The Stewardship Gap Project: Relationships Between Stewardship Gap Areas
January 1, 2017

The following table provides extracts of text or paraphrased or combined ideas from articles, reports, and other works relevant to the stewardship gap, as well as citations to their sources. The table indicates the gap area that the text relates to and the influences between gaps interpreted from the text. A bibliography follows at the end of the document, and is also linked to from http://stewardshipgap.net.

Notes:

1. “CMP” stands for Curation, Management, and Preservation
2. "Other" designations (e.g., Knolwedge: Other) refer to a miscellaneous sub-category of the main gap area.

Relevant text and ideas	Gap Area	Gap Influence	Source
“Reusable” and reproducible are rarely feasible in any absolute sense	Reuse: Difficulty of Reusing Data	CMP->Reuse	Borgman 2012
A core challenge for staffing is that most of the labor involved in data management is invisible work. Information professionals, software engineers, scientific programmers, instrument builders, and other technical experts underpin the foundations of scholarship. These skills often are undervalued and career paths are unclear. It is difficult to recruit talented people under “soft money” contracts that lack job security and routes for promotion. Investments in the human infrastructure will be crucial to the success of keeping and exploiting research data.	Funding: Lack of Funding	Culture->Human Resources	Borgman 2015; Ember and Hanisch 2013; Lynch 2008
A core challenge for staffing is that most of the labor involved in data management is invisible work. Information professionals, software engineers, scientific programmers, instrument builders, and other technical experts underpin the foundations of scholarship. These skills often are undervalued and career paths are unclear. It is difficult to recruit talented people under “soft money” contracts that lack job security and routes for promotion. Investments in the human infrastructure will be crucial to the success of keeping and exploiting research data.	Funding: Lack of Funding	Funding->Human Resources	Borgman 2015; Ember and Hanisch 2013; Lynch 2008

Relevant text and ideas	Gap Area	Gap Influence	Source
a key difficulty in assessing digital preservation costs lies in distinguishing between the preservation stage and other stages, given that the preservation costs may be impacted by decisions made much earlier in the lifecycle of the digital resource, perhaps even at the time of its creation	Knowledge: Costs of Stewardship	Knowledge->Funding	Eakin et al. n.d.; Hendley 1998, quoted in Eakin et al; Russell and Weinberger 2000, quoted in Eakin et al.
a tension between two needs that cannot easily be reconciled: providing production systems for the current generation of high-end users, and moving to the next highest level of computing capability	Infrastructure and Tools: Difficulty meeting generalized and special needs	Planning->Infrastructure and Tools	Atkins et al. 2003
Access to data (sharing; benefits; value proposition to not withhold)	Legal and Policy: Incentives	Culture->Sharing and Access	Berman 2008; Borgman 2015; Science Staff 2011
Achieving balance between domain scientists and computer scientists in an organization [about cyberinfrastructure]	Human Resources: Uneven Distribution of Skills	Human Resources->CMP	Atkins et al. 2003
Achieving balance between domain scientists and computer scientists in an organization [about cyberinfrastructure]	Human Resources: Uneven Distribution of Skills	Human Resources->Infrastructure and Tools	Atkins et al. 2003
An investigator may be part of multiple, overlapping communities of interest, each of which may have different notions of what are data and different data practices	Culture: Data Definition	Culture->Sharing and Access	Borgman 2012; Esanu et al. 2004; Read et al. 2015
and maintaining a sufficiently long time horizon for access in the face of continual change in data definitions, digital data media and formats, and hardware and software obsolescence	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->CMP	Burda and Teuteberg 2013; Esanu et al. 2004
and maintaining a sufficiently long time horizon for access in the face of continual change in data definitions, digital data media and formats, and hardware and software obsolescence	Sustainability Planning: Planning for Dynamic and	Planning->Infrastructure and Tools	Burda and Teuteberg 2013; Esanu et al. 2004

Relevant text and ideas	Gap Area	Gap Influence	Source
	Adaptable Infrastructure		
And yet the system of academic research achievement does not yet recognize or reward researchers for sharing their data	Legal and Policy: Incentives	Legal/Policy->CMP	Thompson Reuters 2013
And yet the system of academic research achievement does not yet recognize or reward researchers for sharing their data	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	Thompson Reuters 2013
appropriate retention timing is hard to identify	Knowledge: Value	Knowledge->CMP	Borgman 2015; Esanu et al. 2004
Archiving requirements no longer apply solely to post-research publication activities. Properly documenting data for long-term preservation and access must become part of the daily practice of scientists.	Culture: Other	CMP->Access	Esanu et al. 2004
Astronomers reported varying practices for preserving these types of data [secondary data, data on backup tape], usually storing only what they expected to use in the future. None reported having preservation resources for their work outside their group of collaborators	Funding: Lack of Funding	Funding->CMP	STC 2007
Astronomers reported varying practices for preserving these types of data [secondary data, data on backup tape], usually storing only what they expected to use in the future. None reported having preservation resources for their work outside their group of collaborators	Funding: Lack of Funding	Funding->Culture	Wynholds et al. 2011
authors don’t want to relinquish control of large databases that can be mined given the large amount of time to put them together	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Human Resources->Legal/Policy	NRC 2003
authors don’t want to relinquish control of large databases that can be mined given the large amount of time to put them together	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Human Resources->Sharing	NRC 2003

Relevant text and ideas	Gap Area	Gap Influence	Source
building both free universal (general use) specialized cyberinfrastructure	Infrastructure and Tools: Difficulty meeting generalized and special needs	Planning->Infrastructure and Tools	Cummings et al. 2008; Lynch 2008
building the institutional commitments to make them into resource or reference collections is harder.	Commitment: Lack of Commitment	Commitment->Human Resources	Borgman 2015
building the institutional commitments to make them into resource or reference collections is harder.	Commitment: Lack of Commitment	Commitment->Infrastructure and Tools	Borgman 2015
But today, no media, hardware or software exists that can reasonably assure long-term accessibility to digital assets. A dynamic approach that anticipates failures and obsolescence will be essential.	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->Infrastructure and Tools	Rosenthal 2005
CENS researchers are more willing to invest effort in documenting data for people they know and trust	Culture: Research and Development	Culture->CMP	Wallis et al. 2013
Certain development challenges also exist, including the tension between customization and shared infrastructure as well as the deployment, maintenance, and support of infrastructure.	Infrastructure and Tools: Difficulty meeting generalized and special needs	Planning->Infrastructure and Tools	Cummings et al. 2008
certification can be very labor intensive for archives already stretched thin	Human Resources: Time	Human Resources->CMP	Edwards et al. 2013
Challenge of “Interpretable”: presumes sufficient expertise to assess the integrity of the data and to grasp their meaning and adequate documentation of the context of the data creation, processing, and provenance.	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Borgman 2012
Challenge of “Interpretable”: presumes sufficient expertise to assess the integrity of the data and to grasp their meaning and adequate	Reuse: Difficulty of Reusing Data	Human Resources->Reuse	Borgman 2012

Relevant text and ideas	Gap Area	Gap Influence	Source
documentation of the context of the data creation, processing, and provenance.			
challenge of cyberinfrastructure is primarily a challenge to our own academic cultures	Culture: Research and Development	Culture->Human Resources	NSF 2007
challenge of cyberinfrastructure is primarily a challenge to our own academic cultures	Culture: Research and Development	Culture->Infrastructure and Tools	NSF 2007
challenge of establishing trustworthiness of data for reuse	Reuse: Difficulty of Reusing Data	CMP->Reuse	Rauber 2012; Wynholds et al. 2011
Challenge of free riders for common-pool resources	Culture: Other	Funding->Infrastructure and Tools	Borgman 2015
Challenge of free riders for common-pool resources	Culture: Other	Planning->Funding	Borgman 2015
Challenge of sustainability for common-pool resources			
challenge to develop knowledge infrastructures that serve the diversity of ideas, questions, methods, and resources that each contributes to scholarship	Infrastructure and Tools: Lack of Infrastructure	Infrastructure and Tools->Scholarship	Borgman 2015
challenges in data organization; poor organization	CMP: Difficulty Managing Data for Reuse	CMP->Reuse (by implication)	Berman 2008; Science Staff 2011
Challenges in linking; the robustness of knowledge infrastructures to link publications and data for discoverability depends on investments in information professionals to do the requisite work	Human Resources: Other	Human Resources->CMP	Borgman 2015
Challenges in linking; the robustness of knowledge infrastructures to link publications and data for discoverability depends on investments in information professionals to do the requisite work	Human Resources: Other	Human Resources->Infrastructure and Tools	Borgman 2015
challenges in preserving interpretive information - protocols, codebooks, software, specimens, metadata, standards, and so on (size of space available to researchers)	Funding: Lack of Funding	Funding->CMP	Borgman 2015
challenges in preserving interpretive information - protocols, codebooks, software, specimens, metadata, standards, and so on (size of space available to researchers)	Funding: Lack of Funding	Infrastructure and Tools->CMP	Borgman 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
Challenges of finding a place in shifting landscape where research universities are looking more like corporations.			
Challenges of implementing formal ("differential privacy") and in(formal methods of allowing data to be queried while protecting privacy	Reuse: Difficulty of Reusing Data	Legal/Policy->Reuse	Marchionini, Lee, Bowden, 2012
challenges of providing uniform access to a broad variety of research outputs—the technical and computational limitations in making the data available, searchable and retrievable	Infrastructure and Tools: Other	Infrastructure and Tools->Discovery	Thompson Reuters 2013
challenges of using computational approaches: loss of context, loss of provenance, and the difficulty of maintaining the relationships among objects that are necessary to interpret the results	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Borgman 2015
challenges of using computational approaches: loss of context, loss of provenance, and the difficulty of maintaining the relationships among objects that are necessary to interpret the results	Reuse: Difficulty of Reusing Data	CMP->Reuse	Borgman 2015
clarify challenges of providing tools to enable reuse of those data	Knowledge: Challenges of Enabling Reuse	Knowledge->Infrastructure and Tools	Brown et al. 2015
clarify challenges of providing tools to enable reuse of those data	Knowledge: Challenges of Enabling Reuse	Infrastructure and Tools->Reuse	Brown et al. 2015
clarify challenges of providing tools to enable reuse of those data	Knowledge: Challenges of Enabling Reuse	Knowledge->Reuse	Brown et al. 2015
Clear legal frameworks for the sharing of publications and reuse of data sets are needed at the national and international levels.	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Reuse	OECD 2015
Clear legal frameworks for the sharing of publications and reuse of data sets are needed at the national and international levels.	Legal/Policy: Deficiencies That Inhibit	Legal/Policy->Sharing	OECD 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
	Stewardship, Access, and Use		
Collaboration for preservation: In addition, as archives, repositories, and data centers increasingly provide services that integrate data from multiple sources, the need for sustainable capabilities will become even more pressing, as a failure by one pro	Collaboration: Need for Collaboration	Services->Planning	Downs and Chen 2013; Ubaldi 2013
commons-based exchange of research data impeded by a perceived ownership of data (right to publish first) and need for control (reflected in fear of misuse)	Culture: Sharing attitudes and practices	Culture->Sharing and Access	Fecher et al. 2015
community objectives associated with digital preservation and the organizational incentives that might motivate their realization (Lavoie 2003)	Legal and Policy: Incentives	Culture->CMP	Bradley 2005, quoted in Eakin et al. n.d.; Eakin et al. n.d.; Lavoie 2003, quoted in Eakin et al., n.d.
community objectives associated with digital preservation and the organizational incentives that might motivate their realization (Lavoie 2003)	Legal and Policy: Incentives	Culture->Legal/Policy	Bradley 2005, quoted in Eakin et al. n.d.; Eakin et al. n.d.; Lavoie 2003, quoted in Eakin et al., n.d.
community objectives associated with digital preservation and the organizational incentives that might motivate their realization (Lavoie 2003)	Legal and Policy: Incentives	Legal/Policy->CMP	Bradley 2005, quoted in Eakin et al. n.d.; Eakin et al., n.d.; Lavoie 2003, quoted in Eakin et al., n.d.
Concern by data managers that the tools and infrastructure available to them will suffice for the digital preservation objectives they have to achieve	Infrastructure and Tools: Lack of Infrastructure	Infrastructure and Tools->CMP	Kuipers and van der Hoeven 2009
Concerns from publishers about the sustainability of data when the current custodian of the data ceases to exist in the future	Sustainability Planning: Lack of	Planning->CMP	Kuipers and van der Hoeven 2009

Relevant text and ideas	Gap Area	Gap Influence	Source
	Strategy and Planning		
Considerable expertise, effort, restructuring, and proprietary software may be necessary to reuse data	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Borgman 2012; Wallis et al. 2013
Considerable human expertise, both technical and scientific, is required to support active use of these science data archives	Human Resources: Other	Human Resources->Reuse	Wynholds et al. 2011
Considerable human expertise, both technical and scientific, is required to support active use of these science data archives	Human Resources: Other	Human Resources->Sharing and Access	Wynholds et al. 2011
contextualization and description time-consuming	Human Resources: Time	Human Resources->CMP	McDonough 2012
costs associated with preserving data longer term	Knowledge: Costs of Stewardship	Knowledge->CMP	Brown et al. 2015
costs associated with preserving data longer term	Support: Costs	Knowledge->Planning	Brown et al. 2015
criteria for identifying data and for sharing them are not yet well understood	Culture: Data Definition	Culture->Sharing	Borgman 2012
criteria (for appraisal) are hard to identify	Knowledge: Value	Knowledge->CMP	Esanu et al. 2004
data cleaning and curation (for example, by developing metadata) is a time-consuming activity that is rarely acknowledged in evaluation mechanisms or grant allocation procedures.	Legal and Policy: Incentives	Human Resources->CMP	OECD 2015
data cleaning and curation (for example, by developing metadata) is a time-consuming activity that is rarely acknowledged in evaluation mechanisms or grant allocation procedures.	Human Resources: Time	Human Resources->CMP	OECD 2015
Data curation faces the challenge that the data must be housed, managed, and made accessible prior to use, but actual uses may not be known until after sizeable investments are made.	Knowledge: Reuse	Reuse->CMP	Wynholds et al. 2012
Data curation faces the challenge that the data must be housed, managed, and made accessible prior to use, but actual uses may not be known until after sizeable investments are made.	Knowledge: Reuse	Reuse->Funding	Wynholds et al. 2012

Relevant text and ideas	Gap Area	Gap Influence	Source
Data curation faces the challenge that the data must be housed, managed, and made accessible prior to use, but actual uses may not be known until after sizeable investments are made.	Knowledge: Reuse	Reuse->Human Resources	Wynholds et al. 2012
Data curation faces the challenge that the data must be housed, managed, and made accessible prior to use, but actual uses may not be known until after sizeable investments are made.	Knowledge: Reuse	Reuse->Infrastructure and Tools	Wynholds et al. 2012
Data curation faces the challenge that the data must be housed, managed, and made accessible prior to use, but actual uses may not be known until after sizeable investments are made.	Knowledge: Reuse	Reuse->Planning	Wynholds et al. 2012
data curation for archiving is funded by the same mechanisms that fund principle aims of a research project	Legal/Policy: Institutional structures and pressures	Funding->CMP	Ember and Hanisch 2013
data curation for archiving is funded by the same mechanisms that fund principle aims of a research project	Legal/Policy: Institutional structures and pressures	Legal/Policy->Funding	Ember and Hanisch 2013
data is not fit for use; it has not undergone a certain level of refinement before being shared (e.g., interpretation of raw data, correction of statistical errors, anonymization); or is out of date by the time it is published.	CMP: Insufficient Data Curation or Management	CMP->Reuse	RECODE project; Sturges et al. 2015
data management efforts inadequate to the amounts of data being generated	CMP: Insufficient Data Curation or Management	Human Resources->CMP	NAS 2009
Data producers have few incentives besides the NIJ archiving requirement to deposit their data, let alone expend extra effort to prepare their data for deposit, and several disincentives are in place	Legal and Policy: Incentives	Legal/Policy->Human resources	Hedstrom and Niu 2008
Data producers have few incentives besides the NIJ archiving requirement to deposit their data, let alone expend extra effort to prepare their data for deposit, and several disincentives are in place	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	Hedstrom and Niu 2008

Relevant text and ideas	Gap Area	Gap Influence	Source
data quality issues can be difficult to monitor, as the process by which data were originally collected and managed may be uncertain	CMP: Quality	Knowledge->CMP	Tenopir et al. 2015
data repositories with significant structural funding feel that core services are valued by the funder and relatively secure, the settlement is reasonable, but leaves little room for innovation, R&D or additional services unless these are specifically con	Funding: Lack of Funding	Funding->Planning	Dillo et al. 2015
data sets are not as easily identified and defined as scholarly research articles.	Culture: Data Definition	Culture->Sharing and Access	OECD 2015
data sharing is even less systematic in domains where few common-pool resources exist	Infrastructure and Tools: Lack of Infrastructure	Infrastructure and Tools->Sharing and Access	Borgman 2015
datasets quickly fall out of synchronization with versions of hardware and software used to create and analyze them	CMP: Difficulty Managing Data for Reuse	CMP->Reuse (by implication)	Borgman 2015
De-funded projects (VO Registry, VAO), also service for large, homogenous datasets, put doubt in the minds of astronomers thinking about doing extra work to share their data	Legal and Policy: Incentives	CMP->Sharing and Access	Pepe et al 2014
De-funded projects (VO Registry, VAO), also service for large, homogenous datasets, put doubt in the minds of astronomers thinking about doing extra work to share their data	Legal and Policy: Incentives	Knowledge->CMP	Pepe et al 2014
De-funded projects (VO Registry, VAO), also service for large, homogenous datasets, put doubt in the minds of astronomers thinking about doing extra work to share their data	Legal and Policy: Incentives	Knowledge->Planning	Pepe et al 2014
De-funded projects (VO Registry, VAO), also service for large, homogenous datasets, put doubt in the minds of astronomers thinking about doing extra work to share their data	Legal and Policy: Incentives	Knowledge->Sharing	Pepe et al 2014
De-funded projects (VO Registry, VAO), also service for large, homogenous datasets, put doubt in the minds of astronomers thinking about doing extra work to share their data	Legal and Policy: Incentives	Planning->CMP	Pepe et al 2014

Relevant text and ideas	Gap Area	Gap Influence	Source
De-funded projects (VO Registry, VAO), also service for large, homogenous datasets, put doubt in the minds of astronomers thinking about doing extra work to share their data	Legal and Policy: Incentives	Planning->Sharing	Pepe et al 2014
dealing with the diversity of sources, formats, and documentation	CMP: Difficulty Managing Data for Reuse	CMP->Reuse (by implication)	Esanu et al. 2004; Ubaldi 2013
designing and staffing of organizations that work with faculty to provide access services; manage data, prepare data for handoff; data mining, make decision about whether services centralized or modeled on individual departments, etc.	Planning: Design and Staffing of Organizations	Human Resources->CMP	Lynch 2008
designing and staffing of organizations that work with faculty to provide access services; manage data, prepare data for handoff; data mining, make decision about whether services centralized or modeled on individual departments, etc.	Planning: Design and Staffing of Organizations	Human Resources->Infrastructure and Tools	Lynch 2008
designing and staffing of organizations that work with faculty to provide access services; manage data, prepare data for handoff; data mining, make decision about whether services centralized or modeled on individual departments, etc.	Planning: Design and Staffing of Organizations	Human Resources->Planning	Lynch 2008
determining how much, and what, digital content to produce (talking about scholarly publishing)	Sustainability Planning: Other	Planning->Amount of data	NSF 2007
development of viable economic models for ensuring that the resources needed for long-term stewardship are put in place, while at the same time addressing the needs of the scientific community and society more generally for open access to scientific data	Sustainability Planning: Business and Economic Models	Funding->Human Resources	Downs and Chen 2013
development of viable economic models for ensuring that the resources needed for long-term stewardship are put in place, while at the same time addressing the needs of the scientific community and society more generally for open access to scientific data	Sustainability Planning: Business and Economic Models	Funding->Infrastructure and Tools	Downs and Chen 2013
development of viable economic models for ensuring that the resources needed for long-term stewardship are put in place, while at the same time	Sustainability Planning: Business	Planning->Funding	Downs and Chen 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
addressing the needs of the scientific community and society more generally for open access to scientific data	and Economic Models		
Differences among archives in nomenclature and taxonomy of data exist within, as well as across, scientific communities	Culture: Data Definition	Culture->CMP	Esanu et al. 2004
differences among the mandates and objectives of individual archives	Culture: Mandates and Objectives	Culture->CMP	Esanu et al. 2004
Differences arise over privacy and confidentiality, from cultural attitudes to ownership and use, in attitudes to intellectual property protection and its limits and exceptions, and because of national security concerns.	Culture: Sharing attitudes and practices Legal/Policy	Legal/Policy->Reuse	NSF 2007; Ubaldi 2013
Different pieces of infrastructure evolve on different time scales (technology fast, universities and publishers more slowly). Startup companies can move quickly but may not be investing for long-term sustainability	Infrastructure and Tools: Different timescales of infrastructure development and maturity	Planning->CMP	Borgman 2015
differing criteria among archives for appraising value and selecting data for preservation	Culture: Identifying what is valuable	Culture->CMP	Esanu et al. 2004
Difficult to convince governments and funding agencies to invest in shared technology, human, services infrastructure	Culture: Priority	Culture->Funding	Borgman 2015
Difficult to convince governments and funding agencies to invest in shared technology, human, services infrastructure	Culture: Priority	Culture->Human Resources	Borgman 2015
Difficult to convince governments and funding agencies to invest in shared technology, human, services infrastructure	Culture: Priority	Culture->Infrastructure and Tools	Borgman 2015
difficult to define the array of skills and qualifications required in data science programs and in data jobs	Knowledge: Skills	Knowledge->CMP	Borgman 2015; Lynch 2008
difficult to define the array of skills and qualifications required in data science programs and in data jobs	Knowledge: Skills	Knowledge->Human resources	Borgman 2015; Lynch 2008
difficult to find and retain skilled professionals	Funding: Lack of Funding	Human Resources->CMP	Borgman 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
difficult to find and retain skilled professionals	Funding: Lack of Funding	Human Resources->Sharing and Access	Borgman 2015
difficulties in estimating data creation/collection costs from depositor surveys	Knowledge: Costs of Stewardship	Knowledge->Planning	Houghton and Gruen 2014
Difficulties of maintaining evidentiary record as "scholarly communications are fragmented, dispersed, and reaggregated in new ways"	CMP: Data Management Fragmented	CMP->CMP	Borgman 2015
difficulty of resolving cultural differences in broad collaboration around infrastructure	Culture: Research and Development	Collaboration->Infastructure	Finholt and Birnholtz
difficulty of resolving cultural differences in broad collaboration around infrastructure	Culture: Research and Development	Culture->Collaboration	Finholt and Birnholtz
difficulty of resolving cultural differences in broad collaboration around infrastructure	Culture: Research and Development	Culture->Infrastructure	Finholt and Birnholtz
Difficulty with definition of "data"	Culture: Data Definition	Culture->Legal/Policy	Sturges et al. 2015
Difficulty with definition of "data"	Culture: Data Definition	Culture->Sharing and Access	Sturges et al. 2015
disparate procedures and metrics among archives for data quality	Culture: Evaluation of Quality	Culture->CMP	Esanu et al. 2004
disparate procedures and metrics among archives for data quality	Culture: Preservation Practices	CMP->Reuse	Esanu et al. 2004
distinguishing between the costs of preservation and the costs of access can be problematic	Knowledge: Costs of Stewardship	Knowledge->Funding	Eakin et al. n.d.
Diversity of scientific data and differing traditions and standards in their treatment are also issues.	Culture: Standards	Culture->Sharing and Access	OECD 2015
documentation of data is often a time-consuming task that is competing with many other activities on the researchers priority list	Human Resources: Time	Human Resources->CMP	Pampel and Dallmeier-Tiessen 2014
Domain repositories face an uncertain future despite the growing demand for data sharing and access	Sustainability Planning: Other	Sharing and Access->Planning	Kupiainen 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
Each field develops its own inscriptions to document, describe, and represent what it considers to be data.	Culture: Standards	Culture->Sharing and Access	Borgman 2015
evidence may be lost because the origin and authenticity of the data may be uncertain	Knowledge: Provenance and Authenticity	Knowledge->CMP	Kuipers and van der Hoeven 2009
few long-term strategies for OGD that take into account context within the society and in the public administration, and comprehensively address technical, economic, social, legal, institutional and implementation-related aspects	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->Funding	Ubaldi 2013
few long-term strategies for OGD that take into account context within the society and in the public administration, and comprehensively address technical, economic, social, legal, institutional and implementation-related aspects	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->Infrastructure and Tools	Ubaldi 2013
few long-term strategies for OGD that take into account context within the society and in the public administration, and comprehensively address technical, economic, social, legal, institutional and implementation-related aspects	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->Legal/Policy	Ubaldi 2013
financial trade-off between creating new data and preserving old data	Legal/Policy: Institutional structures and pressures	Funding->CMP	NAS 2009
find ways of preserving and exchanging research data that are not based on the commodification of research data (make a case for preserving public goods)	Sustainability Planning: Business and Economic Models	Planning->CMP	ARL 2006

Relevant text and ideas	Gap Area	Gap Influence	Source
find ways of preserving and exchanging research data that are not based on the commodification of research data (make a case for preserving public goods)	Sustainability Planning: Business and Economic Models	Planning->Sharing and Access	ARL 2006
Focusing on data publication could fix publication data into units that prevents broader use	CMP: Quality and Usability	Culture->CMP	Borgman 2015
Forming these partnerships with other units at the institution] is listed as the biggest challenge by respondents, and in some cases has led to uncertain roles at the institution-level over which units have primacy over RDM.	Collaboration: Challenges Forming Partnerships	Collaboration->CMP	Fearon et al. 2013
Forming these partnerships with other units at the institution] is listed as the biggest challenge by respondents, and in some cases has led to uncertain roles at the institution-level over which units have primacy over RDM.	Collaboration: Challenges Forming Partnerships	Collaboration->Human Resources	Fearon et al. 2013
Forming these partnerships with other units at the institution] is listed as the biggest challenge by respondents, and in some cases has led to uncertain roles at the institution-level over which units have primacy over RDM.	Collaboration: Challenges Forming Partnerships	Collaboration->Infrastructure and Tools	Fearon et al. 2013
Forming these partnerships with other units at the institution] is listed as the biggest challenge by respondents, and in some cases has led to uncertain roles at the institution-level over which units have primacy over RDM.	Collaboration: Challenges Forming Partnerships	Collaboration->Responsibility	Fearon et al. 2013
fundamental issue in the debate over the sharing of publication related data, information, and materials is whether exceptions should be made to a community standard if the progress of science might be advanced by them	Culture: Standards	Culture->Sharing and Access	NRC 2003
funding for preservation	Funding: Lack of Funding	Funding->CMP	Pepe et al 2014
Generally speaking, the more handcrafted the data collection and the more labor-intensive the postprocessing for interpretation, the less likely	Human Resources: Time	Human Resources->Sharing and Access	Borgman 2012

Relevant text and ideas	Gap Area	Gap Influence	Source
that researchers will share their data (goes on to say generalizations difficult to make)			
Government agencies often have strong technical capabilities, but are subject to political and budgetary pressures and competing priorities	Legal/Policy: Institutional structures and pressures	Culture->CMP	Downs and Chen 2013
guard against financial threats	Sustainability Planning: Business and Economic Models	Planning->Funding	Downs and Chen 2013
hard drives not as reliable as vendors say	Infrastructure and Tools: Other	Infrastructure and Tools->CMP	STC 2007
HEI research data holdings are often distributed amongst schools, departments, groups and individuals. Several Universities have done surveys that show their holdings are very fragmented, for example residing on USB drives or local servers within individual research groups	CMP: Data Management Fragmented	CMP->Knowledge	Addis 2015
How are data used and how are they reused? What types of data are most likely to be reused, by whom, and for what purposes? Answers to these questions are needed to inform information policy and the design of digital libraries to support the capture, management, use, and reuse of data	Knowledge: Reuse	Knowledge->CMP	Wynholds et al. 2012
How are data used and how are they reused? What types of data are most likely to be reused, by whom, and for what purposes? Answers to these questions are needed to inform information policy and the design of digital libraries to support the capture, management, use, and reuse of data	Knowledge: Reuse	Knowledge->Infrastructure and Tools	Wynholds et al. 2012
How are data used and how are they reused? What types of data are most likely to be reused, by whom, and for what purposes? Answers to these questions are needed to inform information policy and the design of	Knowledge: Reuse	Knowledge->Legal/Policy	Wynholds et al. 2012

Relevant text and ideas	Gap Area	Gap Influence	Source
digital libraries to support the capture, management, use, and reuse of data			
How are data used and how are they reused? What types of data are most likely to be reused, by whom, and for what purposes? Answers to these questions are needed to inform information policy and the design of digital libraries to support the capture, management, use, and reuse of data	Knowledge: Reuse	Knowledge->Planning	Wynholds et al. 2012
How will publishers who traditionally regard themselves as content providers negotiate the transition to a role focusing on service provision? How will businesses and systems accustomed to a pay-to-read model adapt to a pay-to-publish model?	Culture: Other	Planning->Funding	Thompson Reuters 2013
If DMPs not accompanied by means of sharing, may not promote sharing and reuse	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->Reuse	Borgman 2015
If DMPs not accompanied by means of sharing, may not promote sharing and reuse	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->Sharing and Access	Borgman 2015
If left to their own devices, most scholars will document their data just well enough to serve the purposes of their immediate and planned research projects	Human Resources: Time	Human Resources->CMP	Borgman 2015
If left to their own devices, most scholars will document their data just well enough to serve the purposes of their immediate and planned research projects	Human Resources: Time	Human Resources->Sharing and Access	Borgman 2015
Imbalance in structure and distribution of knowledge (study of "charismatic megafuna" over other biota that are more fundamental ecological process	Knowledge: Challenges of Enabling Reuse	Knowledge->?	Edwards et al. 2013
Improving discovery	Discovery	CMP->Discovery Infrastructure->Discovery These are mainly implied relationships having to do with curation affect	Borgman 2015; Read et al. 2015; Wallis et al. 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
		on interoperability (tied to discovery but not necessarily) and fact that you need infrastructure to enable discovery.	
in many disciplines there are no standards in which the data can be described	Culture: Standards	Culture->CMP	Pampel and Dallmeier-Tiessen 2014
in practice it is rare to find detailed information regarding preservation commitments offered by individual repositories	Knowledge: Stewardship Commitments	Knowledge-Planning	Lavoie and Malpas 2015
In preparing secondary data products for access, repositories need to address issues of trust, documentation, interoperability, and ongoing value.	CMP: Insufficient Data Curation or Management	CMP->Reuse	Wynholds et al. 2011
in research-funding orgs, funding for infrastructure competes with funding for research without expansion of overall funding	Legal/Policy: Institutional structures and pressures	Legal/Policy->Funding	Berman 2014
In the academic sector, university libraries are natural foci for the stewardship of digital research data. But they need financial support to evolve in this direction at a time when many budgets are being cut.	Funding: Lack of Funding	Funding->Human Resources	Berman et al. 2010
In the academic sector, university libraries are natural foci for the stewardship of digital research data. But they need financial support to evolve in this direction at a time when many budgets are being cut.	Funding: Lack of Funding	Funding->Infrastructure and Tools	Berman et al. 2010
In the present climate, domain repositories have difficulty supporting technical systems and staff expertise required for long-term preservation	Funding: Lack of Funding	Funding->CMP	Ember and Hanisch 2013
In the present climate, domain repositories have difficulty supporting technical systems and staff expertise required for long-term preservation	Funding: Lack of Funding	Funding->Human Resources	Ember and Hanisch 2013
In the present climate, domain repositories have difficulty supporting technical systems and staff expertise required for long-term preservation	Funding: Lack of Funding	Funding->Infrastructure and Tools	Ember and Hanisch 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
Inadequate incentives for scientist participation in data preservation	Legal and Policy: Incentives	Legal/Policy->CMP	ARL 2006
inadequate knowledge of researchers of where to upload/deposit data	Knowledge: Where to Deposit Data	Knowledge->CMP	Sturges et al. 2015; Wallis et al. 2013
inadequate knowledge of researchers of where to upload/deposit data	Knowledge: Where to Deposit Data	Knowledge->Sharing and Access	Sturges et al. 2015; Wallis et al. 2013
inconsistencies and a lack of standardisation in journal policies about data deposit	Legal and Policy: Lack of Consistency and Alignment	Legal/Policy->CMP	Sturges et al. 2015
increased pressure on researchers to share, but little motivation	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	Wallis et al. 2013
infrastructures are long-term initiatives that must navigate short-term funding cycles and political turnover.	Commitment: Duration of Commitment	Commitment->Infrastructure	Borgman 2015; Ember and Hanisch 2013
infrastructures are long-term initiatives that must navigate short-term funding cycles and political turnover.	Commitment: Duration of Commitment	Commitment->Funding	Borgman 2015; Ember and Hanisch 2013
Institutional repositories cannot offer the features and visibility to researchers that disciplinary repositories specializing in data sets for a specific audience are able to provide	Infrastructure and Tools: Difficulty meeting generalized and special needs	Infrastructure and Tools->Discovery	Reeves et al. 2015
Institutional repositories cannot offer the features and visibility to researchers that disciplinary repositories specializing in data sets for a specific audience are able to provide	Infrastructure and Tools: Difficulty meeting generalized and special needs	Infrastructure and Tools->Sharing and Access	Reeves et al. 2015
Insufficient data are released to justify developing a repository, and insufficient demand exists to justify releasing and depositing data	Culture: Demand for Data	Culture->Sharing and Access	Wallis et al. 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
Insufficient data are released to justify developing a repository, and insufficient demand exists to justify releasing and depositing data	Culture: Demand for Data	Sharing->Funding	Wallis et al. 2013
Insufficient data are released to justify developing a repository, and insufficient demand exists to justify releasing and depositing data	Culture: Demand for Data	Sharing->Infrastructure and Tools	Wallis et al. 2013
Insufficient data are released to justify developing a repository, and insufficient demand exists to justify releasing and depositing data	Culture: Demand for Data	Sharing->Planning	Wallis et al. 2013
insufficient resources, will, and leadership to build cyberinfrastructure for the humanities and social sciences	Culture: Priority	Culture->Human Resources	NSF 2007
insufficient resources, will, and leadership to build cyberinfrastructure for the humanities and social sciences	Culture: Priority	Culture->Infrastructure and Tools	NSF 2007
integrating storage and compute resources between local and national infrastructures	Infrastructure and Tools: Difficulty meeting generalized and special needs	Infrastructure and Tools->Research Collaboration	Lynch 2008
Interoperability allows some data and stakeholders in and keeps others out. Policy, practice, standards, business models, and vested interests often are greater determinants of interoperability than is technology	CMP: Interoperability	Culture->CMP	Borgman 2015
Interoperability allows some data and stakeholders in and keeps others out. Policy, practice, standards, business models, and vested interests often are greater determinants of interoperability than is technology	CMP: Interoperability	Funding->CMP	Borgman 2015
Interoperability allows some data and stakeholders in and keeps others out. Policy, practice, standards, business models, and vested interests often are greater determinants of interoperability than is technology	CMP: Interoperability	Legal/Policy->CMP	Borgman 2015
it is difficult to reuse others' data. The further you are away, the harder it is	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Borgman 2012
It is not sufficient to document data pre-processing steps; the processes themselves must be captured	CMP: Insufficient Data Curation or Management	CMP->Reuse	Rauber 2012

Relevant text and ideas	Gap Area	Gap Influence	Source
It is time for a financial review of the situation: we need to understand the full costs of curating data so that we can plan and budget for the required infrastructure more effectively	Knowledge: Costs of Stewardship	Knowledge->Infrastructure and Tools	Lyon 2007
It is time for a financial review of the situation: we need to understand the full costs of curating data so that we can plan and budget for the required infrastructure more effectively	Knowledge: Costs of Stewardship	Knowledge->Planning	Lyon 2007
knowledge-workers' institutions act without vision/strategy; piecemeal	Sustainability Planning: Lack of Strategy and Planning	Planning->Sharing and Access	Borgman 2012
lack a standardized mechanism to reference resources (from astronomy)	Culture: Standards	Culture->Reuse?	Pepe et al 2014
lack of a clear OGD strategy that provides incentives to public officials to invest in OGD activities	Legal and Policy: Incentives	Legal/Policy->Funding	Ubaldi 2013
lack of a clear OGD strategy that provides incentives to public officials to invest in OGD activities	Legal and Policy: Incentives	Planning->Funding	Ubaldi 2013
lack of a clear OGD strategy that provides incentives to public officials to invest in OGD activities	Legal and Policy: Incentives	Planning->Legal/Policy	Ubaldi 2013
lack of an ongoing and efficient flow of resources to preservation throughout the digital lifecycle	Funding: Lack of Funding	Funding->CMP	Berman et al. 2010; Ember and Hanisch 2013
lack of business model	Sustainability Planning: Business and Economic Models	Planning->CMP	Thompson Reuters 2013
lack of certainty about how much it will all cost and how to sustain resources in the long term...The difficulty here is to estimate how much data there will be and how more storage to budget for	Knowledge: Costs of Stewardship	Knowledge->CMP	Weigert 2015
lack of certainty about how much it will all cost and how to sustain resources in the long term...The difficulty here is to estimate how much data there will be and how more storage to budget for	Knowledge: Costs of Stewardship	Knowledge->Planning	Weigert 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
lack of certainty about how much it will all cost and how to sustain resources in the long term...The difficulty here is to estimate how much data there will be and how more storage to budget for	Knowledge: Costs of Stewardship	Planning->CMP	Weigert 2015
lack of certainty about how much it will all cost and how to sustain resources in the long term...The difficulty here is to estimate how much data there will be and how more storage to budget for [Duplicate]	Knowledge: Amount of data	Knowledge->CMP	Weigert 2015
lack of clarity about ownership of data; ownership of large-scale data sets, potentially collected by machines or software providers	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Reuse	OECD 2015
lack of clear guidelines around text and data mining	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Reuse	OECD 2015
lack of common requirements and enforcement practices for data sharing across agencies	Legal and Policy: Lack of Consistency and Alignment	Legal/Policy->Sharing	Reeves et al. 2015
lack of common standards for data management and archiving	Culture: Standards	Culture->CMP	Reeves et al. 2015
Lack of comprehensive plan for research data stewardship and preservation due to primarily bad news, a weaker value proposition, and a challenging business model	Sustainability Planning: Lack of Strategy and Planning	Culture->Planning	Berman 2014
Lack of comprehensive plan for research data stewardship and preservation due to primarily bad news, a weaker value proposition, and a challenging business model	Sustainability Planning: Lack of Strategy and Planning	Planning->CMP	Berman 2014

Relevant text and ideas	Gap Area	Gap Influence	Source
Lack of comprehensive plan for research data stewardship and preservation due to primarily bad news, a weaker value proposition, and a challenging business model	Sustainability Planning: Lack of Strategy and Planning	Planning->Human Resources	Berman 2014
Lack of comprehensive plan for research data stewardship and preservation due to primarily bad news, a weaker value proposition, and a challenging business model	Sustainability Planning: Lack of Strategy and Planning	Planning->Infrastructure and Tools	Berman 2014
Lack of comprehensive plan for research data stewardship and preservation due to a more challenging business model	Sustainability Planning: Business and Economic Models	Planning->CMP	Berman 2014
Lack of comprehensive plan for research data stewardship and preservation due to a more challenging business model	Sustainability Planning: Business and Economic Models	Planning->Human Resources	Berman 2014
Lack of consistency from funders about how data is handled (whether made accessible or not)	Legal and Policy: Lack of Consistency and Alignment	Legal/Policy->Sharing and Access	Borgman 2012
lack of coordination of roles, responsibilities, and funding sources among those best positioned to preserve data (researchers)	Responsibility: Coordinate Stewardship Activities	Collaboration->CMP	Berman et al. 2010
lack of coordination of roles, responsibilities, and funding sources among those best positioned to preserve data (researchers)	Responsibility: Coordinate Stewardship Activities	Responsibility->CMP	Berman et al. 2010
lack of critical mass to develop and sustain shared data resources	Collaboration: Lack of Critical Mass	Collaboration->Infrastructure and Tools	Borgman 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
lack of demonstrated demand for research data outside of genomics, climate science, astronomy, social science surveys, and a few other areas	Culture: Demand for Data	Culture->Reuse?	Borgman 2012
lack of effective incentive mechanisms to promote data-sharing practices among researchers	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	OECD 2015
lack of expertise, resources to make metadata available	Human Resources: Skills	Human Resources->CMP	Eakin et al. n.d.
lack of expertise, resources to make metadata available	Human Resources: Skills	Human Resources->Sharing and Access	Eakin et al. n.d.
lack of funding and resources (libraries and museums) needed for large scale, long-term digital data stewardship	Funding: Lack of Funding	Funding->CMP	Downs and Chen 2013; Erway 2015
lack of funding and resources (libraries and museums) needed for large scale, long-term digital data stewardship	Funding: Lack of Funding	Funding->Human Resources	Downs and Chen 2013; Erway 2015
lack of funding and resources (libraries and museums) needed for large scale, long-term digital data stewardship	Funding: Lack of Funding	Funding->Infrastructure and Tools	Downs and Chen 2013; Erway 2015
lack of incentives for competition transparency; impacts	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	Manyika et al. 2011
lack of incentives to experiment with cyberinfrastructure in the humanities and social sciences;	Legal and Policy: Incentives	Legal/Policy->CMP	NSF 2007
lack of incentives to experiment with cyberinfrastructure in the humanities and social sciences;	Legal and Policy: Incentives	Legal/Policy->Human resources	NSF 2007
lack of incentives to experiment with cyberinfrastructure in the humanities and social sciences;	Legal and Policy: Incentives	Legal/Policy->Infrastructure and Tools	NSF 2007
lack of information about costs for decision making	Knowledge: Costs of Stewardship	Knowledge->Planning	STC 2007
lack of information about costs of keeping systems up to date as new technologies emerge	Knowledge: Costs of Stewardship	Knowledge->CMP	Brown et al. 2015
lack of information about costs of keeping systems up to date as new technologies emerge	Knowledge: Costs of Stewardship	Knowledge->Planning	Brown et al. 2015
lack of information about costs of keeping systems up to date as new technologies emerge	Knowledge: Costs of Stewardship	Planning->CMP	Brown et al. 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
lack of information about who has previously enhanced, annotated or had access to the data	Knowledge: Provenance and Authenticity	Knowledge->CMP (quality, trustworthiness)	Thaesis and van der Hoeven 2010
lack of infrastructure for large scale, long-term digital data stewardship; lack of repositories for CENS research areas	Infrastructure and Tools: Lack of Infrastructure	Infrastructure and Tools->CMP	Downs and Chen 2013; Wallis et al. 2013
lack of insight into a reasonable fee structure/amount: “Trying to figure out the right cost for a Terabyte of data, or 7 hours of curation	Knowledge: Costs of Stewardship	Knowledge->Planning	Kupiainen 2015
lack of interoperability (among silos of information)	CMP: Interoperability	Infrastructure->Sharing and Access	Borgman 2015; Thaesis and van der Hoeven 2010; Ubaldi 2013
Lack of knowledge about preservation practices	Knowledge: How to preserve	Knowledge->CMP	STC 2007
lack of liability arrangements may prevent researchers from storing data	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->CMP	Thaesis and van der Hoeven 2010
lack of liability arrangements may prevent sharing	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Sharing	Borgman 2015
lack of perceived demand; CENS investigators invest little effort in metadata creation because they cannot foresee who might use their data	Culture: Demand for Data	Knowledge->Curation	Wallis et al. 2013
lack of reward or credit for sharing	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	Borgman 2012
lack of standardized archival practices and data formats (curation)	Culture: Standards	Culture->CMP	STC 2007
lack of standards (including Internet of Things)	Culture: Standards	Culture->Sharing and Access	NAS 2009

Relevant text and ideas	Gap Area	Gap Influence	Source
lack of standards for infrastructure	Culture: Standards	Culture->Sharing and Access	Borgman 2015
lack of standards for infrastructure	Culture: Standards	Culture->Infrastructure and Tools	Borgman 2015
lack of support for data management for researchers in short and long term	Human Resources: Support for Data Management	Human Resources->CMP	Giarlo 2012; Tenopir et al. 2011
lack of understanding of preservation issues related to datasets	Knowledge: How to preserve	Knowledge->Planning	Brown et al. 2015
lack of vocabularies (for quality) applicable across domains	Culture: Standards	Culture->Reuse	McDonough 2012
legal contracts by which data were obtained may prevent sharing	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Sharing	Borgman 2015
legislation for sharing databases has been problematic; it is problematic for researchers to have to agree to restrictive terms to use a database	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Reuse	NRC 2003
legislation for sharing databases has been problematic; it is problematic for researchers to have to agree to restrictive terms to use a database	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Sharing	NRC 2003
Libraries are not often set up with the proper infrastructure and support to quickly adapt to new digital strategies and business models	Infrastructure and Tools: Lack of Infrastructure	Infrastructure and Tools->Planning	Johnson et al. 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
likelihood that recipient will not reciprocate	Culture: Sharing attitudes and practices	Culture->Sharing and Access	NRC 2003
links to “personal datasets”, i.e., links to potential data hosted on astronomers' personal websites, become unreachable much faster than links to curated “institutional datasets”	CMP: Data Management Fragmented	Culture->CMP	Pepe et al 2014
links to “personal datasets”, i.e., links to potential data hosted on astronomers' personal websites, become unreachable much faster than links to curated “institutional datasets”	CMP: Data Management Fragmented	Culture->Sharing and Access	Pepe et al 2014
Little known about beginning, middle, and end of knowledge infrastructures and what happens to the data they held. When research funding ends, collaborations may disband gracefully, taking resources and expertise to the next projects.	Knowledge: Other	Knowledge->CMP	Borgman 2015
Lots of collections, resources out there; not a programmatic way of dealing with them	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->CMP	Palmer 2015
Low funding priority for data management and preservation in many scientific disciplines; lack of funding	Culture: Priority	Culture->Data Management and Preservation	Esanu et al. 2004; NAS 2009; NRC 2003
Low funding priority for data management and preservation in many scientific disciplines; lack of funding	Culture: Priority	Culture->Funding	Esanu et al. 2004; NAS 2009; NRC 2003
management of metadata is critical for archivists, but not priority for manufacturers or users	Culture: Priority	Culture->CMP	STC 2007
Managing data requires a much different architecture than systems for publications or other textual documents. Rarely are data self-describing, nor do they stand alone as independent units	CMP: Other	Infrastructure and Tools->CMP	Borgman et al. 2014

Relevant text and ideas	Gap Area	Gap Influence	Source
Managing research data is a knowledge infrastructure problem beyond the scope of individual researchers or projects	Collaboration: Need for Collaboration	Collaboration->CMP	Borgman et al. 2014
Many bemoaned the lack of common metadata and archives as a main impediment to using and storing data	Culture: Standards	Culture->CMP	Science Staff 2011
Many bemoaned the lack of common metadata and archives as a main impediment to using and storing data	Culture: Standards	Culture->Reuse	Science Staff 2011
Many bemoaned the lack of common metadata and archives as a main impediment to using and storing data	Culture: Standards	Culture->Sharing and Access	Science Staff 2011
many existing repositories will not scale to future demands, interoperate well among disciplines, nor guarantee long-term access	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Infrastructure and Tools->Sharing and Access	Atkins et al. 2003
Many governments and research institutions are currently bearing the costs of offering open access to articles and to data, as well as the costs of storage and the preservation of data sets on line. Given the rapidly increasing amounts of data, public institutions will be challenged to find sustainable funding and business models. Public-private partnerships with private service providers may offer innovative solutions.	Sustainability Planning: Business and Economic Models	Collaboration->Funding	OECD 2015
Many governments and research institutions are currently bearing the costs of offering open access to articles and to data, as well as the costs of storage and the preservation of data sets on line. Given the rapidly increasing amounts of data, public institutions will be challenged to find sustainable funding and business models. Public-private partnerships with private service providers may offer innovative solutions.	Sustainability Planning: Business and Economic Models	Planning->Funding	OECD 2015
Many traditional information preservation and access institutions such as libraries and museums are struggling to develop the skills, resources, and infrastructure needed for large scale, long-term digital data stewardship	Human Resources: Skills	Human Resources->CMP	Downs and Chen 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
misalignment between the scale of the need for on-campus data curation and the level of commitment by academic libraries to address this need (as measured by the amount of resources allocated to this need vs. other needs)	Commitment: Extent of Commitment	Commitment->CMP	Giarlo 2012
misalignment between the scale of the need for on-campus data curation and the level of commitment by academic libraries to address this need (as measured by the amount of resources allocated to this need vs. other needs)	Commitment: Extent of Commitment	Commitment->Funding	Giarlo 2012
misalignment between the scale of the need for on-campus data curation and the level of commitment by academic libraries to address this need (as measured by the amount of resources allocated to this need vs. other needs)	Commitment: Extent of Commitment	Commitment->Human Resources	Giarlo 2012
misalignment between the scale of the need for on-campus data curation and the level of commitment by academic libraries to address this need (as measured by the amount of resources allocated to this need vs. other needs)	Commitment: Extent of Commitment	Commitment->Infrastructure and Tools	Giarlo 2012
mismatch between job classifications like archivist and librarian, and the actual demands of the data world. The implication is that repositories cannot operate effectively on soft money; there needs to be an underlying sustainable funding base	Funding: Lack of Funding	Funding->Human Resources	Ember and Hanisch 2013
mismatch between the long-term commitment to preservation inherent in the work of archiving, and the short-term and episodic funding upon which this work is based; long-term preservation now only covered for first 5 years (RDA)	Commitment: Duration of Commitment	Commitment->CMP	Kupiainen 2015; Ember and Hanisch 2013
mismatch between the long-term commitment to preservation inherent in the work of archiving, and the short-term and episodic funding upon which this work is based; long-term preservation now only covered for first 5 years (RDA)	Commitment: Duration of Commitment	Commitment->Funding	Kupiainen 2015; Ember and Hanisch 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
mismatch between the long-term commitment to preservation inherent in the work of archiving, and the short-term and episodic funding upon which this work is based; long-term preservation now only covered for first 5 years (RDA)	Commitment: Duration of Commitment	Commitment->Human Resources	Kupiainen 2015; Ember and Hanisch 2013
mismatch between the long-term commitment to preservation inherent in the work of archiving, and the short-term and episodic funding upon which this work is based; long-term preservation now only covered for first 5 years (RDA)	Commitment: Duration of Commitment	Commitment->Infrastructure and Tools	Kupiainen 2015; Ember and Hanisch 2013
More labor is required to describe unique items or to merge them into common structures. Each research data set is unique and requires its own metadata and provenance records.	Human Resources: Time	Human Resources->CMP	Borgman 2015
Most evaluations of universities and researchers are almost entirely based on teaching and bibliometric indicators, attributing little value to the sharing of pre-publication inputs and post-publication outcomes, such as data and other relevant information	Legal and Policy: Incentives	Legal/Policy->CMP	OECD 2015
Most evaluations of universities and researchers are almost entirely based on teaching and bibliometric indicators, attributing little value to the sharing of pre-publication inputs and post-publication outcomes, such as data and other relevant information	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	OECD 2015
Most researchers will share more readily with their peers, given the concerns for labor, interpretation, and likelihood of reuse	Human Resources: Time	Human Resources->Sharing and Access	Borgman 2012
Much of the data being sought is “preserved” in ad hoc and fragmented ways, and all too often ends up in “data mortuaries” rather than archives	CMP: Other	CMP->Reuse	Atkins et al. 2003; Wynholds et al. 2011
Much of the effort under way to use cyberinfrastructure for collaborative research is not giving adequate attention to sociological and culture barriers to technology adoption that may cause failure, even after large investments	Culture: Research and Development	Culture->Collaboration	Atkins et al. 2003
Much of the effort under way to use cyberinfrastructure for collaborative research is not giving adequate attention to sociological and culture	Culture: Research and Development	Culture->Human Resources	Atkins et al. 2003

Relevant text and ideas	Gap Area	Gap Influence	Source
barriers to technology adoption that may cause failure, even after large investments			
Much of the effort under way to use cyberinfrastructure for collaborative research is not giving adequate attention to sociological and culture barriers to technology adoption that may cause failure, even after large investments	Culture: Research and Development	Culture->Infrastructure and Tools	Atkins et al. 2003
Much of the effort under way to use cyberinfrastructure for collaborative research is not giving adequate attention to sociological and culture barriers to technology adoption that may cause failure, even after large investments	Culture: Research and Development	Infrastructure->Collaborative Research	Atkins et al. 2003
need to ensure data are accessible, usable by general public; this takes more work than making data accessible to fellow specialists and may require focused efforts	Human Resources: Time	Human Resources->Reuse	Wynholds et al. 2011
need to ensure data are accessible, usable by general public; this takes more work than making data accessible to fellow specialists and may require focused efforts	Human Resources: Time	Human Resources->Sharing and Access	Wynholds et al. 2011
Neither the producers of data nor the agencies that require sharing can agree on what are the data	Legal and Policy: Lack of Consistency and Alignment	Legal/Policy->Sharing and Access	Borgman 2012
No enterprise yet has a long-term strategy or solution [for digital preservation] that does not require significant and ongoing capital investment and operational expense.	Sustainability Planning: Business and Economic Models	Funding->CMP	STC 2007
No enterprise yet has a long-term strategy or solution [for digital preservation] that does not require significant and ongoing capital investment and operational expense.	Sustainability Planning: Business and Economic Models	Funding->Human Resources	STC 2007
No enterprise yet has a long-term strategy or solution [for digital preservation] that does not require significant and ongoing capital investment and operational expense.	Sustainability Planning: Business	Funding->Infrastructure and Tools	STC 2007

Relevant text and ideas	Gap Area	Gap Influence	Source
	and Economic Models		
No funding for investigating feasible business models; not part of contracted deliverables.	Funding: Lack of Funding	Funding->Planning	Kupiainen 2015
no group whose mainstream mission it is to plan and coordinate the data infrastructure needed for the research community (reason for no comprehensive plan)	Responsibility: Coordinate Stewardship Activities	Commitment->Infrastructure and Tools	Berman 2014
no group whose mainstream mission it is to plan and coordinate the data infrastructure needed for the research community (reason for no comprehensive plan)	Responsibility: Coordinate Stewardship Activities	Commitment->Planning	Berman 2014
no group whose mainstream mission it is to plan and coordinate the data infrastructure needed for the research community (reason for no comprehensive plan)	Responsibility: Coordinate Stewardship Activities	Planning->Infrastructure and Tools	Berman 2014
No mechanism for expressing or knowing if access to datasets is required on a time-frame faster than normal quality control processes allow (to enforce quality standards on data before they are used)	Discovery	Discovery->Reuse	Ashley 2012
no one knows what will have value; difficult to determine what will be of value;	Knowledge: Value	Knowledge->CMP	Borgman 2015; Brown et al. 2015; Esanu et al. 2004
no one sector will be able to take on the responsibility for stewardship of all research data	Responsibility: Conduct Stewardship Activities	Collaboration->Responsibility	Berman 2010
no single private foundation in the United States—with the exception of the Bill & Melinda Gates Foundation, which primarily funds health initiatives—has annual funding that equals the budget of CISE. Among the	Funding: Imbalance in Funding	Funding->Human Resources	NSF 2007

Relevant text and ideas	Gap Area	Gap Influence	Source
large private foundations, few are focused on humanities and social sciences			
no single private foundation in the United States—with the exception of the Bill & Melinda Gates Foundation, which primarily funds health initiatives—has annual funding that equals the budget of CISE. Among the large private foundations, few are focused on humanities and social sciences	Funding: Imbalance in Funding	Funding->Infrastructure and Tools	NSF 2007
No standard policy about open access and archiving among funders	Legal and Policy: Lack of Consistency and Alignment	Legal/Policy->Sharing and Access	Thaesis and van der Hoeven 2010
not all researchers are necessarily aware of the possibilities offered by open science.	Knowledge: Reuse Possibilities	Knowledge->Reuse	OECD 2015
not clear how such longterm stewardship should be funded and supported	Responsibility: Support Stewardship Activities	Commitment->Responsibility	Downs and Chen 2013
not clear how such longterm stewardship should be funded and supported	Responsibility: Support Stewardship Activities	Responsibility->Funding	Downs and Chen 2013
not clear what organizations or institutions can and should maintain and store such data, ensuring their long-term integrity and usability	Responsibility: Conduct Stewardship Activities	Responsibility->CMP	Berman et al. 2010; Borgman 2015; Downs and Chen 2013
Not possible to know costs until a preservation strategy is chosen	Knowledge: Costs of Stewardship	Planning-Knowledge	Eakin et al., n.d.; Lavoie 2006, (quoted in Eakin et al. n.d.)
not yet agreed upon standards (e.g., in brain imaging) for how to share, annotate, etc	Culture: Standards	Culture->Reuse	NRC 2003

Relevant text and ideas	Gap Area	Gap Influence	Source
not yet agreed upon standards (e.g., in brain imaging) for how to share, annotate, etc	Culture: Standards	Culture->Sharing and Access	NRC 2003
OGD: A critical element of accessibility is to know the source of what the user is searching for, and in many instances, where to begin searching can pose a challenge	Discovery	Discovery->Reuse	Ubaldi 2013
OGD: Certain legal requirements, as well as fragmented and diverse legislation concerning privacy (e.g. UK Data Protection Act, re-use and sometimes fees (e.g. in Sweden and Germany36) can create confusion for end-users, or can make it more difficult for	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Reuse	Ubaldi 2013
OGD: challenges when federal government seeks to impose coordination or consistency across the broad range of rulemaking processes, data and portals enabling access to government data (need to collaborate and coordinate)	Legal/Policy: Institutional structures and pressures	Collaboration->Sharing and Access	Ubaldi 2013
OGD: Converting large volumes of data into re-useable formats has cost implications	Knowledge: Costs of Stewardship	CMP->Sharing and Access	Ubaldi 2013
OGD: difficult from the user perspective to know which piece of data is valid or should be trusted	Reuse: Difficulty of Reusing Data	CMP->Reuse	Ubaldi 2013
OGD: every public agency has its own set of data, formats and standards [Duplicate]	Culture: Standards	Culture->Reuse	Ubaldi 2013
OGD: every public agency has its own set of data, formats and standards. This can make it difficult from the user perspective to know which piece of data is valid or should be trusted. A critical element of accessibility is to know the source of what the user is searching for, and in many instances, where to begin searching can pose a challenge.	Culture: Data Definition	Culture->Reuse	Ubaldi 2013
OGD: how to develop a new business case and financing model for collecting, converting, and diffusing public data in light of the accessibility principle that open government data should be free or provided at cost	Sustainability Planning: Business and Economic Models	Planning->CMP	Ubaldi 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
OGD: how to develop a new business case and financing model for collecting, converting, and diffusing public data in light of the accessibility principle that open government data should be free or provided at cost	Sustainability Planning: Business and Economic Models	Planning->Funding	Ubaldi 2013
OGD: how to develop a new business case and financing model for collecting, converting, and diffusing public data in light of the accessibility principle that open government data should be free or provided at cost	Sustainability Planning: Business and Economic Models	Planning->Sharing and Access	Ubaldi 2013
OGD: internal culture of public sector institutions is not immediately conducive to data sharing and requires additional efforts	Culture: Sharing attitudes and practices	Culture->Sharing and Access	Ubaldi 2013
OGD: lack of appropriate institutional structures to ensure accountability, transparency, quality of data and responsibility in a context of collaboration	Collaboration: Support Structures	Culture->Collaboration	Ubaldi 2013
OGD: lack of integration of OGD tools and applications	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->Reuse	Ubaldi 2013
OGD: lack of procedures and standards on how to deal with open data in governments (e.g. lack of tools available to make data open, of validation structures and guidelines, guidelines on data collection that can compromise the quality of the data and even	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->CMP	Ubaldi 2013
OGD: Need for engagement models and strategies to allow two-way dialogue to take place between the public sector and the users of government data	Collaboration: Lack of Collaboration	Planning->Reuse	Ubaldi 2013
OGD: Proliferation of regulatory compliance requirements: federal web designers in a compliance minefield that makes it hard for them to avoid breaking the rules – while diverting energy from innovation towards compliance	Legal and Policy: Incentives	Legal/Policy->Infrastructure and Tools	Ubaldi 2013
OGD: There is a substantial commitment and investment on the part of public agencies as they need to acquire new skills, train employees,	Commitment: Other	Commitment->Funding	Ubaldi 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
purchase technologies, and upgrade network infrastructure, which need to be accounted for.			
OGD: There is a substantial commitment and investment on the part of public agencies as they need to acquire new skills, train employees, purchase technologies, and upgrade network infrastructure, which need to be accounted for.	Commitment: Other	Commitment->Human Resources	Ubaldi 2013
OGD: There is a substantial commitment and investment on the part of public agencies as they need to acquire new skills, train employees, purchase technologies, and upgrade network infrastructure, which need to be accounted for.	Commitment: Other	Commitment->Infrastructure and Tools	Ubaldi 2013
OGD: Use of data restricted due to disclosure policies that limit data transparency or copyrights that result in lack of clarity over who owns government data.	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Reuse	Ubaldi 2013
OGD: value and transparency might be hindered if data are not simple to access due to its formats	Human Resources: Other	Access->Knowledge	Ubaldi 2013
Open science efforts involve different communities and different actors: researchers, governmental institutions, universities and research centres, libraries and data centres, private non-profit organisations, business sector organisations including private academic publishers, supra-national entities, citizens, etc.	Culture: Other	Collaboation->Infrastructure and Tools	OECD 2015
Open science efforts involve different communities and different actors: researchers, governmental institutions, universities and research centres, libraries and data centres, private non-profit organisations, business sector organisations including private academic publishers, supra-national entities, citizens, etc.	Culture: Other	Collaboration->Human Resources	OECD 2015
Open science efforts involve different communities and different actors: researchers, governmental institutions, universities and research centres, libraries and data centres, private non-profit organisations, business	Culture: Other	Collaboration->Planning	OECD 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
sector organisations including private academic publishers, supra-national entities, citizens, etc.			
Open science efforts involve different communities and different actors: researchers, governmental institutions, universities and research centres, libraries and data centres, private non-profit organisations, business sector organisations including private academic publishers, supra-national entities, citizens, etc.	Culture: Other	Collaboration->Sharing and Access	OECD 2015
organizational structures are not robust enough to support resources moving between local and national resources or to allow design of tomorrow's architectures, tools, and services (see domain repository problem of having resources to investigate business	Collaboration: Support Structures	Collaboration->Infrastructure and Tools	Lynch 2008
organizational structures are not robust enough to support resources moving between local and national resources or to allow design of tomorrow's architectures, tools, and services (see domain repository problem of having resources to investigate business	Collaboration: Support Structures	Culture->Collaboration	Lynch 2008
particular domains may have valuable systems set up that clash with the expectations of certification—for instance, in oceanography researchers are required to put their data in certain repositories and certification would require the repositories to reje	Legal/Policy: Institutional structures and pressures	Legal/Policy->CMP	Edwards et al. 2013
Planning and developing requirements for data archives must accommodate the continual change and evolution in the practice of science; the local variability in focus, practice, and available technology and other physical and human infrastructure; the differing mandates and objectives of different data producers, as well as a diversity of potential users, including scientists, educational institutions, businesses, policymakers, and ordinary citizens	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Culture->Planning	Esanu et al. 2004
Planning and developing requirements for data archives must accommodate the continual change and evolution in the practice of science; the local variability in focus, practice, and available technology and other physical and human infrastructure; the differing mandates and	Sustainability Planning: Planning for Dynamic and	Human Resources->Planning	Esanu et al. 2004

Relevant text and ideas	Gap Area	Gap Influence	Source
objectives of different data producers, as well as a diversity of potential users, including scientists, educational institutions, businesses, policymakers, and ordinary citizens	Adaptable Infrastructure		
Planning and developing requirements for data archives must accommodate the continual change and evolution in the practice of science; the local variability in focus, practice, and available technology and other physical and human infrastructure; the differing mandates and objectives of different data producers, as well as a diversity of potential users, including scientists, educational institutions, businesses, policymakers, and ordinary citizens	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Infrastructure and Tools->Planning	Esanu et al. 2004
Planning and developing requirements for data archives must accommodate the continual change and evolution in the practice of science; the local variability in focus, practice, and available technology and other physical and human infrastructure; the differing mandates and objectives of different data producers, as well as a diversity of potential users, including scientists, educational institutions, businesses, policymakers, and ordinary citizens	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Sharing and Access->Planning	Esanu et al. 2004
plans for the reevaluation of resources also need to protect digital resources that happen to be reviewed when the represented topics or disciplines are unpopular or when budgets are being scrutinized injudiciously	Sustainability Planning: Other	Planning->CMP	Downs and Chen 2013
Policies to promote open data are less mature than those to promote open access to scientific publications.	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Sharing and Access	OECD 2015
Policy makers need to promote openness in science while at the same time preserving competition; pushing for open access and open data too early may be counterproductive in some cases.	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	OECD 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
poor interoperability of data sources and steep learning curve	CMP: Interoperability	CMP->Reuse	Wynholds et al. 2011
Popular attention drawn to extreme futures; miss middle ground of getting there	Sustainability Planning: Other	Culture->Planning	Edwards et al. 2013
preservation is an organizational as well as a technical challenge and the responsibility, as has been widely recognized, spans a broad range of stakeholders	Responsibility: Support Stewardship Activities	Collaboration->Responsibility	ARL 2006; Turner et al. 2015
privacy protections make data harder to use or less useful	Reuse: Difficulty of Reusing Data	Legal/Policy->Reuse	Fiore 2012
Private foundations are important sources of support in the humanities and the social sciences [for cyberinfrastructure], but they cannot make up for the low level of federal funding.	Funding: Imbalance in Funding	Funding->CMP	NSF 2007
Private foundations are important sources of support in the humanities and the social sciences [for cyberinfrastructure], but they cannot make up for the low level of federal funding.	Funding: Imbalance in Funding	Funding->Human Resources	NSF 2007
Private foundations are important sources of support in the humanities and the social sciences [for cyberinfrastructure], but they cannot make up for the low level of federal funding.	Funding: Imbalance in Funding	Funding->Infrastructure and Tools	NSF 2007
Problems in using big data (privacy, hides complexities and expertise involved)	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Borgman 2012
Process for managing changes over time in needs and responsibilities of data authors, managers, users, funding agencies at the level of the discipline or community is no longer adequate given the substantial cost of creating data collections and managing	Collaboration: Need for Collaboration	Collaboration->Human Resources	NSB 2005
Process for managing changes over time in needs and responsibilities of data authors, managers, users, funding agencies at the level of the discipline or community is no longer adequate given the substantial cost of creating data collections and managing	Collaboration: Need for Collaboration	Collaboration->Planning	NSB 2005

Relevant text and ideas	Gap Area	Gap Influence	Source
proliferation of MTAs is an impediment to sharing pub-related materials	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Sharing	NRC 2003
properly managing data volume (the number of bits is enormous and growing)	CMP: Difficulty Managing Data for Reuse	CMP->Reuse (by implication)	Esanu et al. 2004; STC 2007
Providing access in ways that sustain their value for research requires computational power	Infrastructure and Tools: Other	Infrastructure and Tools->Reuse	Borgman 2015
Providing access in ways that sustain their value for research requires computational power	Infrastructure and Tools: Other	Infrastructure and Tools->Sharing and Access	Borgman 2015
Providing access in ways that sustain their value for research requires continual investment in curation	Funding: Other	Funding->CMP	Borgman 2015
Providing access in ways that sustain their value for research requires continual investment in curation	Funding: Other	Funding->Sharing and Access	Borgman 2015
Providing access in ways that sustain their value for research requires staff expertise to assist users in exploiting the resources	Reuse: Difficulty of Reusing Data	Knowledge->CMP	Borgman 2015
Providing access in ways that sustain their value for research requires staff expertise to assist users in exploiting the resources	Reuse: Difficulty of Reusing Data	Knowledge->Sharing and Access	Borgman 2015
publishers appreciated the benefits of sharing data, but felt that their servers would have difficulty holding the quantity of data included in each article and that repositories were the right place	Responsibility: Conduct Stewardship Activities	Responsibility->CMP	Downs and Chen 2013; Sturges et al. 2015
publishers appreciated the benefits of sharing data, but felt that their servers would have difficulty holding the quantity of data included in each article and that repositories were the right place	Responsibility: Conduct Stewardship Activities	Responsibility->Infrastructure and Tools	Downs and Chen 2013; Sturges et al. 2015
push towards open access, while creating more equity of access for the community of users, creates a Burdan for domain repositories because it	Culture: Other	Legal/Policy->Funding	Ember and Hanisch 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
narrows their funding possibilities. Without care, this shift may create a different kind of inequity—less well-			
RDMS requires a diverse range of skills, many outside the typical expertise of library staff and not all libraries can afford to hire new and/or retrain staff. Creating archiving infrastructure and curating research data are also expensive endeavors.	Funding: Lack of Funding	Funding->CMP	Fearon et al. 2013
RDMS requires a diverse range of skills, many outside the typical expertise of library staff and not all libraries can afford to hire new and/or retrain staff. Creating archiving infrastructure and curating research data are also expensive endeavors.	Funding: Lack of Funding	Funding->Human Resources	Fearon et al. 2013
RDMS requires a diverse range of skills, many outside the typical expertise of library staff and not all libraries can afford to hire new and/or retrain staff. Creating archiving infrastructure and curating research data are also expensive endeavors.	Funding: Lack of Funding	Funding->Infrastructure and Tools	Fearon et al. 2013
relatively few formal digital communities and relatively few institutional platforms for online collaboration in the humanities	Collaboration: Lack of Collaboration	Infrastructure and Tools->Collaboration	NSF 2007
Reluctance to use datasets for which the researcher lacked prerequisite expertise.	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Wynholds et al. 2011
researchers do not provide data managers with all types of data they make use of, e.g. source code and raw data sets	Culture: Sharing attitudes and practices	Culture->CMP	Thaesis and van der Hoeven 2010
researchers don't realize planning early [for data management] could save expense and pain later	Sustainability Planning: Lack of Strategy and Planning	Planning->Funding	Lynch 2008
researchers don't realize planning early [for data management] could save expense and pain later	Sustainability Planning: Lack of Strategy and Planning	Planning->Human Resources	Lynch 2008

Relevant text and ideas	Gap Area	Gap Influence	Source
researchers lack expertise in data management and data curation	Human Resources: Skills	Human Resources->CMP	Wallis et al. 2013
Researchers produce large amounts of data, some of which may be useful to others. Making those data useful to others requires a substantial investment in documentation, and often in interpersonal negotiation, above and beyond the conduct of the research per se. It is not possible to justify making that level of investment in all data just in case someone, somewhere, at some future time, might wish to use them.	Human Resources: Time	Human Resources->CMP	Wallis et al. 2013
Researchers produce large amounts of data, some of which may be useful to others. Making those data useful to others requires a substantial investment in documentation, and often in interpersonal negotiation, above and beyond the conduct of the research per se. It is not possible to justify making that level of investment in all data just in case someone, somewhere, at some future time, might wish to use them.	Human Resources: Time	Human Resources->Reuse	Wallis et al. 2013
Researchers worry about valuable earlier work that they still store, but that in reality is unavailable for use because it is recorded on outdated media (e.g., floppy disks), formatted for obsolete software, or relies on antique programming languages. Som	CMP: Insufficient Data Curation or Management	CMP->Reuse	Kroll and Forsman 2010
Researchers' skills needed for sharing articles or data sets openly on line are unevenly distributed (Disciplines like CS and physics have longer tradition of uploading to repositories and curating large datasets; researchers in other disciplines may need	Human Resources: Skills	Human Resources->Sharing and Access	OECD 2015
Reticence to trust data: The astronomers were wary of the dangers of treating the data as given or as an aggregation of facts rather than as the output of complex observational instruments.	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Wynholds et al. 2011
Scholars need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later	Infrastructure and Tools: Lack of Tools	Human Resources->CMP	Borgman 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
Scholars need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later	Infrastructure and Tools: Lack of Tools	Human Resources->Reuse	Borgman 2015
Scholars need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later	Infrastructure and Tools: Lack of Tools	Human Resources->Sharing and Access	Borgman 2015
Scholars need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->CMP	Borgman 2015
Scholars need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->Reuse	Borgman 2015
Scholars need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->Sharing and Access	Borgman 2015
scholarship is exploratory, local, diverse, and lacks shared community resources	Infrastructure and Tools: Lack of Infrastructure		
Scientific data are heterogeneous, and both interpretation and management require deep expertise in the domain.	Human Resources: Other	Human Resources->CMP	Wynholds et al. 2011
Scientific data are heterogeneous, and both interpretation and management require deep expertise in the domain.	Human Resources: Other	Human Resources->Reuse	Wynholds et al. 2011
Scientific data are heterogeneous, and both interpretation and management require deep expertise in the domain.	Human Resources: Other	Human Resources->Sharing and Access	Wynholds et al. 2011
shortage of qualified people to work at repositories, particularly in developing data management systems, defining data models, and programming	Human Resources: Skills	Human Resources->CMP	Ember and Hanisch 2013
shortage of qualified people to work at repositories, particularly in developing data management systems, defining data models, and programming	Human Resources: Skills	Human Resources->Infrastructure and Tools	Ember and Hanisch 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
Should a primary and backup archive system be connected, or not be connected, by a network?	Knowledge: How to preserve	Knowledge->Infrastructure and Tools	STC 2007
since it can be difficult to trust data quality, etc., is it not better to collect new or additional data?	Reuse: Difficulty of Reusing Data	Knowledge->Reuse	Tenopir et al. 2015
skills required for preservation are in short supply	Human Resources: Skills	Human Resources->CMP	Brown et al. 2015
Some institutions have not fully considered how to sustain RDM for the long term	Sustainability Planning: Business and Economic Models	Planning->CMP	Brown et al. 2015
Some institutions have not fully considered how to sustain RDM for the long term	Sustainability Planning: Business and Economic Models	Planning->Infrastructure and Tools	Brown et al. 2015
Stakeholders may disagree on what kinds of data in any domain might be worthy of that degree of investment or to whom those data might be valuable	Commitment: Other	Culture->Commitment	Borgman 2015
Stakeholders may disagree on what kinds of data in any domain might be worthy of that degree of investment or to whom those data might be valuable	Commitment: Other	Culture->Reuse	Borgman 2015
Standards can improve the flow of information within communities but also can create boundaries between them	Culture: Standards	Culture->Knowledge	Borgman 2015
students and citizens need to acquire the skills to take advantage of, use and reuse data set shared by the research community.	Human Resources: Skills	Human Resources->Reuse	OECD 2015
tension between a desire to store data in formats which render it fit for long-term access and a desire to store data in formats which render it fit for immediate use	CMP: Tradeoffs between data management for short or long term	CMP->Reuse	McDonough 2012

Relevant text and ideas	Gap Area	Gap Influence	Source
the ability of the NEA, NEH, and IMLS to fund cyberinfrastructure directly is diminished because much of the money in these agency budgets goes to states through block grants over which the agencies have little control.	Legal/Policy: Institutional structures and pressures	Legal/Policy->Funding	NSF 2007
the construction and operation of VOs is limited by either the lack of technology or the lack of organizational understanding that could be gained from more systematic codifying and sharing of information.	Collaboration: Support Structures	Knowledge->Infrastructure and Tools	Cummings et al. 2008
the construction and operation of VOs is limited by either the lack of technology or the lack of organizational understanding that could be gained from more systematic codifying and sharing of information.	Collaboration: Support Structures	Sharing and Access->Infrastructure and Tools	Cummings et al. 2008
the construction and operation of VOs is limited by either the lack of technology or the lack of organizational understanding that could be gained from more systematic codifying and sharing of information.	Collaboration: Support Structures	Sharing and Access->Knowledge	Cummings et al. 2008
the cost of opening up government data, although such cost – as well as the cost of data production – has not been sufficiently appraised so far	Knowledge: Costs of Stewardship	Knowledge->Sharing and Access	Ubaldi 2013
The growth of new ways of knowing may come at the expense of old or alternative ones. Efforts to expand access to knowledge for some groups may curtail or limit the effective access of others	Knowledge: Other	Knowledge->Sharing and Access	Edwards et al. 2013
the most robust, important and adopted infrastructure-related efforts of the VAO, like the VO “Registry” essential for tools to find data, are not at all secure from funding cuts	Funding: Lack of Funding	Funding->Infrastructure and Tools	Pepe et al 2014
the most serious challenge to data curation arises at the semantic level, ensuring the authenticity and correct interpretability of data	CMP: Difficulty Managing Data for Reuse	CMP->Reuse	Rauber 2012
the NVO and VAO efforts skewed towards large, homogenous datasets	Culture: Priority	Culture->CMP	Pepe et al 2014
The real challenges lie in designing digital libraries to assist in the capture, management, interpretation, use, reuse, and stewardship of research data. Opportunities and challenges for the digital library community are plentiful. (8)	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->CMP	Borgman et al. 2014

Relevant text and ideas	Gap Area	Gap Influence	Source
The real challenges lie in designing digital libraries to assist in the capture, management, interpretation, use, reuse, and stewardship of research data. Opportunities and challenges for the digital library community are plentiful. (8)	Sustainability Planning: Planning for Dynamic and Adaptable Infrastructure	Planning->Reuse	Borgman et al. 2014
The science and technology studies literature is more concerned with the processes by which publications and data are created than with subsequent curation activities [6–8].	Culture: Priority	Culture->CMP	Wynholds et al. 2011
the size, privacy, longevity, preservation, access, and other requirements for research data preclude a "one-size-fitsall" approach to creation of supporting data cyberinfrastructure	Infrastructure and Tools: Other	Culture->Infrastructure	Berman 2010
the size, privacy, longevity, preservation, access, and other requirements for research data preclude a "one-size-fitsall" approach to creation of supporting data cyberinfrastructure	Infrastructure and Tools: Other	CMP->Infrastructure	Berman 2010
the size, privacy, longevity, preservation, access, and other requirements for research data preclude a "one-size-fitsall" approach to creation of supporting data cyberinfrastructure	Infrastructure and Tools: Other	Sharing and Access->Infrastructure and Tools	Berman 2010
the substantial amount of labor required to document data in reusable forms	Human Resources: Time	Human Resources->CMP	Borgman 2012; Wallis et al. 2013
the substantial amount of labor required to document data in reusable forms	Human Resources: Time	Human Resources->Reuse	Borgman 2012; Wallis et al. 2013
the third biggest challenge reported is faculty (non)engagement due to a lack of awareness of services that the library provides, low perceived value of services, and resistance to data sharing.	Culture: Sharing attitudes and practices	Knowledge->Culture?	Fearon et al. 2013
The vast majority of information created is "unstructured"	CMP: Data Management Fragmented	CMP->Sharing and Access	Gantz 2007
The weakness of NSF strategies and policies governing long-lived data collections is that they have been developed incrementally and have not been considered collectively.	Sustainability Planning: Lack of	Planning->CMP	NSB 2005

Relevant text and ideas	Gap Area	Gap Influence	Source
	Strategy and Planning		
The weakness of NSF strategies and policies governing long-lived data collections is that they have been developed incrementally and have not been considered collectively.	Sustainability Planning: Lack of Strategy and Planning	Planning->Human Resources	NSB 2005
The weakness of NSF strategies and policies governing long-lived data collections is that they have been developed incrementally and have not been considered collectively.	Sustainability Planning: Lack of Strategy and Planning	Planning->Infrastructure and Tools	NSB 2005
there are few robust systems for making decisions about what to preserve	Infrastructure and Tools: Lack of Tools	Infrastructure and Tools->CMP	Berman et al. 2010
there is a lot of workforce turnover, so substantial effort is going into training and retraining; once people are trained they become desirable in the market and there is competition for them. This leads to a lack of continuity and a loss of institutional knowledge	Human Resources: Skills	Human Resources->CMP	Ember and Hanisch 2013
there is not the same interest in maximizing the sharing of research datasets as there is for published research articles, especially at the pre-publication stage.	Culture: Sharing attitudes and practices	Culture->Sharing and Access	OECD 2015
To improve data access, more data must be kept in reusable forms, which requires a change in incentives. For most scholars, the fundamental problem is better management of their own data. They need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later.	Human Resources: Skills	Human Resources->CMP	Borgman 2015
To improve data access, more data must be kept in reusable forms, which requires a change in incentives. For most scholars, the fundamental problem is better management of their own data. They need tools, services, and assistance in archiving their own data in ways they can reuse	Human Resources: Skills	Human Resources->Reuse	Borgman 2015

Relevant text and ideas	Gap Area	Gap Influence	Source
them, which increases the likelihood that their data will be useful to others later.			
To improve data access, more data must be kept in reusable forms, which requires a change in incentives. For most scholars, the fundamental problem is better management of their own data. They need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later.	Human Resources: Skills	Human Resources->Sharing and Access	Borgman 2015
To improve data access, more data must be kept in reusable forms, which requires a change in incentives. For most scholars, the fundamental problem is better management of their own data. They need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later.	Human Resources: Skills	Legal/Policy->CMP	Borgman 2015
To improve data access, more data must be kept in reusable forms, which requires a change in incentives. For most scholars, the fundamental problem is better management of their own data. They need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later.	Human Resources: Skills	Legal/Policy->Reuse	Borgman 2015
To improve data access, more data must be kept in reusable forms, which requires a change in incentives. For most scholars, the fundamental problem is better management of their own data. They need tools, services, and assistance in archiving their own data in ways they can reuse them, which increases the likelihood that their data will be useful to others later.	Human Resources: Skills	Legal/Policy->Sharing and Access	Borgman 2015
To obtain funding, investigators must pitch radically new, rather than incremental, scientific goals. As a result, they may risk reinvention rather than learning from the strengths and weaknesses of projects that went before – digital library and otherwis	Culture: Research and Development	Culture->Infrastructure and Tools	Borgman et al. 2014

Relevant text and ideas	Gap Area	Gap Influence	Source
uncertainty about the future mechanisms, forms, and economics of scholarly publishing and scholarly communication more generally; [impacts creation of cyberinfrastructure]	Knowledge: Scholarly Publishing	Knowledge->CMP	NSF 2007
uncertainty about the future mechanisms, forms, and economics of scholarly publishing and scholarly communication more generally; [impacts creation of cyberinfrastructure]	Knowledge: Scholarly Publishing	Knowledge->Human resources	NSF 2007
uncertainty about the future mechanisms, forms, and economics of scholarly publishing and scholarly communication more generally; [impacts creation of cyberinfrastructure]	Knowledge: Scholarly Publishing	Knowledge->Infrastructure and Tools	NSF 2007
uncertainty about the future mechanisms, forms, and economics of scholarly publishing and scholarly communication more generally; [impacts creation of cyberinfrastructure]	Knowledge: Scholarly Publishing	Knowledge->Planning	NSF 2007
uncertainty of long-term costs remains problematic	Knowledge: Costs of Stewardship	Knowledge->Planning	Bradley 2005, quoted in Eakin et al. n.d.
uncertainty persists among researchers regarding the legal or proprietary distinctions in terms of which shared data is licensed or copyrighted versus what becomes part of the public domain	Legal/Policy: Deficiencies That Inhibit Stewardship, Access, and Use	Legal/Policy->Reuse	Thompson Reuters 2013
undeveloped notion of time on the Web; difficult to assign creation times to content	Infrastructure and Tools: Other	Infrastructure and Tools->CMP	Nelson 2012
Unequal access to the resources and expertise necessary to create and operate a digital data collection (for individuals establishing research collections, and those trying trying to establish resource and reference collections).	Human Resources: Unequal Access to Resources and Expertise	Human Resources->CMP	NSB 2005
Unequal access to the resources and expertise necessary to create and operate a digital data collection (for individuals establishing research collections, and those trying trying to establish resource and reference collections).	Human Resources: Unequal Access to Resources and Expertise	Human Resources->Infrastructure and Tools	NSB 2005

Relevant text and ideas	Gap Area	Gap Influence	Source
users may be unable to understand or use the data e.g. the semantics, format or algorithms involved	CMP: Insufficient Data Curation or Management	CMP->Reuse	Kuipers and van der Hoeven 2009
Value proposition for stewardship	Culture: Priority	Culture->Human Resources	Berman et al. 2010
Value proposition for stewardship	Culture: Priority	Culture->Infrastructure and Tools	Berman et al. 2010
Value proposition for stewardship	Culture: Priority	Culture->Legal/Policy	Berman et al. 2010
We view the challenge of stewarding research as a mutually beneficial partnership among scholars, universities, and the public, and this perspective influences what follows.	Collaboration: Need for Collaboration	Collaboration->Human Resources	UNC-CH 2012
What are the incentives and benefits, such as recognition, for making their research public?	Legal and Policy: Incentives	Legal/Policy->Sharing and Access	Thompson Reuters 2013
What data should be saved, and what should not be saved?	Knowledge: Value	Knowledge->CMP	STC 2007
What is the best digital preservation strategy?	Knowledge: How to preserve	Knowledge->Planning	STC 2007
What level of geographical separation should be achieved; i.e., how far apart is far enough?	Knowledge: How to preserve	Knowledge->Infrastructure and Tools	STC 2007
What policymakers need and what scientists find interesting are often too different; mismatch in what infrastructures have been built to do for political reasons and what is most needed (e.g., global climate knowledge and regional climate knowledge).	Legal/Policy: Institutional structures and pressures	Culture->Infrastructure and Tools	Edwards et al. 2013
Where funding for domain repositories does exist, it is often associated with a particular [short-term] project or program and lacks any long-term commitment	Commitment: Duration of Commitment	Commitment->Funding	Dillo et al. 2015; Ember and Hanisch 2013
Where funding for domain repositories does exist, it is often associated with a particular [short-term] project or program and lacks any long-term commitment	Commitment: Duration of Commitment	Commitment->Human Resources	Dillo et al. 2015; Ember and Hanisch 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
Where funding for domain repositories does exist, it is often associated with a particular [short-term] project or program and lacks any long-term commitment	Commitment: Duration of Commitment	Commitment->Infrastructure and Tools	Dillo et al. 2015; Ember and Hanisch 2013
While a long-term preservation strategy is generally acknowledged to be a vital component of research data management plans, institutional compliance concerns have thus far centered on access rather than archiving	Legal and Policy: Lack of Consistency and Alignment	Legal/Policy->CMP	Lavoie and Malpas 2015
While a long-term preservation strategy is generally acknowledged to be a vital component of research data management plans, institutional compliance concerns have thus far centered on access rather than archiving	Legal and Policy: Lack of Consistency and Alignment	Legal/Policy->Sharing and Access	Lavoie and Malpas 2015
While a long-term preservation strategy is generally acknowledged to be a vital component of research data management plans, institutional compliance concerns have thus far centered on access rather than archiving	Legal and Policy: Lack of Consistency and Alignment	Planning->CMP	Lavoie and Malpas 2015
While the principle of open access to scientific data is well established in OECD countries, the scope of access varies greatly across countries.	CMP: Other	Legal/Policy->Sharing and Access	OECD 2015
who will pay the costs associated with this curation and quality control	Responsibility: Support Stewardship Activities	Responsibility->CMP	Thompson Reuters 2013
who will pay the costs associated with this curation and quality control	Responsibility: Support Stewardship Activities	Responsibility->Funding	Thompson Reuters 2013
With both content-area and digital curation expertise, domain repositories are uniquely capable of ensuring that data and other research products are adequately preserved, enhanced, and made available for replication, collaboration, and cumulative knowledge building. However,	Funding: Lack of Funding	Funding->CMP	Ember and Hanisch 2013

Relevant text and ideas	Gap Area	Gap Influence	Source
the systems currently in place for funding domain repositories in the US are inadequate for these tasks			
With both content-area and digital curation expertise, domain repositories are uniquely capable of ensuring that data and other research products are adequately preserved, enhanced, and made available for replication, collaboration, and cumulative knowledge building. However, the systems currently in place for funding domain repositories in the US are inadequate for these tasks	Funding: Lack of Funding	Funding->Human Resources	Ember and Hanisch 2013
With both content-area and digital curation expertise, domain repositories are uniquely capable of ensuring that data and other research products are adequately preserved, enhanced, and made available for replication, collaboration, and cumulative knowledge building. However, the systems currently in place for funding domain repositories in the US are inadequate for these tasks	Funding: Lack of Funding	Funding->Infrastructure and Tools	Ember and Hanisch 2013
With both content-area and digital curation expertise, domain repositories are uniquely capable of ensuring that data and other research products are adequately preserved, enhanced, and made available for replication, collaboration, and cumulative knowledge building. However, the systems currently in place for funding domain repositories in the US are inadequate for these tasks	Funding: Lack of Funding	Funding->Research Collaboration	Ember and Hanisch 2013
With both content-area and digital curation expertise, domain repositories are uniquely capable of ensuring that data and other research products are adequately preserved, enhanced, and made available for replication, collaboration, and cumulative knowledge building. However, the systems currently in place for funding domain repositories in the US are inadequate for these tasks	Funding: Lack of Funding	Funding->Sharing and Access	Ember and Hanisch 2013
with the exception of leading journals in Economics, there are few cases in which these normative statements [policies for data sharing] are coupled with penalties or incentives to reinforce them	Legal and Policy: Incentives	Legal/Policy does not affect data sharing (needs something behind it)	Pienta et al. 2010

Relevant text and ideas	Gap Area	Gap Influence	Source
Without explicit incentives and credits, it is challenging for companies to step forward and partner productively to support the common good (referring to stewardship)	Legal and Policy: Incentives	Collaboration->CMP	Berman and Cerf 2013
Without explicit incentives and credits, it is challenging for companies to step forward and partner productively to support the common good (referring to stewardship)	Legal and Policy: Incentives	Collaboration->Funding	Berman and Cerf 2013
Without explicit incentives and credits, it is challenging for companies to step forward and partner productively to support the common good (referring to stewardship)	Legal and Policy: Incentives	Collaboration->Human Resources	Berman and Cerf 2013
Without explicit incentives and credits, it is challenging for companies to step forward and partner productively to support the common good (referring to stewardship)	Legal and Policy: Incentives	Collaboration->Infastructure	Berman and Cerf 2013
Without explicit incentives and credits, it is challenging for companies to step forward and partner productively to support the common good (referring to stewardship)	Legal and Policy: Incentives	Collaboration->Sharing and Access	Berman and Cerf 2013
Without explicit incentives and credits, it is challenging for companies to step forward and partner productively to support the common good (referring to stewardship)	Legal and Policy: Incentives	Legal/Policy->Collaboration	Berman and Cerf 2013
Without explicit incentives and credits, it is challenging for companies to step forward and partner productively to support the common good (referring to stewardship)	Legal and Policy: Incentives	Legal/Policy->Funding	Berman and Cerf 2013

References

Addis, M. (2015). *Estimating Research Data Volumes in UK HEI*. Retrieved from http://figshare.com/articles/Estimating_Research_Data_Volumes_in_UK_HEI/1575831

Ashley, K. (2012). Generic Data Quality Metrics – what and why. Presented at the Prepared for the NSF III #1247471 workshop, Curating for Quality: Ensuring Data Quality to Enable New Science. Retrieved from http://dl.acm.org/citation.cfm?id=2582001

Association of Research Libraries (ARL). (2006). *To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering* (A Report to the National Science Foundation from the ARL Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe). Arlington, VA. Retrieved from http://www.arl.org/storage/documents/publications/digital-data-report-2006.pdf

Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-molina, H., Klein, M. L., Messerschmitt, D. G., … Wright, M. H. (2003). *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure*. Retrieved from https://www.nsf.gov/cise/sci/reports/atkins.pdf

Berman, F. (2008). Got data?: a guide to data preservation in the information age. *Communications of the ACM*, *51*(12). Retrieved from http://portal.acm.org/citation.cfm?doid=1409360.1409376

Berman, F. (2010). We need a research data census, *53*(12). Retrieved from http://cacm.acm.org/magazines/2010/12/102121-we-need-a-research-data-census/fulltext

Berman, F. (2014). Despite Growing Data, Infrastructure Stands Still – Why the gap puts research data at risk. *IEEE*. Retrieved from http://theinstitute.ieee.org/ieee-roundup/members/achievements/despite-growing-data-infrastructure-stands-still

Berman, F., & Cerf, V. (2013). Who Will Pay for Public Access to Research Data? *Science*, *341*(6146). Retrieved from http://www.sciencemag.org/content/341/6146/616.full.pdf?keytype=ref&siteid=sci&ijkey=.e2Ezowko%2FxF2

Berman, F., Lavoie, B., Ayris, P., Choudhury, G. S., Cohen, E., Courant, P., … Van Camp, A. (2010). Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information. Retrieved from http://blueribbontaskforce.sdsc.edu/biblio/BRTF_Final_Report.pdf

Borgman, C. L. (2012). The conundrum of sharing research data, *63*(6). Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/asi.22634/epdf

Borgman, C. L. (2015). *Big data, little data, no data: scholarship in the networked world*.

Borgman, C. L., Darch, P. T., Sands, A. E., Wallis, J. C., & Traweek, S. (2014). The Ups and Downs of Knowledge Infrastructures in Science: Implications for Data Management. *Proceedings of the Joint Conference on Digital Libraries, 2014 (DL2014)*. Retrieved from http://works.bepress.com/borgman/321

Bradley, K. (2005). *APSR Sustainability Issues Discussion Paper* (Australian Partnership for Sustainable Repositories). Australian National Library. Retrieved from https://pdfs.semanticscholar.org/183a/e2dae1eaccac8481d4d3c8a81226c5f7c236.pdf

Brown, S., Bruce, R., & Kernohan, D. (2015). Directions for Research Data Management in UK Universities. Retrieved from http://repository.jisc.ac.uk/5951/

Burda, D., & Teuteberg, F. (2013). Sustaining accessibility of information through digital preservation: A literature review. *Journal of Information Science*, *39*(4). Retrieved from http://jis.sagepub.com/content/39/4/442

Cummings, J., Finholt, T. A., Foster, I., Kesselman, C., & Lawrence, K. A. (2008). Beyond Being There: A Blueprint for Advancing the Design, Development, and Evaluation of Virtual Organizations. *Technology*, *3*(2). Retrieved from http://web.ci.uchicago.edu/events/VirtOrg2008/VO_report.pdf

Dillo, I., Hodson, S., & de Waard, A. (2015). *Income Streams for Data Repositories*. White paper by RDA Interest Group on Publishing Data Cost Recovery for Data Centres.

Downs, R. R., & Chen, R. S. (2013). Towards Sustainable Stewardship of Digital Collections of Scientific Data. Presented at the GSDI World Conference (GSDI 13) Proceedings. Retrieved from https://academiccommons.columbia.edu/catalog/ac:199366

Eakin, L., Friedlander, A., Schonfeld, R., & Choudhury, G. S. (January 1, 1). *A Selective Literature Review on Digital Preservation Sustainability*. Retrieved from http://brtf.sdsc.edu/biblio/Cost_Literature_Review.pdf

Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D., … Calvert, S. (2013). *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Retrieved from http://deepblue.lib.umich.edu/handle/2027.42/97552

Ember, C., & Hanisch, R. (2013). *Sustaining Domain Repositories for Digital Data*. Retrieved from http://datacommunity.icpsr.umich.edu/sites/default/files/WhitePaper_ICPSR_SDRDD_121113.pdf

Erway, R. (2015). What is actually happening out there in terms of institutional data repositories? | hangingtogether.org. *Hanging Together*. Retrieved from http://hangingtogether.org/?p=5342

Esanu, E., Davidson, J., Ross, S., & Anderson, W. (2004). Selection, Appraisal, and Retention of Digital Scientific Data: Highlights of an ERPANET/CODATA Workshop. *Data Science Journal*, *3*(December 2004).

Fearon, D., Gunia, B., Pralle, B., Lake, S., & Sallans, A. (2013). *ARL Spec Kit 334: Research data management services* (No. 9781594079023). Washington, DC: Association of Research Libraries. Retrieved from http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/

Fecher, B., Friesike, S., & Hebing, M. (2015). What Drives Academic Data Sharing? *PLoS ONE*, *10*(2). Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0118053

Finholt, T. A., & Birnholtz, J. P. (2006). If We Build It, Will They Come? The Cultural Challenges of Cyberinfrastructure Development. Retrieved from http://socialmedia.northwestern.edu/files/2013/10/BainbridgeBook.pdf

Fiore, A. T. (2012). Data Quality at Web Scale: Examining Context and Privacy. Presented at the Prepared for the NSF III #1247471 workshop, Curating for Quality: Ensuring Data Quality to Enable New Science. Retrieved from http://dl.acm.org/citation.cfm?id=2582001

Gantz, J. F., McArthur, J., Minton, S., Reinsel, D., Chute, C., Schlichting, W., … Manfrediz, A. (2007). The Expanding Digital Universe [White Paper]. Retrieved from http://www.tobb.org.tr/BilgiHizmetleri/Documents/Raporlar/Expanding_Digital_Universe_IDC_WhitePaper_022507.pdf

Giarlo, M. J. (2012). Academic Libraries as Data Quality Hubs. Presented at the Prepared for the NSF III #1247471 workshop, Curating for Quality: Ensuring Data Quality to Enable New Science. Retrieved from http://dl.acm.org/citation.cfm?id=2582001

Hedstrom, M., & Niu, J. (2008). Incentives for Data Producers to Create “Archive-Ready” Data: Implications for Archives and Records Management. Presented at the Society of American Archivists Research Forum. Retrieved from http://files.archivists.org/conference/2008/researchforum/M-HedstromJ-Niu-SAA-ResearchPaper-2008.pdf

Hendley, T. (1998). *Comparison of Methods & Costs of Digital Preservation*. Cimtech Ltd, University of Hertfordshire. Retrieved from http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley.html

Houghton, J., & Gruen, N. (2014). *Open Research Data Report* (Report to the Australian National Data Service (ANDS)). Retrieved from http://ands.org.au/resource/open-research-data.html

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). *NMC Horizon Report: 2015 Library Edition*. Austin, Texas: The New Media Consortium. Retrieved from http://www.nmc.org/publication/nmc-horizon-report-2015-library-edition/

Kroll, S., & Forsman, R. (2010). *A Slice of Research Life: Information Support for Research in the United States*. Retrieved from http://www.oclc.org/content/dam/research/publications/library/2010/2010-15.pdf

Kuipers, T., & Hoeven, J. van der. (2009). *PARSE.Insight: Insight into Digital Preservation of Research Output in Europe: Survey Report*. Retrieved from http://docplayer.net/127428-Parse-insight-deliverable-d3-4-survey-report-of-research-output-europe-title-of-deliverable-survey-report.html

Kupiainen, I. (Ed.). (2015). RDA/WDS IG Cost Recovery Models. *Research Data Alliance Plenary 5*.
Lavoie, B. (2003). *The Incentives to Preserve Digital Materials: Roles, Scenarios, and Economic Decision-Making*. Dublin, Ohio: OCLC Online Computer Library Center, Inc. Retrieved from http://www.oclc.org/content/dam/research/activities/digipres/incentives-dp.pdf
Lavoie, B. (2006). The Costs of Digital Preservation. In M. Deegan & S. Tanner (Eds.), *Digital Preservation* (pp. 106–132). London, UK: Facet Publishing.
Lavoie, B., & Malpas, C. (2015). *Stewardship of the Evolving Scholarly Record: From the Invisible Hand to Conscious Coordination*. Dublin, Ohio: OCLC Research. Retrieved from http://www.oclc.org/research/publications/2015/oclcresearch-esr-stewardship-2015.html
Lynch, C. (2008, January 1). The Institutional Challenges of Cyberinfrastructure and E-Research. *Educause Review*, *43*(6). Retrieved from http://er.educause.edu/articles/2008/10/the-institutional-challenges-of-cyberinfrastructure-and-eresearch
Lyon, L. (2007). *Dealing with Data: Roles, Rights, Responsibilities and Relationships: Consultancy Report*. University of Bath. Retrieved from http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing_with_data_report-final.pdf
Manyika, J., Byers, A. H., Chui, M., Brown, B., Bughin, J., Dobbs, R., … McKinsey Global Institute. (2011). Big data: The next frontier for innovation, competition, and productivity. Retrieved from http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
Marchionini, G., Lee, C. A., Bowden, H., & Lesk, M. (2012). Curating for Quality: Ensuring Data Quality to Enable New Science, 119. Retrieved from http://dl.acm.org/citation.cfm?id=2582001
McDonough, J. (2012). Start Making Sense: Quality, Context & Meaning. Presented at the Prepared for the NSF III #1247471 workshop, Curating for Quality: Ensuring Data Quality to Enable New Science. Retrieved from http://dl.acm.org/citation.cfm?id=2582001
National Academy of Sciences (NAS). (2009). Ensuring the integrity, accessibility, and stewardship of research data in the digital age, *325*(5939). Retrieved from http://www.nap.edu/catalog.php?record_id=12615
National Research Council (NRC). (2003). *Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences*. Washington, D.C.: National Academies Press. Retrieved from http://www.nap.edu/catalog/10613
National Science Board (NSB). (2005). *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century*. Retrieved from http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf
National Science Foundation (NSF). (2007). Cyberinfrastructure Vision for 21st Century Discovery. *Director*, (March). Retrieved from http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf
Nelson, M. L. (2012). A Plan for Curating “Obsolete Data or Resources.” Presented at the Prepared for the NSF III #1247471 workshop, Curating for Quality: Ensuring Data Quality to Enable New Science. Retrieved from http://dl.acm.org/citation.cfm?id=2582001
Organization for Economic Co-operation and Development (OECD). (2015). *Making Open Science A Reality*. Retrieved from https://www.innovationpolicyplatform.org/content/open-science
Palmer, C., & Palmer, C. (2015, January). *Realizing the Potential of Research Data*. Presented at the Coalition for Networked Information (CNI) Spring 2015 Membership Meeting, Seattle, Washington. Retrieved from https://www.cni.org/news/video-realizing-the-potential-of-research-data/
Pampel, H., & Dallmeier-Tiessen, S. (2014). Open Research Data: From Vision to Practice. In *Opening Science* (pp. 213–224). Cham: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-00026-8_14

Pepe, A., Goodman, A., Muench, A., Crosas, M., & Erdmann, C. (2014). How Do Astronomers Share Data? Reliability and Persistence of Datasets Linked in AAS Publications and a Qualitative Study of Data Practices among US Astronomers. *PLoS ONE*, *9*(8). Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0104798

Pienta, A. M., Alter, G. C., & Lyle, J. A. (2010). The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data. Retrieved from http://deepblue.lib.umich.edu/handle/2027.42/78307

Rauber, A. (2012). Digital Preservation in Data-Driven Science: On the Importance of Process Capture, Preservation and Validation. Proceedings of the 2nd International Workshop on Semantic Digital Archives. Paphos, Cyprus. Retrieved from http://ceur-ws.org/Vol-912/paper0.pdf

Read, K. B., Sheehan, J. R., Huerta, M. F., Knecht, L. S., Mork, J. G., Humphreys, B. L., & Group, N. B. D. A. (2015). Sizing the Problem of Improving Discovery and Access to NIH-Funded Data: A Preliminary Study. *PLoS ONE*, *10*(7). Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0132735

Reeves, J., Brodeur, J. J., Daniels, M. G., Nicholls, N., & Turnator, E. (2015). Libraries and the Research Data Management Landscape. In *The Process of Discovery: The CLIR Postdoctoral Fellowship Program and the Future of the Academy*. Council on Library and Information Resources. Retrieved from http://www.clir.org/pubs/reports/pub167/

Rosenthal, D. S. H., Robertson, T., Lipkis, T., Reich, V., & Morabito, S. (2005). Requirements for Digital Preservation Systems: A Bottom-Up Approach. *D-Lib Magazine*, *11*(11). Retrieved from http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html

Russell, K., & Weinberger, E. (2000). *Cost elements of digital preservation*. Leeds: Cedars Project. Retrieved from https://www.scribd.com/document/7345161/RUSSELL-Kelly-Cost-elements-of-digital-preservation

Science and Technology Council. (2007). The Digital Dilemma: Strategic Issues in Archiving and Accessing Digital Motion Picture Materials. *Academy of Motion Picture Arts and Sciences*. Retrieved from http://www.oscars.org/science-technology/sci-tech-projects/digital-dilemma

Science Staff. (2011, January 1). Challenges and Opportunities. *Science*, *331*(6018). Retrieved from http://science.sciencemag.org/content/331/6018/692

Sturges, P., Bamkin, M., Anders, J. H. S., Hubbard, B., Hussain, A., & Heeley, M. (2015). Research data sharing: developing a stakeholder-driven model for journal policies. *Journal of the Association for Information Science and Technology*. Retrieved from http://eprints.nottingham.ac.uk/id/eprint/3185

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., … Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE*, *6*(6), e21101. https://doi.org/10.1371/journal.pone.0021101

Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., … Dorsett, K. (2015). Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. *PLoS ONE*, *10*(8). Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0134826

Thaesis, & van der Hoeven, J. (2010). *PARSE.Insight: Insight into issues of Permanent Access to the Records of Science in Europe. Final Report*.

Thompson Reuters. (2013). *Unlocking the Value of Research Data: A Report from the Thompson Reuters Industry Forum*. Retrieved from http://researchanalytics.thomsonreuters.com/m/pdfs/1003903-1.pdf

Tsoukala, V., Angelaki, M., Kalaitzi, V., Wessels, B., Price, L., Taylor, M. J., … Wadhwa, K. (2015). *Policy guidelines for open access and data dissemination and preservation*. Policy RECommendations for Open access to research Data in Europe (RECODE). Retrieved from http://recodeproject.eu/wp-content/uploads/2015/02/RECODE-D5.1-POLICY-RECOMMENDATIONS-_FINAL.pdf

Ubaldi, B. (2013). *Open Government Data* (10.1787/5k46bj4f03s7-en). Retrieved from http://www.oecd-ilibrary.org/governance/open-government-data_5k46bj4f03s7-en

University of North Carolina-Chapel Hill (UNC-CH). (2012). *Research Data Stewardship at UNC: Recommendations for Scholarly Practice and Leadership*. Chapel Hill, North Carolina: University of North Carolina Chapel Hill. Retrieved from http://sils.unc.edu/sites/default/files/general/research/UNC_Research_Data_Stewardship_Report.pdf

Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology. *PLoS ONE*, *8*(7). Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0067332

Weigert, V. (2015). Supporting universities to implement effective research data management. *Research Information*. Retrieved from http://www.researchinformation.info/news/news_story.php?news_id=1893

Wynholds, L. A., Wallis, J. C., Borgman, C. L., Sands, A., & Traweek, S. (2012). Data, Data Use, and Scientific Inquiry: Two Case Studies of Data Practices. Presented at the Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM. Retrieved from http://works.bepress.com/cgi/viewcontent.cgi?article=1284&context=borgman

Wynholds, L., Fearon Jr, D. S., Borgman, C. L., & Traweek, S. (2011). When Use Cases Are Not Useful: Data Practices, Astronomy, and Digital Libraries. Presented at the Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries, ACM. Retrieved from http://escholarship.org/uc/item/4tk5d7hx