海外事例に見る オープンサイエンス推進 における図書館の役割

鹿児島大学附属図書館 西薗由依



(本発表の位置付け)

研究データ管理(RDM)をテーマとした国際会議への参加報告を基軸とします。

- 10th International Digital Curation Conference (2015.2)
 http://www.dcc.ac.uk/events/idcc15/
- DCC/Jisc合同WS "Research Data Management readiness" (2015.2)

http://www.dcc.ac.uk/events/other-dcc-events/RDM-readiness

COAR-SPARC Conference 2015 (2015.4)

https://www.coar-repositories.org/community/events/annual-meeting-2015/

Cf.) 報告書: http://id.nii.ac.jp/1280/00000083/ http://id.nii.ac.jp/1280/00000107/

機関におけるRDM展開の推進要因

英国の研究志向大学への調査(DCC, 2014)(複数回答可)

要因	ラッセルグループ (N=24)	その他の研究志向大学 (N=24)
研究支援拡大	67%	40%
学術上の利益	54%	40%
財務上の責務	41%	40%
研究の公正性	62%	46%
RCUKのデータポリシー	97%	94%
政府のオープンデータ政策	59%	43%
EU Horizon 2020のRDMポリシー	62%	43%
学術雑誌のデータ登録要請ポリシー	36%	20%
その他	67%	40%

RCUKのポリシー

Full Coverage Partial Coverage No Coverage

	Policy Cov	erage			Policy Sti	pulations		Support Provided			
Research Funders	Published outputs	Data	Time limits	Data plan	Access/ sharing	Long- term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	•	•	•	•	•	•	0	•	0	•	•
BBSRC	•	•	•	•	•	•	•	•	•	•	•
CRUK	•	•	•	•	•	•	•	0	•	0	0
EPSRC	•		•	•	•	•	•	•	0	0	•
ESRC	•	•	•	•	•	•	•	•	•	•	•
MRC	•	•	•	•	•	•	0	•	•	0	•
NERC	•	•	•	•	•	•	•	•	•	•	0
STFC	•	•	•	•	•	•	•	•	•	•	•
Wellcome Trust	•	•	•	•	•	•	•	•	•	•	•

http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies

EPSRCポリシー(2015.5発効)

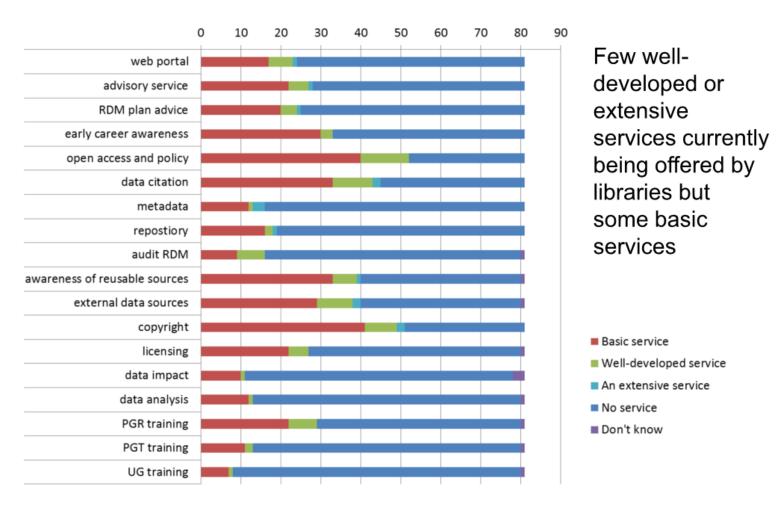
- 機関がRDMの第一義的な 責任を負う(←EPSRC被助成 者の多寡に関わらず、対応 する必要)
- データライフサイクル全体に 対して効率的なデータ管理 を提供
- メタデータ、DOI付与
- 最低10年間の保存
- 非デジタルデータも、アクセス要求へ応える必要
- モニタリング実施、ペナル ティあり etc.

Expectations

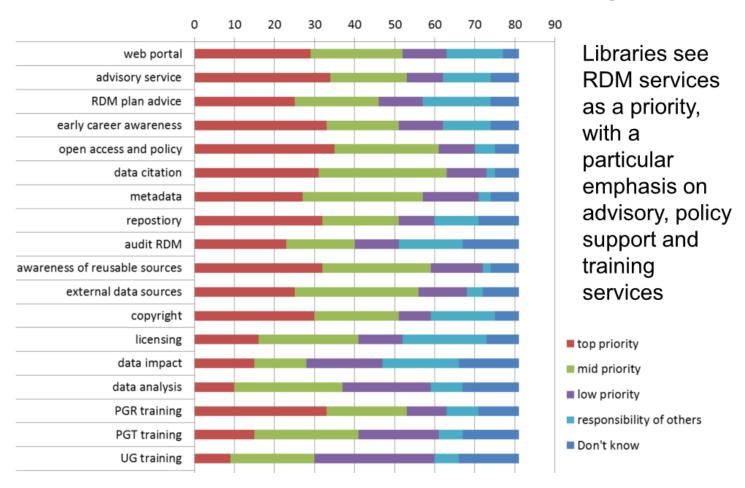
EPSRC has the following clear expectations of organisations in receipt of EPSRC research funding (clarification and guidance on the interpretations of these expectations (PDF 135KB) &):

- Research organisations will promote internal awareness of these principles and expectations and ensure that
 their researchers and research students have a general awareness of the regulatory environment and of the
 available exemptions which may be used, should the need arise, to justify the withholding of research data.
- Published research papers should include a short statement describing how and on what terms any supporting research data may be accessed.
- iii. Each research organisation will have specific policies and associated processes to maintain effective internal awareness of their publicly-funded research data holdings and of requests by third parties to access such data; all of their researchers or research students funded by <u>EPSRC</u> will be required to comply with research organisation policies in this area or, in exceptional circumstances, to provide justification of why this is not possible.
- iv. Publicly-funded research data that is not generated in digital format will be stored in a manner to facilitate it being shared in the event of a valid request for access to the data being received (this expectation could be satisfied by implementing a policy to convert and store such data in digital format in a timely manner).
- v. Research organisations will ensure that appropriately structured metadata describing the research data they hold is published (normally within 12 months of the data being generated) and made freely accessible on the internet; in each case the metadata must be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it. Where the research data referred to in the metadata is a digital object it is expected that the metadata will include use of a robust digital object identifier (For example as available through the DataCite organisation \(\varphi \).
- vi. Where access to the data is restricted the published metadata should also give the reason and summarise the conditions which must be satisfied for access to be granted. For example 'commercially confidential' data, in which a business organisation has a legitimate interest, might be made available to others subject to a suitable legally enforceable non-disclosure agreement.
- vii. Research organisations will ensure that <u>EPSRC</u>-funded research data is securely preserved for a minimum of 10 years from the date that any researcher 'privileged access' period expires or, if others have accessed the data, from last date on which access to the data was requested by a third party; all reasonable steps will be taken to ensure that publicly-funded data is not held in any jurisdiction where the available legal safeguards provide lower levels of protection than are available in the UK.
- viii. Research organisations will ensure that effective data curation is provided throughout the full data lifecycle, with 'data curation' and 'data lifecycle' being as defined by the Digital Curation Centre. The full range of responsibilities associated with data curation over the data lifecycle will be clearly allocated within the research organisation, and where research data is subject to restricted access the research organisation will implement and manage appropriate security controls; research organisations will particularly ensure that the quality assurance of their data curation processes is a specifically assigned responsibility.
- ix. Research organisations will ensure adequate resources are provided to support the curation of publicly-funded research data; these resources will be allocated from within their existing public funding streams, whether received from Research Councils as direct or indirect support for specific projects or from higher education funding councils as block grants.

RDMへの取り組み状況(2012英)



Priorities for the next 3 years



RDMへの取り組み状況(2014英)

英国の研究志向大学への調査(DCC, 2014)

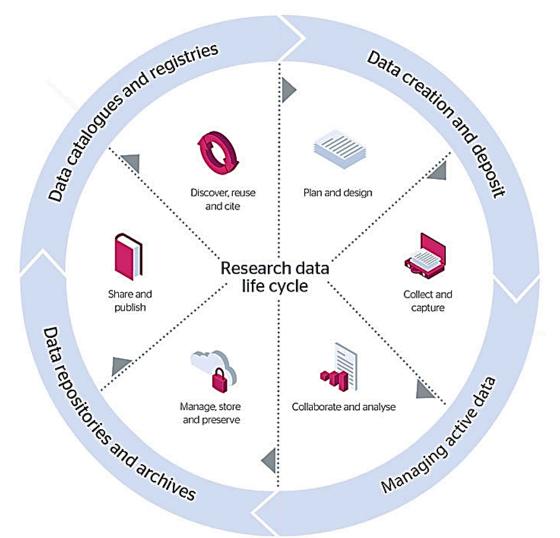
RDMポリシー制定状況	ラッセルグループ (N=24)	その他の研究志向大学 (N=24)
上層部の公式対応は未着手	4%	8%
委員会での要件審議・案策定承認	25%	17%
案を策定し審議中	4%	17%
裁定に向けた最終段階	13%	33%
制定済	54%	25%

長期間(10年以上)のアクセス保証 への対応	ラッセルグループ (N=24)	その他の研究志向大学 (N=24)
対応済	18%	3%
1年以内に対応予定	44%	40%
1年後以降に対応予定	31%	43%
わからない	8%	14%

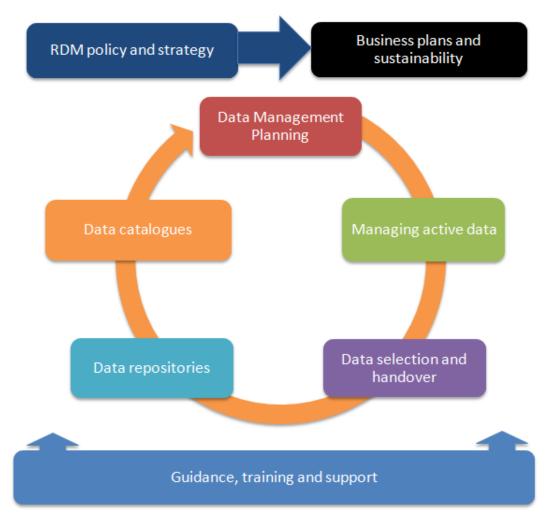
助成団体ポリシーの要件

- 欧州経済領域(EEA)内での研究データ管理・保存 の義務付け
- データライフサイクル全体について、効率的なデータ管理を要求
- 要件を満たすインフラの整備が必要

研究データライフサイクル



RDMサービス構成要素



RDMサービスの分類



Data Management Support

データ管理計画

Data Management Planning

アクティブデータインフラ

Active Data Infrastructure データ受託

Data Stewardship



研究前 Before research 研究中 During research 研究後 After research

機関のRDMポリシー

	Date	Definitions	Support	DMP	Scope	Ownership	External	Criteria	Retain	Ethics
HE Institution in chronological order	Date the policy was adopted	Offers definitions of key RDM terms, i.e. "data", "documentation"	Broadly defines the role institution plays in supporting RDM and outlines RDM responsibility of researchers	Specifies a requirement to complete a DMP (either institutional or funder)	Policy specifies who it covers	Contains a statement on institutional ownership of research data	Contains a statement on the primacy of external funding requirements		Contains a statement on length of time data should be kept	Contains a state the ethical use, data, particular affects potentia
University of Birmingham	2009 October	0	•	0	•	•	•	•	•	•
University of Edinburgh	2011 May	0	•	•	•	0	0	•	0	•
University of Northampton	2011 June	•	•	•	•	0	•	0	•	0
University of Hertfordshire	2011 September	•	•	•	•	•	0	•	•	•
University of Warwick	2011 November	0	•	•	•	0	•	•	•	•
Glyndwr University	2011 December	0	0	0	0	0	0	0	0	•
University of Southampton	2012 February	•	0	0	•	•	•	•	•	•
University of East London	2012 March	0	•	•	0	0	0	•	•	•
Brunel University	2012 March	0	•	•	•	0	ŏ	ŏ	0	•
University of Essex	2012 April	0	•	•	•	0	0	0	0	•
Queen Mary, University of London	2012 June	•	•	•	•	0	•	0	•	•
University of Sheffield	2012 July	•	•	•	•	•	•	Ŏ	0	•
University of Leeds	2012 July	0	•	•	0	0	•	O	0	•
University of Oxford	2012 July	ě	•	•	•	Ö	•	0	•	•
University of the Arts London	2012 October	•	•	0	0	0	0	•	0	0
Goldsmiths University	2013 January	0	•	•	Ö	Ö	•	0	•	
The University of Lancaster	2013 February			•		Ö		0		
University for the Creative Arts	2013 February	0	•	•	0	Ö	0	Ŏ	0	
Oxford Brooks University	2013 February	•			•	•	ŏ	ŏ	•	0
University of Durham	2013 March	0	•	•	0	O	0	•	•	Ĭ
Edge Hill University	2013 May	•			•	Ö	•			
University of Exeter*	2013 June	0			•*	Ö		0	0	
University of Nottingham	2013	•	Ŏ		0	Ö	Ŏ	•	Ö	
University of Manchester	c2013	0	ŏ		•	ě	•	Ō	Ō	
King's College London	2013 June	0	•			Ö		ŏ	Ö	0
University of Lincoln	2013 June	0		0	Ö	Ö	Ö	ŏ	ŏ	ŏ
University of Liverpool	2014 January	•	Ŏ	ŏ		•	ŏ	ŏ		•
University of Bath	2014 April	0	0			o		0		
London School of Hygiene & Tropical Medicine	2014 June	•	•			Ö		•		
University College London	2014 August	·	-			Ö		0	•	Ŏ
University of Dundee	?	0			0	•		Ŏ	0	_
Key								•	•	
•	0	None	None	None	None	None	None	None	None	None
	9	Partialy defined terms or reference to supporting documentation	Institutional support mentioned	Funder required or optional	Staff		Mention of contractual/funding requirements	Non-specified reference, or criteria specified in supporting document	Mentioned but not specified	Mentions of rel legislation on d protection/Fol
	•	Full definitions	Institutional support mentioned and specified	Institutional requirement	Staff and postgraduate students	Yes	Specifies where funder policies take priority	Yes, specified	Specified length	Yes

Horton, L and DCC (2014) 'Overview of UK Institution RDM Policies', Digital Curation Centre. Available at: http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies

機関内の研究データに関する実態の把握

- 効果的・持続的なサービス提供のために
- Data Asset Framework (DAF)
- インタビュー
- 分野ごとのプロファイル蓄積・共有
 - ex.) Data Curation Profiles Toolkit

http://datacurationprofiles.org

Data Curation Profiles Directory

http://docs.lib.purdue.edu/dcp/

- 種別、フォーマット
- 量
- 所有者
- 保管先、保管期間
- メタデータ付与
- バックアップ先、頻度
- 共有
- サポートのニーズ(トレーニングへの関心は高い)
- 資金源、助成団体のポリシーの認知度(往々にして低い)

etc.

データ管理計画(DMP)

Full Coverage Partial Coverage No Coverage

	Policy Cov	erage		Policy Stipulations			Support Provided				
Research Funders	Published outputs	Data	Time limits	Data plan	Access/ sharing	Long- term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC		•	•	•	•	•	0	•	0	•	•
BBSRC	•	•	•	•	•	•	•	•	•	0	•
CRUK	•	•	•	•	•	•	•	0	•	0	0
EPSRC	•	•	•	•	•	•	•	•	0	0	•
ESRC	•	•	•	•	•	•	•	•	•	•	0
MRC	•	•	•	•	•	•	0	•	•	0	•
NERC	•	•	•	•	•	•	•	•	•	•	•
STFC	•	•	•		•	•	•	•	•	•	•
Wellcome Trust	•	•	•		•	•	•	•	•	•	•

http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies

データ管理計画(DMP)作成支援

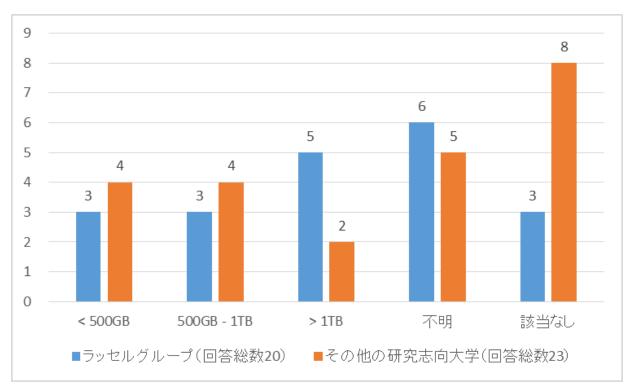
- 研究者のニーズが高い
- 全体的な計画書だけでなく、プロジェクトの構成要素 ごとに詳細な計画書を作成することもあり
- 専門外分野でも対応(サウザンプトン大)
- 個別対応やオンラインツール、ガイダンス
- DMP作成支援
 - →分野に関する知識の習得、人的資源のストック (研究者の持つ知識やスキルの把握、協力的か どうか)
 - →研究者同士をつなぐことも

トレーニング、アドボカシー活動

- 特に若手研究者へ働きかけ
- サウザンプトン大:博士論文の付随データを別途リポジトリに登録することを奨励。さらに永続識別子付与で流通・共有・再利用の促進へ。
- 既存の活動を拡張させるアプローチ: "this is not a project"
- オンラインツール

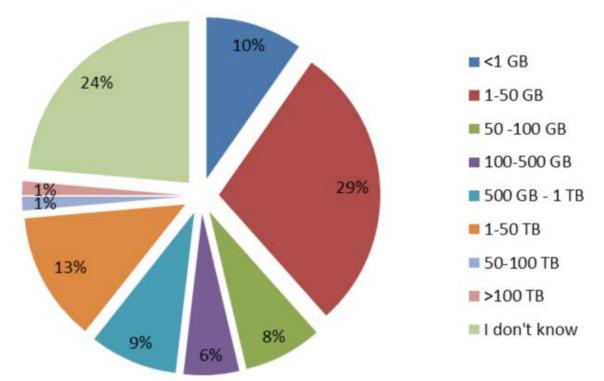
アクティブデータ管理サポート ストレージ

• 英国の研究志向大学への調査(DCC, 2014) 現在無料提供しているストレージの割り当て容量



アクティブデータ管理サポート ストレージ

Cf.) シェフィールド大学のDAF調査結果の例 現在の研究活動中に生成される研究データ量予測



データカタログ/データリポジトリ

- 既存の機関リポジトリとデータリポジトリ
- CRISとデータリポジトリ

Bristol大学: Pure = データカタログ	Lancaster大学:Pure = データリポジトリ
①研究者がPureにメタデータレコードを作成。	①研究者がPureにメタデータレコードを作成。 研究者がデータセットをPureに登録。
②メタデータは自動的に <u>data.bris</u> に取り込まれる。	②研究者が登録完了のメールを受け取る。
③タイトルと予定されるDOIで初期レコードが作られる	③データライブラリアンがメタデータのチェック を行い、DataCiteにDOIの付与を依頼。
④研究者にdata.brisにログインしてデータと関連 付けるよう依頼するメールが送信。	④DOI付与後、データライブラリアンがPureにDOIを登録。
⑤研究者がデータとメタデータを関連付け、公開 依頼。	⑤研究者はデータセットにDOIが付与された旨を メールで受け取る。
⑥データライブラリアンがメタデータのチェック を行い、DataCiteにDOIの付与を依頼。	⑥データが有効になると、Pure portalで公開される。
⑦DOI付与後、データライブラリアンがPureにDOIを登録。	http://id.nii.ac.jp/1280/0000083/

メタデータ

- DataCiteメタデータス キーマへの対応必須項目+推奨/任意 項目
- EPrintsでは研究データリポジトリ用プラグイン
 ReCollect を開発。
 DataCite, INSPIRE, DDI対応。

シンプルかつ一般的な 記述方法を取りつつ、拡 張性を持たせている。

Table 1: DataCite Mandatory Properties

ID	Property	Obligation
1	Identifier (with type sub-property)	М
2	Creator (with name identifier and affiliation sub-properties)	М
3	Title (with optional type sub-properties)	М
4	Publisher	М
5	PublicationYear	М

Table 2: DataCite Recommended and Optional Properties

ID	Property	Obligation
6	Subject (with scheme sub-property)	R
7	Contributor (with type, name identifier, and affiliation sub- properties)	R
8	Date (with type sub-property)	R
9	Language	0
10	ResourceType (with general type description sub-property)	R
11	AlternateIdentifier (with type sub-property)	0
12	RelatedIdentifier (with type and relation type sub-properties)	R
13	Size	0
14	Format	О
15	Version	0
16	Rights	0
17	Description (with type sub-property)	R
18	GeoLocation (with point and box sub-properties)	R

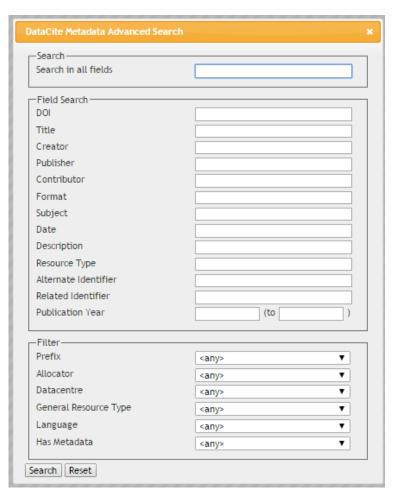
DOI サフィックス

- 'Must be copy-paste, so no mistakes': Bristol
 - 10.5523/BRIS.14S2RMSTMX6AM1K8JKCD3G9UZ3
- · 'Based on an internal ID': UKDS
 - 10.5255/UKDA-SN-7471-1
- 'Sequential': ADS
 - 10.5284/1024569
- Many institutions using a combination of these

DataCite Metadata Search



http://search.datacite.org/ui



分野データリポジトリの位置付け

- 整備状況は分野によりばらつきがある
- 国際的な規模で運用、豊富な人材、各分野の専門家による開発・運営→特定分野のデータ格納先として、機関リポジトリより分野データリポジトリを推奨することも多い
- 分野データリポジトリに登録されたものでも、機関 リポジトリへのメタデータ登録を推奨
- 複数リポジトリ間の効率的な運用の検討

オープンサイエンスとセンシティブデータ

各国のデータ共有ポリシー概要:欧州

_	_								
		***				公開制限データ		機関 保有 のリ	データ 管理計
			基本方針		商業 的· 機密	国家 機密	保存 期間	ポジト リの 有無	画の提出
	RCU	JK Common Principles in Data Policy	・公的助成を受けた研究データは公益財であり、できる限り制限なく、適時にまた知的財産を害することのないよう責任ある方法でオープンに利用できるようにする。	制限有	制限有	明確就なし	NA	NA	規定あり
英国	ポリシー	BBSRC	・BBSRCの支援の結果として生じる研究データは、 次の研究のために、科学界にとって 適時に、また責任をもって提供 できる、 できる限り制限なく 利用できるようにする	NA	制限有	NA	規定あり	Δ	義務化
	/一の策定状況	EPSRC	・EPSRCの助成した研究データは公益のために作られた公益財であり、適時に、責任を持って、自由にかつオープンに使用できるようにする	NA	制限有	制限有	規 定 あり	NA	義務化 されて いない
	況	Wellcome Trust	・できる限り制限なく、適時に責任ある方法で研究コミュニティが研究データを広く利用できるようにする。	制限有	制限有	NA	規 定 あり	Δ	義務化
EU		EU (Horizon 2020) *Horizon2020による助成研究が対象	・正当な利益に反しない限り、論文への記載等の適切な手段により、 できる限り速やかに結果を公開 しなければならない。(結果の保護、さらなる開発を行う場合を除く)	制限有	制限有	制限有	NA	Δ	義務化
		DFG	・研究データの取り扱いに関する原則"Principles for the Handling of Research Data"がドイツ科学 組織連盟により採択。 ・公募要領に、「再利用できる研究データは、その計画を提案に含めること、分野やデータによって、データ共有のコンセプトが異なるため、もっとも効率の良い方法をとる」との記述あり。						

注1: RCUK; Research Council UK, BBSRC; Biological Sciences Research Council, EPSRC; Engineering and Physical Sciences Research Council, DFG; German Research Foundation

注2:「リポジトリの有無」については、O=直営、△=助成、例示、あるいは開発中

対応事例: 英リーズ大学

• 分類スキーマ Unclassified / Confidential / Highly Confidential

センシティブデータに関する、保存場所やアクセス コントロール、共有の指針

Activity	Confidential	Highly Confidential
Storage of data in shared areas of University server	Yes	Only if encrypted
Storage of data in personal area of University server	Yes	Yes
Remote access to the data	Yes, but only via Citrix or other University-approved mechanism	Yes, but only via Citrix or other University-approved mechanism
Storage of data on University-owned laptops or other portable devices	Only on a temporary basis and only if encrypted	Only on a temporary basis and only if encrypted
Storage of data on privately- owned laptops or other portable devices (including memory sticks)	No	No

http://www.dcc.ac.uk/webfm_send/1932

→研究データ管理にどう適用するか

対応事例: 英リーズ大学

- センシティブな研究データの保存場所としての、機関リポジトリの可能性と運用方針を検討
- "Confidential Data"に該当する研究データは、
 EPrintsベースのリポジトリに保存し、アクセスコントロールを行っている
- "Highly Confidential Data"に該当する研究データは、暗号化した上でリポジトリに保存し、アクセスコントロールを行うことを検討(暗号化キーの管理やアクセスコントロールについては、別部署で担当)

非デジタルデータ

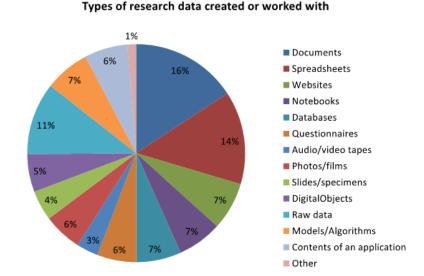
• カタロギング、アクセス保証の要求

Expectations

※抜粋 https://www.epsrc.ac.uk/about/standards/researchdata/expectations/

iv. Publicly-funded research data that is not generated in digital format will be stored in a manner to facilitate it being shared in the event of a valid request for access to the data being received (this expectation could be satisfied by implementing a policy to convert and store such data in digital format in a timely manner).

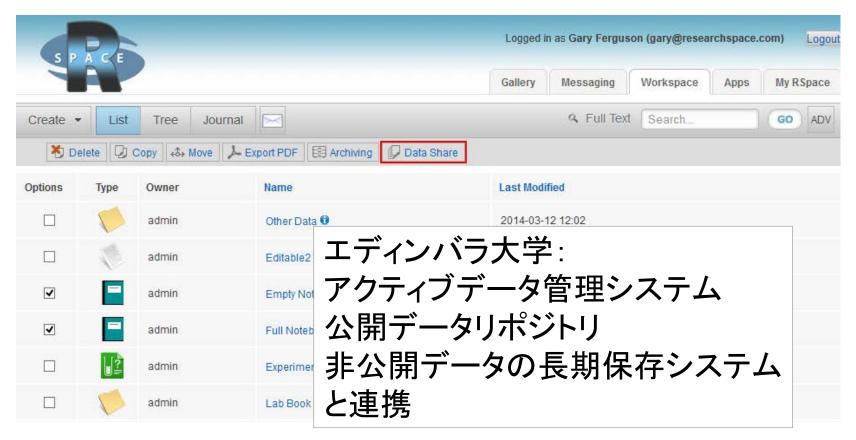
ノッティンガム大学のDAF報告書で挙げられている学内の非デジタルデータの例:手稿、建築図面、実験ノート、質問紙 etc.



http://admire.jiscinvolve.org/wp/files/2013/02/ADMIRe-Survey-Results-and-Analysis-2013.pdf

実験ノートにおける取り組み

• 電子実験ノートシステム(ELN)とRDMとの統合



図書館の役割

英国の研究志向大学への調査(DCC, 2014)「RDM対応の主導的な役割を果たしている部署は?」(複数回答可)

部署	ラッセルグループ (N=24)	その他の研究志向大学 (N=24)
研究支援部署•広報部署	54%	57%
図書館	87%	86%
IT部署	77%	49%
その他	26%	9%

http://dx.doi.org/10.5281/zenodo.10711

図書館の役割

学内連携の有無

部署	ラッセルグループ (N=24)	その他の研究志向大学 (N=24)
研究支援部署•広報部署	54%	57%
(内訳)	単独 0%, 連携 54%	単独 6%, 連携 51%
図書館	87%	86%
(内訳)	単独 15%, 連携 72%	単独 17%, 連携 69%
IT部署	77%	49%
(内訳)	単独 0%, 連携 77%	単独 6%, 連携 43%
その他	26%	9%
(内訳)	単独 5%, 連携 21%	単独 0%, 連携 9%

http://dx.doi.org/10.5281/zenodo.10711

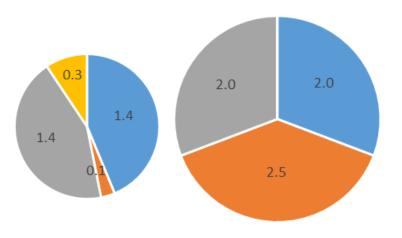
RDMサポートスタッフ体制

英国の研究志向大学への調査(DCC, 2014)

ラッセルグループ

調査時点 3.1 FTE

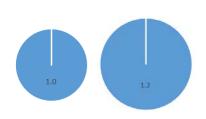
→ 2015.5時点(予想) 6.5 FTE



その他の研究志向大学

調査時点 1.0 FTE

→ 2015.5時点(予想) 1.2 FTE



- 図書館

- 研究支援部署 IT部署 学部、研究者等
- ※数値は中間値

RDMコミュニティ

- プロジェクトを通じた機関間コミュニティ形成
- 地域コンソーシアムによる情報共有や課題検討
- ユーザグループによる課題解決の取り組み ex.) PureやEPrintsにおける研究データ対応
- 大陸を超えた連携例も 南アケープ・ペニンシュラエ業大学 一独 ミュンヘンエ科大学の事例

https://www.coar-repositories.org/files/3_Chiware-COAR-SPARC-Presentation.pdf

DCC

図書館の役割

