

Scenarios_WPs_matrix

	PT1 (Ds)	PT2 (Ap)	PT3 (Ex)	PT4 (PM)	PT5 (RI)	PC1 (C)	PC2 (A)	PC3 (Q)	PW1 (W)	PW2 (Po)	PW3 (PI)
LSDRT1 Unknown preservation risks in large media files	X[1]			X		X			X[2]		
LSDRT2 Digitised TIFFs do not meet storage and access requirements	X[3]			X		X	X	X			X[4]
LSDRT3 Do acquired files conform to an agreed technical profile, are they valid and are they complete?	X[5]					X			X[6]	X[7]	
LSDRT4 Out-of-sync Sound and Video in wmv to Video Format-X Migration Results	X[1]						X	X			X?
LSDRT5 Detect audio files with very bad sound quality	X[8]					X				X?[9]	
LSDRT6 Migrate mp3 to wav	X[10]			X		X	X	X			X[11]
LSDRT7 Characterise and Validate very large video files	X[12]					X				X[13]	
LSDRT9 Characterisation or large amounts of wav audio	X[14]					X				X[15]	
LSDRT10 Camera raw files pose preservation risk	X[16]					X	X[17]	X			X[18]
LSDRT11 Image based document comparison approach	X[19]					X		X			
LSDRT12 Detect undesired influence of lossy JP2 compression on OCR in the absence of ground truth	X[20]						X?	X			
RDST1 Scientific Data Ingest-related Scenario	X[21]					X		X		X	
RDST2 Format Migration of (raw) Scientific Datasets	X[10]						X	(X)		X	X[22]
RDST3 Maintaining understandability and usability of raw data through external resources	X[23]					(X)	X?		X[24]	X[25]	X
RDST4 Preserving the value of raw data and verifiability of processed datasets forming part of a scientific workflow	X[26]										
WCT1 Comparison of Web Archive pages	X[27]							X	X		
WCT2 ARC to WARC migration	X[28]						X	X			
WCT3 Characterise web content in ARC and WARC containers at State and University Library Denmark	X[29]					X			X[30]		

Scenarios_WPs_matrix

	PT1 (Ds)	PT2 (Ap)	PT3 (Ex)	PT4 (PM)	PT5 (RI)	PC1 (C)	PC2 (A)	PC3 (Q)	PW1 (W)	PW2 (Po)	PW3 (PI)
WCT4 Web Archive Mime-Type detection at Austrian National Library	X[31]					X			X		
WCT6 (W)ARC to HBase migration	X[32]						X	X			
WCT7 Format obsolescence detection	X[33]					X		X	X		
WCT8 Huge text file analysis using hadoop	X[34]								X		

1. scalability (size, complexity, heterogeneity), IO efficiency, robustness

2. be able to integrate measures of format registries and content profiling to enable automated validation against control policies.

3. scalability (#objects, size), IO efficiency, robustness

4. efficiency: effort required to evaluate actions and create a plan

automation: ability to deploy directly, ability to specify SLAs directly

maturity of the process framework to create plans

5. scalability (#objects, size), IO efficiency, robustness

6. scalability of the content profiler to analyse large content sets for properties of interest

coverage of policies-formats-properties

7. expressiveness of the policy model to cover aspects of profiles

8. scalability (#objects, size, heterogeneity), IO efficiency, robustness

9. should "bad" be defined in a control policy against a content profile, or is this a one-time run?

10. scalability (#objects, size), IO efficiency, robustness

11. efficiency: effort required to do evaluation and create plan

automation: ability to deploy directly, ability to specify SLAs directly

maturity of the process framework to create plans

12. scalability (size, complexity, heterogeneity), IO efficiency, robustness

13. expressiveness of the policy model to cover aspects of interest

14. scalability (#objects, size, heterogeneity), IO efficiency, robustness

15. expressiveness of the policy model to cover aspects of interest

16. scalability (complexity), robustness

17. in this scenario, the actions essentially are off-the-shelf migration tools. nothing to be developed in scope, probably (but who knows ;)

18. efficiency: effort required to create a plan

automation: ability to deploy directly, ability to specify SLAs directly

maturity of the process framework to create plans

19. scalability (#objects, size, complexity), IO efficiency, parallelization (algorithmic complexity), robustness

- 20. scalability (#objects), IO efficiency, robustness
- 21. scalability (#objects, size), IO efficiency, PT repository integration, robustness
- 22. efficiency: effort required to evaluate actions and create plan

automation: ability to deploy directly, ability to specify SLAs directly

maturity of the process framework to create plans

- 23. not applicable at present ?

- 24. coverage of policies-formats-properties

- 25. expressiveness
- 26. PT: scalability (#objects, size, heterogeneity), IO efficiency, robustness

- 27. scalability (#objects, size, heterogeneity, complexity), IO efficiency, robustness
- 28. scalability (size, complexity), IO efficiency, robustness
- 29. scalability (size, heterogeneity, complexity), IO efficiency, robustness

- 30. scalability of the content profiler to analyse large content sets for properties of interest
- 31. scalability (#objects, heterogeneity), IO efficiency, robustness
- 32. scalability (#objects, size, heterogeneity, complexity), IO efficiency, robustness

- 33. scalability (#objects, heterogeneity, complexity), IO efficiency, robustness

- 34. scalability (size), IO efficiency, robustness