Social science data archive business models: A historical analysis of change over time

What is a business model? Frameworks from the business literature.

“Business model”: term emerged in the 1990s as internet increased options on how to offer services. Little agreement on what a business model actually consists of. A combination of two ideas: (1) the idea of organizations as a bundle of resources and capabilities that can be arranged in different ways and (2) the question of how best to arrange those resources and capabilities to conduct transactions.

SSDA Business Model Framework (combination of Morris and Osterwald frameworks)

1: How does the data archive create value?
2: For whom does the data archive create value?
3: What is the data archive’s source of competence or internal capacity strengths?
4: How has the archive competitively positioned itself vis-à-vis a variety of alternative ways over time?
5: How does the data archive obtain revenue?
6: What have the time, scope, and size ambitions been of the data archive and its governance bodies?
7: How does the archive create and maintain communications with its users, funders, governing bodies and other stakeholders?
8: What kind of relationships does the data archive establish between itself and its different customer segments/stakeholders?
9: How are work activities arranged and resources organized to support activities?

Business models: Activity in the Data Professions

CESSDA Toolkit: Archive Development Canvas

Brainstorming tool - includes many of the above elements: “use the canvas model how the elements would change if you adopted a different approach.”

European Commission Research Infrastructures Project

(2017, Feb) EU Commission Final Report on Stakeholder Workshop Recommends increased attention to industry as a possible user group/financial supporter, and that the EC provide a framework for long term funding commitments.

(2016, Nov) EU Commission. Stakeholder Workshop Long-Term Sustainability of Research Infrastructures – Exploring RI Full Potential (Brussels) Emphasizes the importance of partnerships with commercial industry as a user and as a supporter.


RDA-WDS Cost Recovery Interest Group Project

(2016, Dillo, Hodson, & Wundt) explored current and potential future alternative income streams via structured phone interviews with representatives from 22 mostly science data repositories globally.

OECD Global Science Forum (GSF) and CODATA

(2016, Geneva) Report of Workshop on Strengthening the Sustainability and Effectiveness of Research Infrastructure Confession about what the term business model means, note that it implies profit, where the mission of infrastructure is not to create profit but to create knowledge.

(2015) Sustainable Business Models for Data Repositories documents current revenue streams, assess attitudes to revenue diversification, explore what new income streams are being trialed.

SciDataCon (2016) Session on Sustainable Business Models for Data Repositories witness statements from 14 repositories.

Knowledge Exchange


Repository: ESSRF (2016, Dillo, Hodson, & Waard) explored current and potential future alternative income streams

Curation for all deposits, then release of “as is” data, then addition of contract services, education classes, educational materials, easy ordering, qualitative expertise, started information studies courses to maintain SSRC funding (1970).

2: For whom does the data archive create value?

ICPSR: Expanded from focus on political science to broader set of researchers. Develop contractor service to government agencies, granting agencies. Heavy educational use in 1970s.

LIS: Income researchers and policy analysis internationally, granting agencies, media, organizations interested in inequity

UKDA: UK government agencies (sought to expand to increase support), expanded from focus on political science to broader set of data, consideration of commercial users, qualitative researchers (’99), pressure to provide services to host university.

4: How has archive positioned itself in relation to alternatives?

ICPSR: Alternatives - free data put up by individual researchers via Internet. Other projects such as repositioning in relation to IPUMS free Census Data

LIS: No direct competitors as no one else does the data harmonization, but people do confuse their microdata with other cross national economic data and worry why do we need microdata?

UKDA: 1970s comparison of UKDA to other data banks to argue UKDA’s value to SSRC; archive remained at Essex in part due to free computing cycles (1980s); Rebuilding of ESRC contracts led to greater coordination via the UK Data Service model.

8: What kind of relationships does the data archive establish between itself and its different customer segments/stakeholders?

ICPSR: Historical formal structures of Council, committees, Organizational Representatives, attend conferences as vendor, summer training, sometimes difficult host university support, sought relationships with EU and American data banks

LIS: Attend major international income statistical conferences, cultivate and maintain contacts in member nations formally (via boards) and informally, host universities support in Lux and NYC

UKDA: Sought relationships with EU and American data banks, data exchanges, effort to extend length of ESRC contracts to 5 year or more, host university support weaker when not seen as prestigious, bringing in resources or providing services to the host.

Change Over Time: ICPSR, UKDA, LIS

1: How does the data archive create value?

ICPSR: Curation for all deposits, then release of “as is” data, then addition of contract services, education classes, educational materials

LIS: Take data from national statistical agencies, harmonize it so that it can be compared, maintain codebooks, offer access, education, production of white papers. Considering serving as platform for other groups’ data collections.

UKDA: Curation for UK government agency datasets (sought to expand), broker international exchange, education classes, educational materials, easy ordering, qualitative expertise, started information studies courses to maintain SSRC funding (1970).

Kristin Eschenfelder, University of Wisconsin-Madison | Kalpana Shankar, University College Dublin | and Rachel Williams and Allison Langham of the University of Wisconsin-Madison

Funding provided by Alfred P. Sloan Foundation, the ASS&T History Fund, the Irish Research Council and the

OECD Global Science Forum (GSF) and CODATA (2016, Geneva) Report of Workshop on Strengthening the Sustainability and Effectiveness of Research Infrastructure Emphasizes how “business model” as essential to keep [research infrastructure] sustainable in the long run.

European Commission. Stakeholder Workshop Long-Term Sustainability of Research Infrastructures – Exploring RI Full Potential (Brussels) Emphasizes the importance of partnerships with commercial industry as a user and as a supporter.

UKDA: UK government agencies (sought to expand to increase support), expanded from focus on political science to broader set of data, consideration of commercial users, qualitative researchers (’99), pressure to provide services to host university.


Repository: ESSRF (2016, Dillo, Hodson, & Waard) explored current and potential future alternative income streams

Curation for all deposits, then release of “as is” data, then addition of contract services, education classes, educational materials, easy ordering, qualitative expertise, started information studies courses to maintain SSRC funding (1970).

2: For whom does the data archive create value?

ICPSR: Expanded from focus on political science to broader set of researchers. Develop contractor service to government agencies, granting agencies. Heavy educational use in 1970s.

LIS: Income researchers and policy analysis internationally, granting agencies, media, organizations interested in inequity

UKDA: UK government agencies (sought to expand to increase support), expanded from focus on political science to broader set of data, consideration of commercial users, qualitative researchers (’99), pressure to provide services to host university.

4: How has archive positioned itself in relation to alternatives?

ICPSR: Alternatives - free data put up by individual researchers via Internet. Other projects such as repositioning in relation to IPUMS free Census Data

LIS: No direct competitors as no one else does the data harmonization, but people do confuse their microdata with other cross national economic data and worry why do we need microdata?

UKDA: 1970s comparison of UKDA to other data banks to argue UKDA’s value to SSRC; archive remained at Essex in part due to free computing cycles (1980s); Rebuilding of ESRC contracts led to greater coordination via the UK Data Service model.

8: What kind of relationships does the data archive establish between itself and its different customer segments/stakeholders?

ICPSR: Historical formal structures of Council, committees, Organizational Representatives, attend conferences as vendor, summer training, sometimes difficult host university support, sought relationships with EU and American data banks

LIS: Attend major international income statistical conferences, cultivate and maintain contacts in member nations formally (via boards) and informally, host universities support in Lux and NYC

UKDA: Sought relationships with EU and American data banks, data exchanges, effort to extend length of ESRC contracts to 5 year or more, host university support weaker when not seen as prestigious, bringing in resources or providing services to the host.

Change Over Time: ICPSR, UKDA, LIS

1: How does the data archive create value?

ICPSR: Curation for all deposits, then release of “as is” data, then addition of contract services, education classes, educational materials

LIS: Take data from national statistical agencies, harmonize it so that it can be compared, maintain codebooks, offer access, education, production of white papers. Considering serving as platform for other groups’ data collections.

UKDA: Curation for UK government agency datasets (sought to expand), broker international exchange, education classes, educational materials, easy ordering, qualitative expertise, started information studies courses to maintain SSRC funding (1970).

Kristin Eschenfelder, University of Wisconsin-Madison | Kalpana Shankar, University College Dublin | and Rachel Williams and Allison Langham of the University of Wisconsin-Madison

Funding provided by Alfred P. Sloan Foundation, the ASS&T History Fund, the Irish Research Council and the