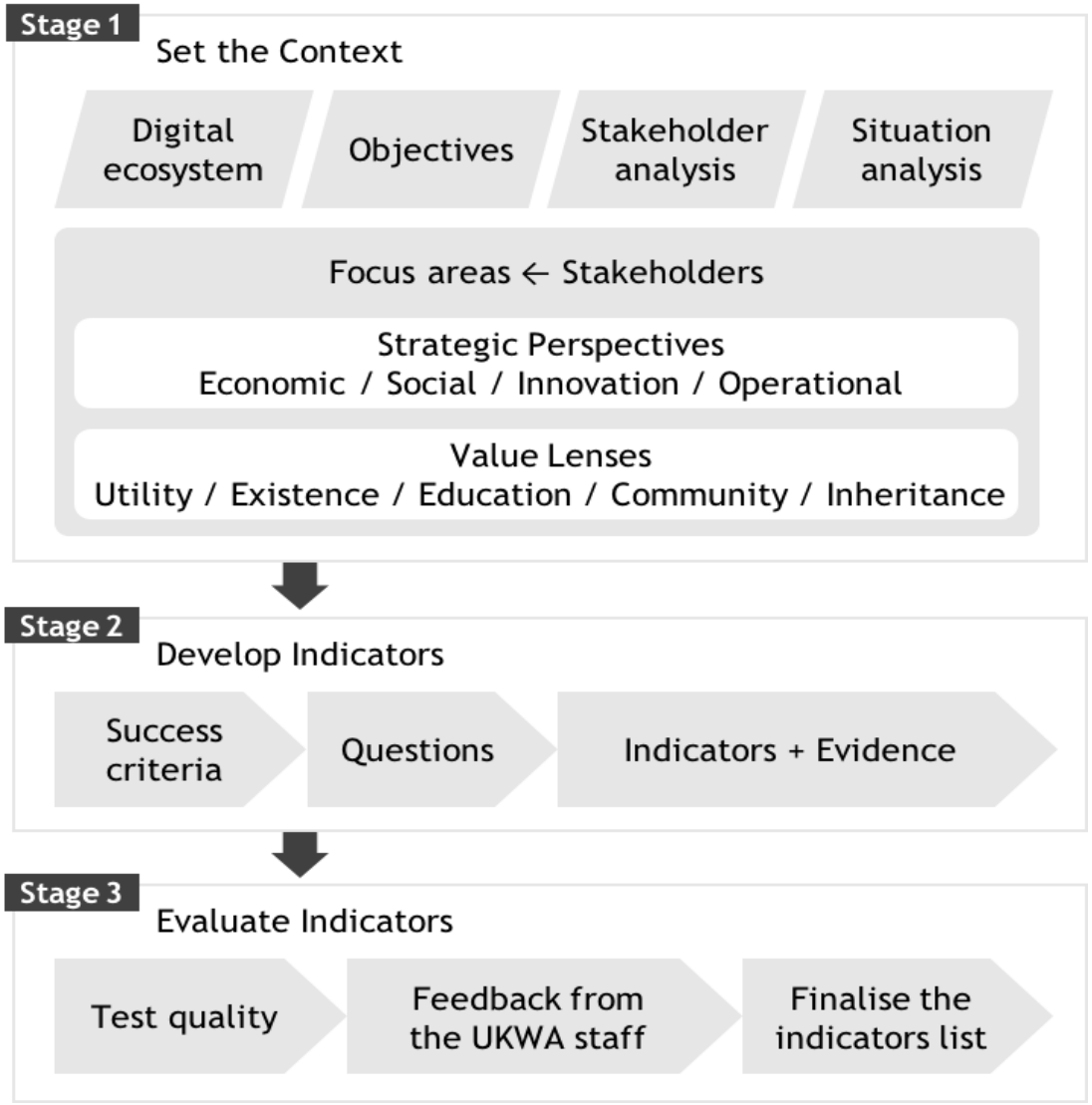
Focus areas of the UKWA: areas to be assessed

*	Values	Stakeholder groups* (*see the legend under the table)	Current activities					Advantage	Importance
			Archiving	Engagement	GUI	R&D	Data sets		
●	Satisfying access needs to resources	per citiz tea	✔					✔	✔
●	Preservation of cultural heritage	per citiz gov tea	✔					✔	✔
	Improvement on knowledge	citiz	✔						
	Contribution to employment	citiz gov IT	✔			✔			
	Support for creativity and inspiration	citiz	✔		✔				
○	Learning and skills development	citiz	✔		✔				✔
○	Encouragement of social/civic engagement	citiz gov tea	✔	✔	✔				✔
●	National collection, UK identity	per citiz own gov jou tea	✔		✔			✔	✔
●	Unique collection	per citiz gov jou tea	✔					✔	✔
●	Trustworthiness, evidence of the past	own gov pol jou tea leg hum sci	✔			✔		✔	✔
●	Accessibility, findability	per citiz own gov pol jou tea leg hum sci IT			✔	✔		✔	✔
●	Foundational data of research	gov pol jou hum sci	✔	✔			✔	✔	✔
●	Providing opportunities for R&D	hum sci IT web		✔		✔		✔	✔
	High reputation/acknowledgement	citiz gov UKWA	✔	✔	✔				
●	Sharing high-level experiences/lessons	web		✔				✔	✔

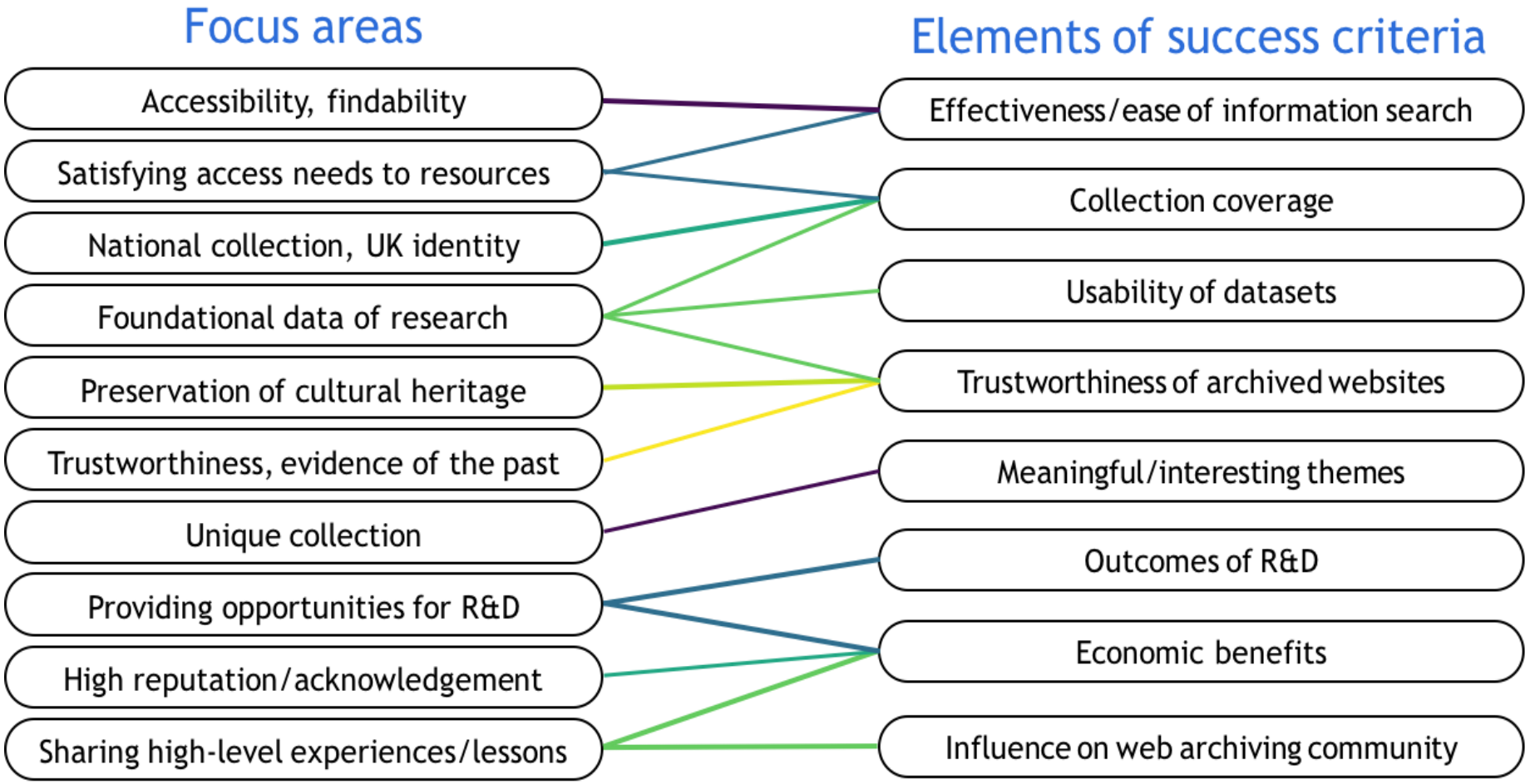
*legend

- ● First priority
- ○ Second priority
- per Persons who are interested in the UK
- citiz UK citizens, communities in the UK
- own UK website owners
- gov UK government
- pol Policy makers, decision makers, strategists
- jou Journalists
- tea Teachers, educators
- leg Persons who take legal actions
- hum Researchers: history, social science, linguistics
- sci Researchers: computer science
- IT IT expertise, developers
- web Web archiving community

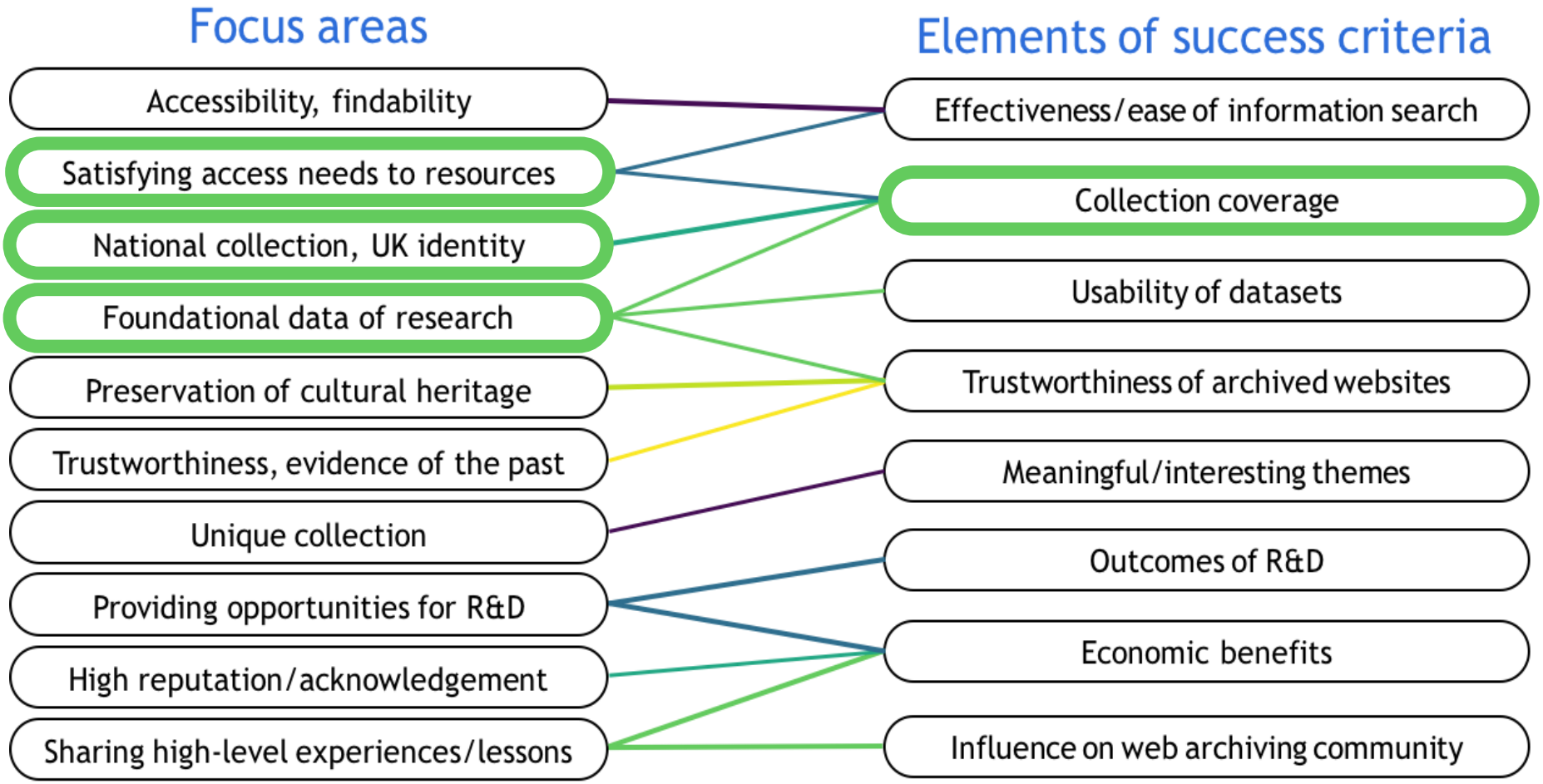
Approach to develop the UKWA IA indicators



Focus areas and elements of success criteria



Focus areas and elements of success criteria



Detailed success criteria of the UKWA

Elements of success criteria	Success criteria
Effectiveness/ease of information search	Users find it effective/easy to search
Collection coverage of existing websites	The UKWA collects the whole UK websites
Usability of datasets	Users find the datasets as usable/meaningful
Trustworthiness of archived websites	Archived websites prove to be unchanged/secure
Meaningful/interesting collections	Users find them meaningful/interesting collections
Outcomes of R&D	Researchers: find research questions, use data, publish the results Developers: find issues, develop tools, tools become widespread
Economic benefits	The UKWA becomes seeds of grants, produces direct/indirect outcome/benefits
Influence on web archiving community	Stakeholders adopt/apply/base the UKWA technologies/research results

Questions for the UKWA

1. Effectiveness/ease of information search
- Do users find it effective/easy to search? - Are users satisfied with the search function? - How much the GUI improved?
2. Collection coverage of existing websites
- Does the UKWA collect the whole UK websites? - What % of the whole UK websites is archived in the UKWA?
3. Usability of datasets
- Do users find the datasets as usable/meaningful? - Are users satisfied with the datasets? - Are users feeling possibilities of using datasets?
4. Trustworthiness of archived websites
- Are the websites securely archived? - Will be the archived websites accessible in the future as well?
5. Meaningful/interesting collections
- Do users find the collections meaningful/interesting collections? - Are users satisfied with the collections?
6. Outcomes of R&D
- Does the UKWA have active partnerships/collaborative projects with researchers? - Did researchers find research questions inspired by the UKWA? - Did researchers use the UKWA for their research? - Did researchers publish the results using the UKWA? - Did developers find issues inspired by the UKWA? - Did developers develop tools for/relevant to/inspired by the UKWA? - Did the tools become widespread?
7. Economic benefits
- How much did the UKWA deliver economic benefits to the society? - How much did the UKWA save money for the society? - How much did the UKWA become seeds of grants? - How much will users accept to pay for the UKWA?
8. Influence on web archiving community
- Did the UKWA have a good effect on web archiving community? - How many web archiving projects adopted/referred to the UKWA technologies/research/experience?

List of the impact assessment indicators

Impact assessment indicator	Aspect measured	Evidence
Effectiveness/ease of information search		
User Satisfaction (B.2.4.2 of ISO 11620:2014)	Effectiveness, ease, accessibility, and usability of the GUI	Users' satisfaction with the information search experience
Collection coverage of existing websites		
Percentage of the UK Websites Collected by the UKWA (modified from A.1.1 of ISO/TR 28118:2009)	Value as the national collection	All websites that fall within the collection criteria of the UKWA
Usability of datasets		
User Satisfaction (B.2.4.2 of ISO 11620:2014)	(1) Usability and (2) value of datasets	Users' satisfaction with (1) usability of datasets and (2) the existence of datasets
Trustworthiness of archived websites		
ISO 16363 Audit and certification of trustworthy digital repositories	Trustworthiness of the archived websites	UKWA's systems and the documentation
Meaningful/interesting collections		
User Satisfaction (B.2.4.2 of ISO 11620:2014)	Users' perceived value of the selective collections	Users' interests about the selective collections
User Assumption	Value of the selective collections for future users	Users' rate for the selective collections based on the assumptions about the future users' satisfaction
Outcomes of R&D		
Number of Contacts with Persons in R&D	UKWA's R&D efforts	Contact records with persons in R&D made by the UKWA in the period
Number of Partnerships and Collaborative Projects	Results of UKWA's R&D effort	Contracts for partnerships and collaborative projects
User Satisfaction (B.2.4.2 of ISO 11620:2014)	R&D users' value of the UKWA	R&D users' satisfaction with the UKWA
User Benefit (modified from 'Question 1' of A.2.4 of ISO 16439:2014)	Reasons for R&D users' value of the UKWA	R&D users' satisfaction reasons in regards to benefits
Economic benefits		
Willingness to Pay	Economic benefits of the UKWA, selective thematic collections, and others	Sample estimates based on contingent valuation method
Influence on web archiving community		
Recognition of Publications	Recognition of the UKWA in web archiving community	Citation count of the publications based on H-index metrics, web traffic analysis of the publications
Number of Publications and Presentations	Visibility of the UKWA to web archiving community	Publications, presentations, blog posts and other appearances
Stakeholder Perception	Contribution of the UKWA to web archiving projects	Stakeholders' experience and thoughts on the impact and contribution of the UKWA to their projects

Test of the quality of the impact assessment indicators

*legend L: Linked to the objectives; R: Reasonable and fair; V: Valid; CD: Not corrupting to disrupt; IF: Not inflexible

Impact assessment indicator	L	R	V	CD	IF
Effectiveness/ease of information search					
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
Collection coverage of existing websites					
Percentage of the UK Websites Collected by the UKWA (modified from A.1.1 of ISO/TR 28118:2009)	✔	✔	✔	✔	✔
Usability of datasets					
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
Trustworthiness of archived websites					
ISO 16363 Audit and certification of trustworthy digital repositories	✔	✔	✔	✔	✔
Meaningful/interesting collections					
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
User Assumption	✔	?	?	?	✔
Outcomes of R&D					
Number of Contacts with Persons in R&D	✔	✔	✔	✔	✔
Number of Partnerships and Collaborative Projects	✔	✔	✔	✔	✔
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
User Benefit (modified from 'Question 1' of A.2.4 of ISO 16439:2014)	✔	✔	✔	✔	✔
Economic benefits					
Willingness to Pay	✔	✔	✔	✔	✔
Influence on web archiving community					
Recognition of Publications	✔	✔	✔	✔	✔
Number of Publications and Presentations	✔	✔	✔	✔	✔
Stakeholder Perception	✔	✔	✔	✔	✔

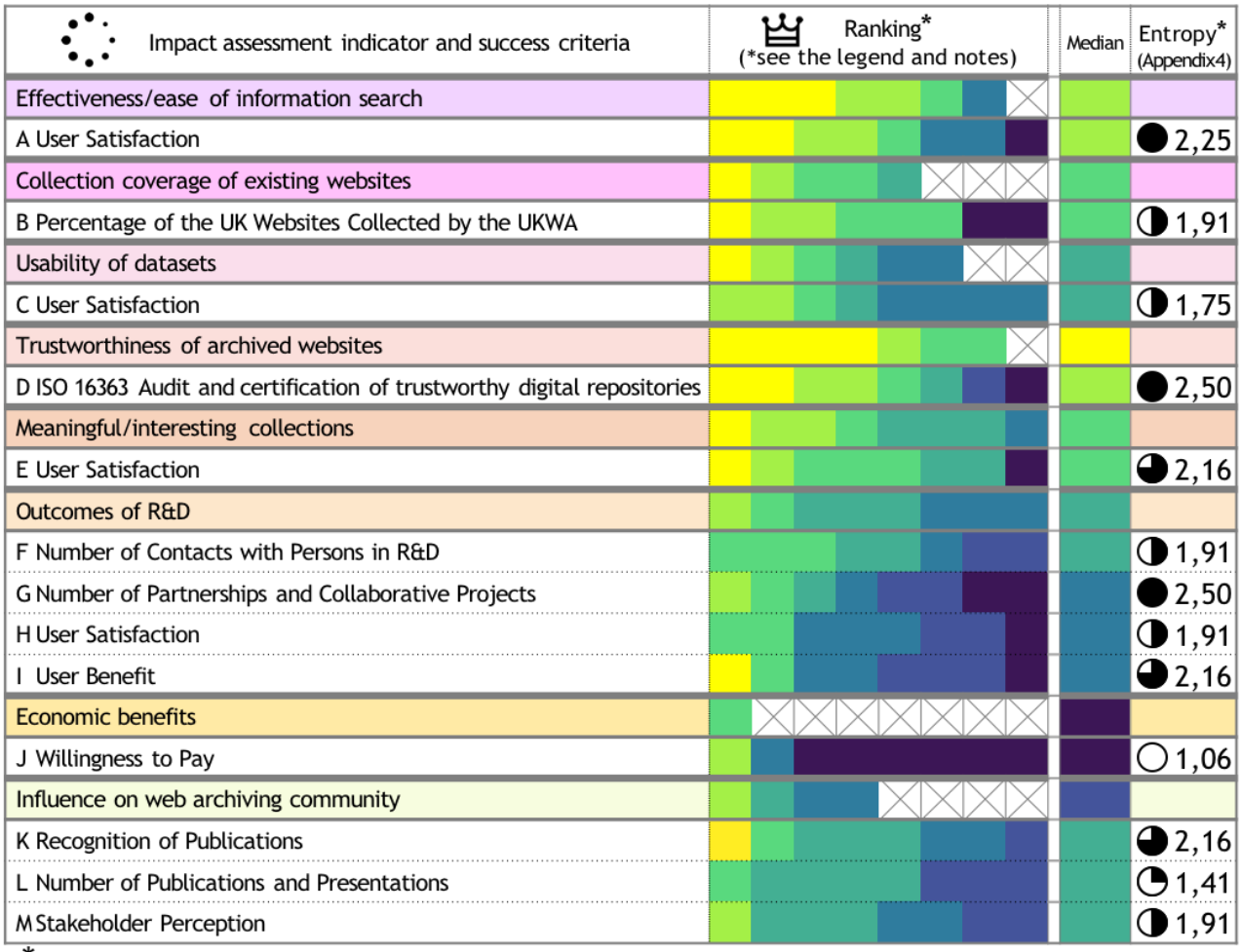
Test of the quality of the impact assessment indicators

*legend L: Linked to the objectives; R: Reasonable and fair; V: Valid; CD: Not corrupting to disrupt; IF: Not inflexible

Impact assessment indicator	L	R	V	CD	IF
Effectiveness/ease of information search					
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
Collection coverage of existing websites					
Percentage of the UK Websites Collected by the UKWA (modified from A.1.1 of ISO/TR 28118:2009)	✔	✔	✔	✔	✔
Usability of datasets					
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
Trustworthiness of archived websites					
ISO 16363 Audit and certification of trustworthy digital repositories	✔	✔	✔	✔	✔
Meaningful/interesting collections					
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
User Assumption	✔	?	?	?	✔
Outcomes of R&D					
Number of Contacts with Persons in R&D	✔	✔	✔	✔	✔
Number of Partnerships and Collaborative Projects	✔	✔	✔	✔	✔
User Satisfaction (B.2.4.2 of ISO 11620:2014)	✔	✔	✔	✔	✔
User Benefit (modified from 'Question 1' of A.2.4 of ISO 16439:2014)	✔	✔	✔	✔	✔
Economic benefits					
Willingness to Pay	✔	✔	✔	✔	✔
Influence on web archiving community					
Recognition of Publications	✔	✔	✔	✔	✔
Number of Publications and Presentations	✔	✔	✔	✔	✔
Stakeholder Perception	✔	✔	✔	✔	✔

Feedback collection

Feedback from UKWA staffs



*legend: The entropy cells' marks signify the levels of entropy as follows.

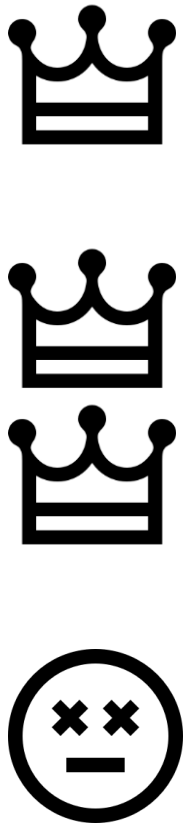
Higher entropy ● ◕ ◑ ◔ ○ Lower entropy

The ranking cells signify the levels of importance as below.

Level	Highest	High	High Middle	Middle	Low Middle	Low	Lowest	
Indicator	1	2-3	4-5	6-7	8-9	10-11	12-13	Not Ranked
Success Criteria	1	2	3	4	5			

Feedback from UKWA staffs

Impact assessment indicator and success criteria	Ranking* (*see the legend and notes)	Median	Entropy* (Appendix4)
Effectiveness/ease of information search			
A User Satisfaction			2,25
Collection coverage of existing websites			
B Percentage of the UK Websites Collected by the UKWA			1,91
Usability of datasets			
C User Satisfaction			1,75
Trustworthiness of archived websites			
D ISO 16363 Audit and certification of trustworthy digital repositories			2,50
Meaningful/interesting collections			
E User Satisfaction			2,16
Outcomes of R&D			
F Number of Contacts with Persons in R&D			1,91
G Number of Partnerships and Collaborative Projects			2,50
H User Satisfaction			1,91
I User Benefit			2,16
Economic benefits			
J Willingness to Pay			1,06
Influence on web archiving community			
K Recognition of Publications			2,16
L Number of Publications and Presentations			1,41
M Stakeholder Perception			1,91



*legend: The entropy cells' marks signify the levels of entropy as follows.

The ranking cells signify the levels of importance as below.

Level	Highest	High	High Middle	Middle	Low Middle	Low	Lowest	Not Ranked
Indicator	1	2-3	4-5	6-7	8-9	10-11	12-13	
Success Criteria	1	2	3	4	5			

国立国会図書館 National Diet Library, Japan
KING'S College LONDON

Feedback from UKWA staffs

Impact assessment indicator and success criteria	Ranking* (*see the legend and notes)	Median	Entropy* (Appendix4)
Effectiveness/ease of information search			
A User Satisfaction			● 2,25
Collection coverage of existing websites			
B Percentage of the UK Websites Collected by the UKWA			◑ 1,91
Usability of datasets			
C User Satisfaction			◑ 1,75
Trustworthiness of archived websites			
D ISO 16363 Audit and certification of trustworthy digital repositories			● 2,50
Meaningful/interesting collections			
E User Satisfaction			◕ 2,16
Outcomes of R&D			
F Number of Contacts with Persons in R&D			◑ 1,91
G Number of Partnerships and Collaborative Projects			● 2,50
H User Satisfaction			◑ 1,91
I User Benefit			◕ 2,16
Economic benefits			
J Willingness to Pay			○ 1,06
Influence on web archiving community			
K Recognition of Publications			◕ 2,16
L Number of Publications and Presentations			◔ 1,41
M Stakeholder Perception			◑ 1,91

2,50

1,06

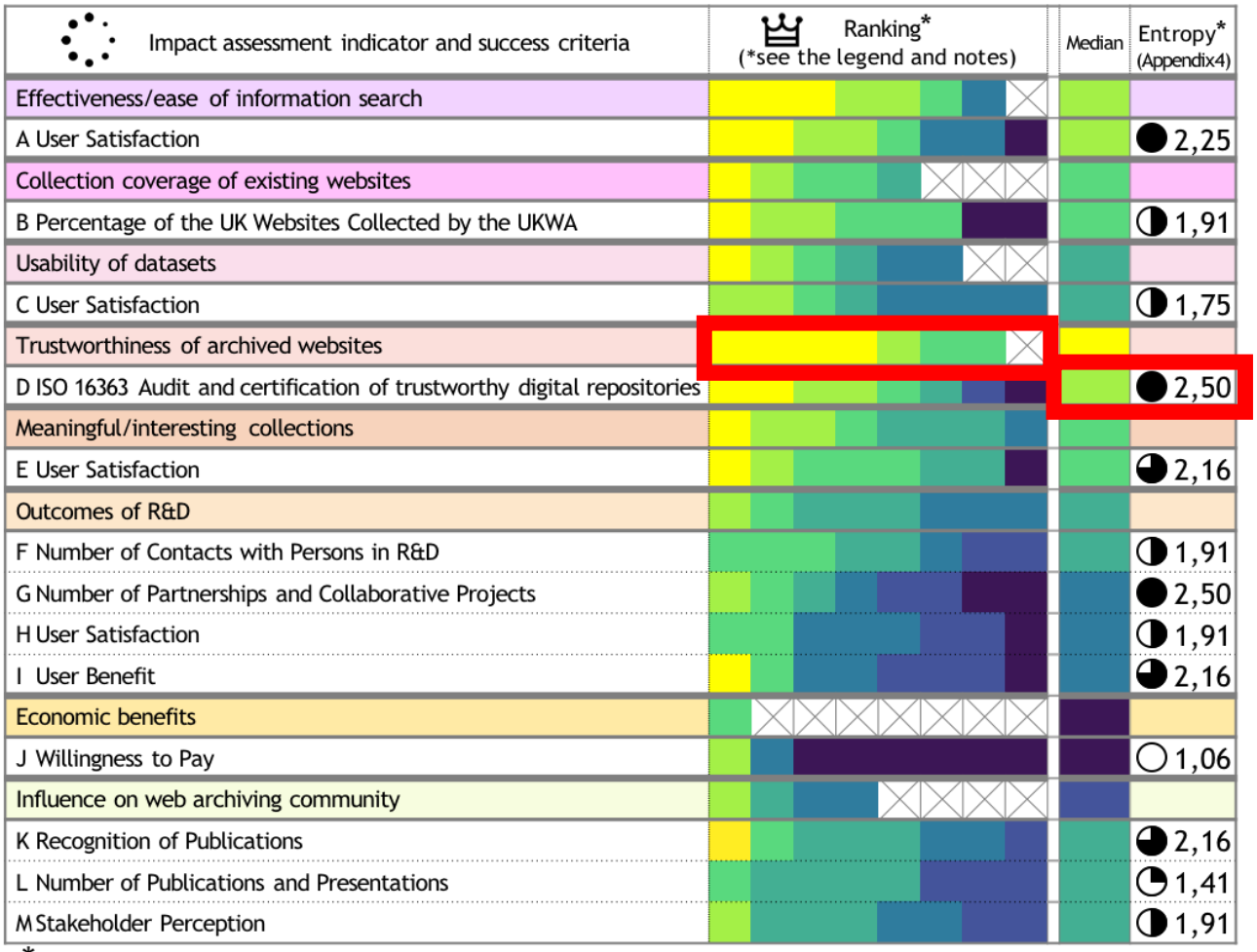
*legend: The entropy cells' marks signify the levels of entropy as follows.

Higher entropy ● ◕ ◑ ◔ ○ Lower entropy

The ranking cells signify the levels of importance as below.

Level	Highest	High	High Middle	Middle	Low Middle	Low	Lowest	Not Ranked
Indicator	1	2-3	4-5	6-7	8-9	10-11	12-13	
Success Criteria	1	2	3	4	5			

Feedback from UKWA staffs



*legend: The entropy cells' marks signify the levels of entropy as follows.

Higher entropy ● ◕ ◑ ◔ ○ Lower entropy

The ranking cells signify the levels of importance as below.

Level	Highest	High	High Middle	Middle	Low Middle	Low	Lowest	
Indicator	1	2-3	4-5	6-7	8-9	10-11	12-13	Not Ranked
Success Criteria	1	2	3	4	5			

The final impact assessment indicators list

Trustworthiness of archived websites
(ISO 16363 Audit rejected, new indicator is needed)
Effectiveness/ease of information search
User Satisfaction
Meaningful/interesting collections
User Satisfaction
Usability of datasets
User Satisfaction
Collection coverage of existing websites
Percentage of the UK Websites collected by the UKWA

Outcomes of R&D
Number of contacts with persons in R&D
Number of Partnerships and Collaborative Projects
User Satisfaction
User Benefit
Influence on web archiving community
Recognition of Publications
Number of Publications and Presentations
Stakeholder Perception
Economic benefits
Willingness to Pay

Recommendations

- To adopt our indicators as a starting point to consider impact indicators
- To discuss among partners and with stakeholders
- To focus upon measuring what you can measure by conducting small scale internal pilot impact assessments

Acknowledgements

Thanks to:

- Nicola Bingham (The British Library)
- Helena Byrne (The British Library)
- The staff of the UK Web Archive