Introduction

* **1. Name of your organisation:**

* **2. Country (of your organisation):**

Policies and Procedures

The following questions address the policies and procedures regarding data management (acquisition, storage and access). With digital research data is meant all digital information regarding scientific research. This ranges from experimental data, theoretical models, to official publications in a journal.

* **3. Which kind of digital material is stored at your organisation? (multiple answers possible)**

   - Journal and e-journal publications

   - e-books

   - Reports

   - Theses

   - Data sets

   - Auxiliary material (e.g. software)

   - Illustrative material (e.g. images, videos, animations and sound)

   - Don't know, please forward this survey to another colleague (provide email address later in this survey)

   - Other (please specify)

Policies and Procedures (cont.)

* 4. Please indicate which of the following types of digital research data you store: (multiple answers possible)

- Standard office documents (text documents, spreadsheets, presentations)
- Network-based data (web sites, e-mail, chat history, etc.)
- Databases (DBASE, MS Access, Oracle, MySQL, etc.)
- Images (JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.)
- Structured graphics (CAD, CAM, 3D, VRML, etc.)
- Audiovisual (multimedia) data (WAVE, MP3, MP4, Flash, etc.)
- Scientific and statistical data formats (SPSS, FITS, GIS, etc.)
- Raw data (device specific output)
- Plain text (TXT in various encodings)
- Structured text (XML, SGML, etc.)
- Archived data (ZIP, RAR, JAR, etc.)
- Software applications (modelling tools, editors, IDE, compilers, etc.)
- Source code (scripting, Java, C, C++, Fortran, etc.)
- Configuration data (parameter settings, logs, library files)
- Other, please specify below.

* 5. Please provide us with an estimate of the volume of stored digital data (excluding back ups) as well an estimate of its volume in 2 and 5 years.

	0MB	1-100MB	100MB-1GB	1GB-1TB	1TB-1PB	1PB-10PB	>10PB	Don't know
Current	j	j	j	j	j	j	j	j
In 2 Years	j	j	j	j	j	j	j	j
In 5 Years	j	j	j	j	j	j	j	j

* 6. Do you have policies and procedures in place which determine what kinds of data is accepted for storage/preservation by your organisation, and how and when it needs to be submitted?

- Yes
- No
- Don't know

Policies and Procedures (cont.)

* **7. Do these procedures address the following?**

	Yes	No	Don't know
Selection criteria regarding what to submit/accept.	j	j	j
Requirements regarding standard formats.	j	j	j
Information about copyrights of data submitted.	j	j	j
The way in which data is submitted.	j	j	j
Responsibilities for data storage and management.	j	j	j
Liability when data is lost or affected.	j	j	j

* 8. Do you have policies in place which require those who submit data to show who has previously enhanced, annotated or had access to the data?

- j⋒ Yes
- j⋒ No
- j⋒ Don't Know

* 9. Do you have policies and an infrastructure to guarantee that data are properly managed and maintained to ensure continued access and usability?

- j⋒ Yes
- j⋒ No
- j⋒ Don't Know

* 10. Do you have security protocols that protect stored data from unauthorized modification, damage or deletion?

- j⋒ Yes
- j⋒ No
- j⋒ Don't Know

* 11. Do you arrange training for/give advice to researchers/research institutes to ensure good practice when they submit digital research data?

- j⋒ Yes
- j⋒ No
- j⋒ Don't Know

* 12. Do you have procedures to determine ownership and for identifying and managing data rights?

- j⋒ Yes
- j⋒ No
- j⋒ Don't Know

* 13. Do you arrange training for/give advice to researchers/research institutes to ensure good practice when they submit digital research data?

- j⋒ Yes
- j⋒ No
- j⋒ Don't Know

* 14. Do you have other arrangements/means for promoting good practice?

ʝ෮ No

ʝ෮ Don't know

ʝ෮ Yes (please specify)

Operational Resources

Preservation means that data is specifically curated to be re-usable in the long term. The following questions address the preservation of digital research data.

* 15. Please indicate which of the following digital research data you PRESERVE for the long term: (multiple answers possible)

- ⊝ Standard office documents (text documents, spreadsheets, presentations)

- ⊝ Network-based data (web sites, e-mail, chat history, etc.)

- ⊝ Databases (DBASE, MS Access, Oracle, MySQL, etc.)

- ⊝ Images (JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.)

- ⊝ Structured graphics (CAD, CAM, 3D, VRML, etc.)

- ⊝ Audiovisual (multimedia) data (WAVE, MP3, MP4, Flash, etc.)

- ⊝ Scientific and statistical data formats (SPSS, FITS, GIS, etc.)

- ⊝ Raw data (device specific output)

- ⊝ Plain text (TXT in various encodings)

- ⊝ Structured text (XML, SGML, etc.)

- ⊝ Archived data (ZIP, RAR, JAR, etc.)

- ⊝ Software applications (modelling tools, editors, IDE, compilers, etc.)

- ⊝ Source code (scripting, Java, C, C++, Fortran, etc.)

- ⊝ Configuration data (parameter settings, logs, library files)

- ⊝ Other, please specify below.

  [   ]

* 16. Does your organisation have any of the following preservation strategies in place?

- ⊝ Migration (periodic conversions of file formats to popular formats of today)

- ⊝ Normalisation (conversion of all data to one standardized file format sustainable over time)

- ⊝ Emulation (no conversions of the original data but capturing the original software used to access it)

- ⊝ Outsourced to a third-party service

- ⊝ No preservation strategies in place

- ⊝ Don't know

- ⊝ Other (please specify)

  [   ]

- ⊝

* 17. Do you think your current infrastructures will scale with future requirements?

　ｊ∩　Yes

　ｊ∩　No

　ｊ∩　Don't Know

* 18. Do you share infrastructures with other organizations?

　ｊ∩　Yes

　ｊ∩　No

　ｊ∩　Don't Know

19. How are (or should) these shared infrastructure (be) managed?

5

6

Operational Resources (cont.)

* 20. Do you think there will be a need for sharing resources?

- ⃝ Yes

- ⃝ No

- ⃝ Don't Know

* 21. Do the tools and infrastructure available to you suffice for the digital preservation objectives you have to achieve?

- ⃝ Yes

- ⃝ No

- ⃝ Don't know

Access

The following questions address if and how data, stored at your organisation, can be accessed by external users.

* **22. If data stored at your facility can be accessed by third party users then do you have procedures in place arranging the level of access (e.g. read-only, annotation rights)?**

    ○ Yes

    ○ No

    ○ N/A

    ○ Don't Know

* **23. Do you have procedures determining what is made available in conformity with legal, ethical and other considerations?**

    ○ Yes

    ○ No

    ○ Don't Know

Access (cont.)

* 24. Are the rules and regulations with regard to access laid down in a 'terms and conditions' statement which has to be accepted by users before access is granted to them?

ʝ Yes

ʝ No

ʝ Don't Know

Access (cont.)

* 25. Is it possible for users of the data stored at your organisation to link to that data when referencing it in a journal?

j̈n Yes

j̈n No

j̈n Don't Know

* 26. Have procedures been established for the peer review of data to be made available to others?

j̈n Yes

j̈n No

j̈n Don't know

Access (cont.)

27. Is it made clear to users that the data has been subject to peer review?

j⌐ Yes

j⌐ No

j⌐ Don't know

Financial Resources

* 28. Is funding for digital preservation currently an issue for you? Will it be an issue in 5 years or 10+ years time?

	Yes	No	Don't know
Current	jn	jn	jn
5 years	jn	jn	jn
10+ years	jn	jn	jn

* 29. What are your main sources of funding?

- ∈ Research funder
- ∈ Research community (researchers collectively)
- ∈ Government (national funding)
- ∈ European Union (European funding)
- ∈ Commercial sponsoring
- ∈ Don't know
- ∈ Other (please specify):

* 30. In your opinion, who is responsible for preservation of digital research data? (multiple answers possible)

- ∈ Research funder
- ∈ Researcher
- ∈ The researcher's institute
- ∈ Research community (researchers collectively)
- ∈ Publisher
- ∈ A coalition of publishers
- ∈ A specialised external organisation (Portico, CLOCKSS, etc.)
- ∈ National library
- ∈ Research library
- ∈ Government
- ∈ European Union
- ∈ Don't know
- ∈ Other (please specify)

* 31. Who, in your opinion, should pay for preservation of digital research data? (multiple answers possible)

- ∈ Research funder

- ∈ Researcher

- ∈ The researcher's institute

- ∈ Research community (researchers collectively)

- ∈ Publisher

- ∈ A coalition of publishers

- ∈ A specialised external organisation (Portico, CLOCKSS, etc.)

- ∈ National library

- ∈ Research library

- ∈ Government (national funding)

- ∈ European Union (European funding)

- ∈ Commercial organisation (profit data centre)

- ∈ Don't know

- ∈ Other (please specify)

* 32. Please indicate how important you think the following reasons for preserving digital data are:

	Very important	Important	Slightly important	Not important
If research is publicly funded, the results should become public property and therefore properly preserved.	j∩	j∩	j∩	j∩
It will stimulate the advancement of science (new research can build on existing knowledge).	j∩	j∩	j∩	j∩
It may serve validation purposes in the future.	j∩	j∩	j∩	j∩
It allows for re-analysis of existing data.	j∩	j∩	j∩	j∩
It may stimulate inter-disciplinary collaborations.	j∩	j∩	j∩	j∩
It potentially has economic value.	j∩	j∩	j∩	j∩
It is unique.	j∩	j∩	j∩	j∩

* 33. How important do you regard the following threats over the next 10 years?

	Very Important	Important	Slightly Important	Not Important	Don't Know
Users may be unable to understand or use the data e.g. the semantics, format or algorithms involved	j∩	j∩	j∩	j∩	j∩
Lack of sustainable hardware, software or support of computer environment may make the information inaccessible	j∩	j∩	j∩	j∩	j∩
Evidence may be lost because the origin and authenticity of the data may be uncertain	j∩	j∩	j∩	j∩	j∩
Access and use restrictions (e.g. Digital Rights Management) may not be respected in the future	j∩	j∩	j∩	j∩	j∩
Loss of ability to identify the location of data	j∩	j∩	j∩	j∩	j∩
The current custodian of the data, whether an organisation or project, may cease to exist at some point in the future	j∩	j∩	j∩	j∩	j∩
The ones we trust to look after the digital holdings may let us down	j∩	j∩	j∩	j∩	j∩

* 34. Do you think that an international infrastructure for data preservation and access should be built to help guard against some of these threats?

     jn   Yes

     jn   No

     jn   Don't Know

35. What do you think this infrastructure should look like?

Reasons for and threats to preservation (cont.)

* 36. Apart from an infrastructure, what do you think is needed to guarantee that valuable digital research data is preserved for access and use in the future? (multiple answers possible)

- ∈ Training

- ∈ More knowledge/expertise

- ∈ More resources (financial and operational)

- ∈ More digital repositories/archives

- ∈ Nothing else

- ∈ Other (please specify)

✳ 37. Regarding current movements in scholarly communication, which of the following scenarios do you think is likely to happen:

ʄ The publication process will not change much. The traditional publication model will not alter in terms of peer review and dissemination via research journals. Publishers and their journal programmes will remain the central player in the certification of quality and the branding excellence.

ʄ Similar as the previous option, except that open access journals will become mainstream via the author-pays model, and most articles are also accessible via institutional repositories. Paid-for subscriptions will each the end of their lifetime.

ʄ A hybrid model, combining subscription-based journals and open access journals, while the journal model remains dominant.

ʄ Most research results will be Open Access and available for free via institutional repositories, leading to fast commoditization of research information. Publishers and their journal programmes will be under strain and face difficulties to keep their business models healthy because of the availability of so much free information. As a consequence, many smaller and more specialised journals will disappear, publishers will face severe challenges to survive in the journal market.

ʄ Publishers will become news aggregators, selecting and combining latest research outcomes residing in institutional repositories (overlaid journals).

ʄ The publisher will become a provider of information management services to free and open access content. Content such as publications and research data are commoditized and freely available, users can have (open) access to it. Publishers offer search tools and evaluation services for the free content as a paid-for service for better findability, retrievability and usability of the free content.

ʄ Other (please specify)

[   ]

✳ 38. Do you think the following initiatives would be useful for raising the level of knowledge about preservation of digital research data?

	Very useful	Useful	Slightly useful	Not useful	Don't know
Workshops on preservation of digital material	ʄ	ʄ	ʄ	ʄ	ʄ
User-oriented training sessions on digital preservation	ʄ	ʄ	ʄ	ʄ	ʄ
International knowledge platform/forum on digital preservation	ʄ	ʄ	ʄ	ʄ	ʄ
Development of guidelines/manuals on how to preserve digital data	ʄ	ʄ	ʄ	ʄ	ʄ

✳ 39. Would you be prepared to be interviewed by telephone after completion of this survey in order to help to shape the European preservation infrastructure?

ʄ No

ʄ Yes. Please provide your email address.

[   ]

* 40. May we share the data collected in this survey with closely related projects in the field of digital preservation?

  j∩ No

  j∩ Yes

* 41. Are there others to whom we should send this questionnaire?

  j∩ No

  j∩ Yes. Please provide the email address of a person to whom we may send this questionnaire.

  [   ]

* 42. Do you want to be informed about the results of this survey?

  j∩ No

  j∩ Yes

  Please provide your email address, unless provided above:

  [   ]

Thank you!

Thank you very much for answering our questions. With your answers we are able to outline the current state of preservation and access of digital data in science.

If you entered your email address we will inform you about the final results of this project. In any case, you can visit our website www.parse-insight.eu.

Please click 'done' to finish this survey.