Data Perspectives Beyond Alliances
Side Event to the RDA Plenary, Tokyo
3 March 2016

Sustainable Business Models for Data Repositories

Dr Simon Hodson
Executive Director, CODATA
www.codata.org

The Challenge: Sustainable Business Models for Data Repositories

- Research funder policies – quite rightly – mandate data stewardship.
  - OECD Principles and Guidelines, 2007
  - G8 Science Ministers Statement, 2013
  - Major funders in US, UK, EC Horizon 2020 data policy etc.
- Increasing need for data repositories and data stewardship.
  - Increasing volume presents a challenge.
  - Requirements for stewardship present a greater challenge.
- **Sustaining digital data infrastructure is a major issue for science policy!**
- Genuine concern that current funding models will prove inelastic and not meet the growing requirements – concern on the part of repositories and funders.
- Witnessing Innovation
  - Changes in funding / business models (ADS, DANS, ICPSR)
  - Innovative business models (Dryad, FigShare)

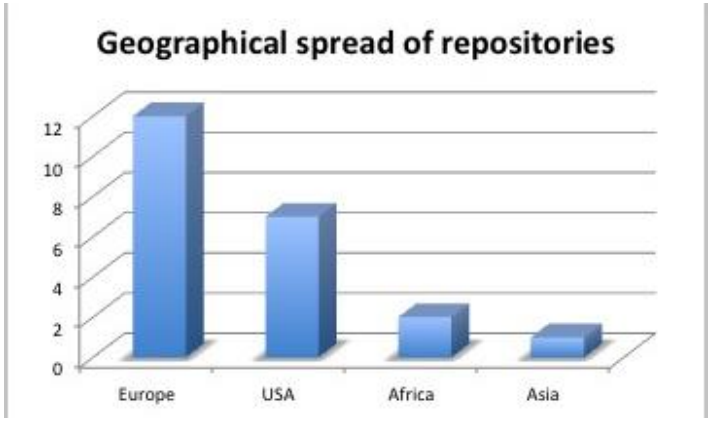
The Challenge: Sustainable Business Models for Data Repositories

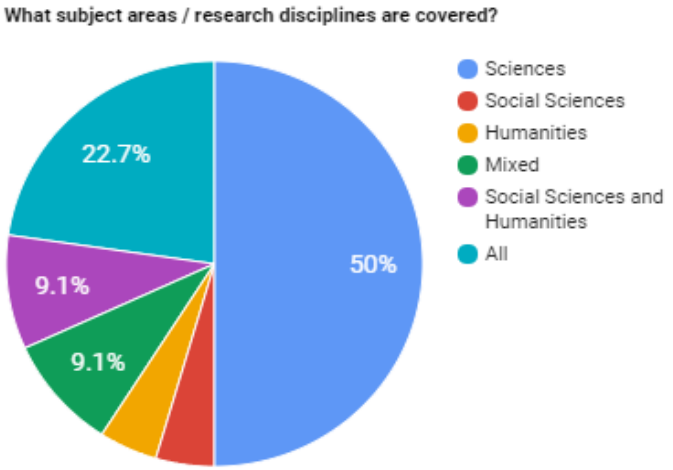
- Policy agreement that the cost of data stewardship is an essential, integral part of the cost of doing research.
- Strong value proposition for data infrastructure and data sharing.
  - CODATA White Paper for GEO: *The Value of Open Data Sharing*: http://dx.doi.org/10.5281/zenodo.33830
- Very little work has been done on the economics and business models of data infrastructure.
  - Blue Ribbon Task Group Report on Sustainable Digital Preservation: http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf
  - Sustaining Domain Repositories for Digital Data: A White Paper (ICPSR): http://datacommunity.icpsr.umich.edu/sites/default/files/WhitePaper_ICPSR_SDRDD_121113.pdf
- Pressing need for work on who pays and how: analysis of income streams, of innovative funding models, of willingness to pay and responsibilities, of business models in general.
- **OECD Global Science Forum is the ideal setting for this work.**

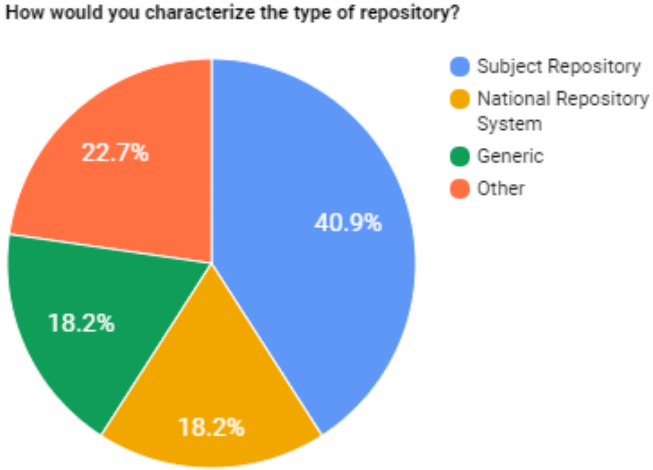
Previous Work on Income Streams/Business Models

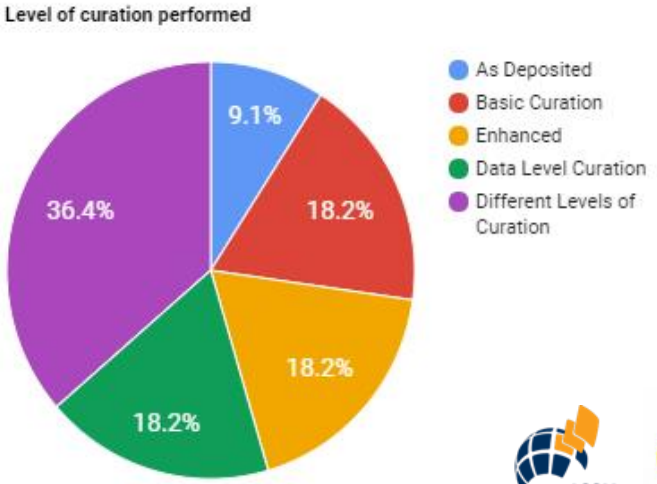
- RDA-WDS WG Draft Report: http://bit.ly/income-streams-draft-P6
- Co-Chairs:
  - Simon Hodson, Executive Director of CODATA
  - Ingrid Dillo, Deputy Director of DANS, WDS SC, RDA TAB
  - Anita de Waard, Elsevier Research
- Landscape survey of 25 data repositories.
- Identified major income streams and funding structure.
- Typology of business models.
- SWOT analysis at RDA workshop in September.

Typology of repositories surveyed:

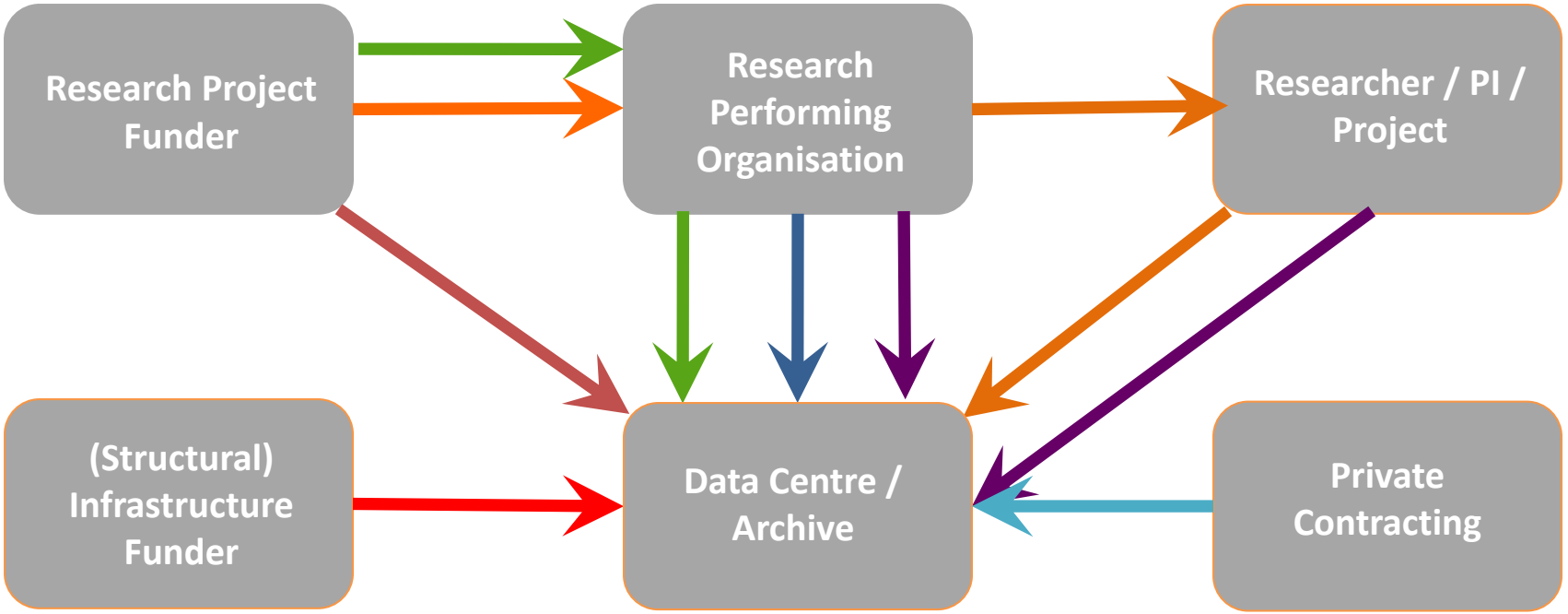






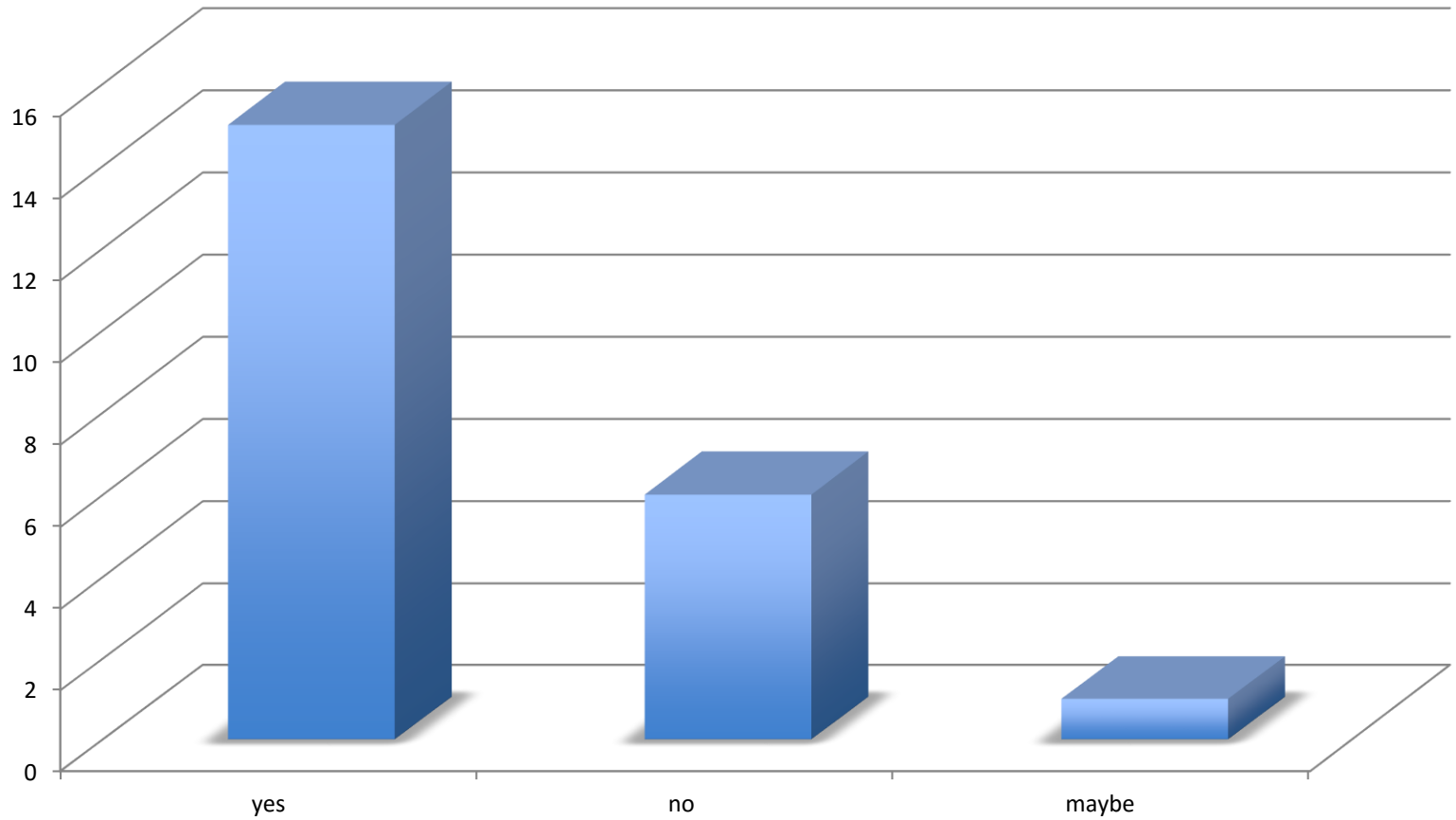


Typology of income streams



1. Structural (central contract)
2. Hosting Support (indirect or direct support through institutional hosting)
3. Annual Contract (from depositing institution)
4. Data Deposit Fee (may be paid by researcher, RPO or publisher; may originate with funder)
5. Access Charge (for the data or for value-adding services)
6. R&D Projects (to develop infrastructure or value-adding services)
7. Private Contracting (services to parties other than core funder)

Exploring Alternative Income Streams



Alternative Income Streams Under Consideration

- Contracts for specific services offered (hosting, archiving, curation)
- Expanding the number of affiliated institutions (services, member benefits)
- Deposit fees
- Increasing core structural funding (given priority for data)
- Charging for value added data or services
- Specific services for the commercial sector
- Sponsorship
- More services for national memory institutions

Typology of Business Models

1. **Largely structurally funded**
2. **Reliant on data access charges or membership fees**
3. **Exploring data deposit fees**
4. **Substantial diversification**

- *Propped up by project funding*
- *Supported by host institution*

ICSU WORLD DATA SYSTEM

RDA RESEARCH DATA ALLIANCE

1: Structural Funding

STRENGTHS

- Longer-term stability: easier planning and achieve efficiency
- Stronger commitments and communication with stakeholders
- Larger chunk of investments can cover operational costs
- Up front funding can help plan budget and build effective organisation
- Immune to marketing and collateral effects
- No need to spend too much time fundraising

WEAKNESSES

- If only renumeration for capital, this is a risk
- Fixed funding is a weakness wrt the context of (immensely) growing volumes of data
- Can reduce the efficiency; no incentive to improve; long evaluation cycles make you lazy!
- Inflexibility of funding, can't adapt easily

OPPORTUNITIES

- Data is hot and funders are more amenable to provide structural funding
- Riding the hype and gaining structural funding can help raise the profile of institutions (win-win)
- Funders have increasing budget for infrastructure
- Data is/can be recognized as infrastructure
- Institutions (universities, RPO, etc.) recognize their responsibility over funding the data infrastructure

THREATS

- "Today it's hot, tomorrow it's not!"
- Not receiving structural funding because of big national initiatives with which you are not aligned
- Increase demand cannot be handled easily
- Not in control of your funding – dependent on small nr of sources
- Funder itself may be descoped (e.g. US)

2: Data Access Charges

STRENGTHS	WEAKNESSES
• Consumer pays for what consumer wants • Cost scales with access if cloud-based storage • Income reflects value of data product • Stable and predictable income from membership / subscription fees; income continues in the long-term • Loyalty of community of members • Members have influence over priorities • Model applicable to similar/related services such as DataCite • Model can accommodate licensing flexibility	• Data worth serving in the long term might not be available • Causes competition amongst repositories for high impact data; data poaching • Not affordable to people at underprivileged organizations; unequal access to what should be a public good • Could lead to data purging • Must think carefully about licensing terms and conditions
OPPORTUNITIES	**THREATS**
• Monitor demand, change/improve services • Create free access for basic products, build consumer base and create demand for value-added services • Most market-oriented approach • Valuable for private sector • Creates captive market	• Vulnerable to economic downturns – although there is actual increased use of archival data when primary funding is down • Expectation is for free access, other providers might undermine business

ICSU WORLD DATA SYSTEM

RDA RESEARCH DATA ALLIANCE

3: Data Deposit Charges

STRENGTHS	WEAKNESSES
Puts charge on data producer (works well with grant funding) OA compatible Scalable Closely linked to the research community – responsive to science need Competition Neutral to value of data to end users (no a priori value judgment) Potentially fair/proportional distribution of funding	Defining the cost (POSF) Does it meet the challenge of diverse data types Market weakness vs structurally funded repositories Administrative overheads Neutral to value of data to end users (data centre has to accept all paid data)
OPPORTUNITIES	**THREATS**
Autonomous generation of revenue Scaled deposit fee model Compatible with subscription as part of business model	PI pushback (vs top-slicing research grant) Rush to cheapest option? Needs very clear policy framework High cost will put off depositors Hostage to future storage and preservation costs Infrastructure costs are estimated too low

ICSU WORLD DATA SYSTEM

RDA RESEARCH DATA ALLIANCE

4: Diversification

STRENGTHS • No single source of failure • Flexibility to experiment with new services and markets • Stimulates innovation • Focuses attention on value to users	WEAKNESSES • Access fees exclude users/limit uses • Funding is short term; obligations long term • Sponsor priorities change • High administrative overhead • Requires highly skilled staff • Host universities are not stakeholders of national repositories • Sustainability of funded projects • Draws attention away from core mission

OPPORTUNITIES
• Research funding is project based
• Data management requirements are creating demand from researchers for services during the project funding
• Sponsor priorities change
	THREATS
	• Competition
	  • Commercial companies
	  • Institutional repositories
	• Variability of funding

Some Conclusions

- Structural funding supports c.50% of repositories surveyed.
- Structural funding suits many repositories although often supplemented and some concerns expressed about flexibility and adaptability.
- Many repositories are interested in charging for value-added services, but very little current exploration of this possibility.
- Data deposit fees are being explored by a small number of repositories.
- Data deposit fees may gain stakeholder acceptance because of similarity to APCs, but concern about administrative overheads and that encourage cheaper, lower levels of curation.
- Many data repositories value participation in research and R&D projects, but many are highly dependent on this income and overheads need to be considered.
- Need for further analysis of stakeholder acceptance of business models and income streams, in addition to:
  - Analysis of innovative income streams;
  - Analysis of means of restraining / mitigating costs.

Sustainable Business Models for Data Repositories

- Clear need for work on sustainable business models.
- Firmly within strategic priorities and role of OECD Global Science Forum.
- Builds on substantial initial work by the RDA-WDS Working Group.
- Analysis of innovative income streams and policy recommendations on sustainable business models can make a substantial, concrete and specific contribution to addressing the challenge.

The Project: Sustainable Business Models for Data Repositories

- Questions to address:
  1. How are data repositories currently funded?
  2. What innovative income streams are available? What means of restraining costs are available?
  3. How do income streams match willingness/ability to pay of various stakeholders?
  4. How do income streams/willingness to pay fit together into a **sustainable** business model?

- Builds on existing work of RDA-WDS Working Group.
- Broader landscape study of current funding models.
- Focus group on innovative income streams and cost restraint.
- Economic analysis of business models.
- Test business models with stakeholder groups.
- Policy recommendations based on concrete business model options.

Impact and Adoption: OECD GSF Project on Business Models

- **International Expert Group:** comprising nominees from GSF delegates and from CODATA, RDA and WDS. First meeting in April or June.

- **Workshop 1:** to identify and appraise possible innovative income streams; and to identify approaches to cost restraint.

- **Workshop 2:** to test possible business models with stakeholders

The Project: Sustainable Business Models for Data Repositories

- **Q1 April-June 2016:** Project set up; Expert Group virtual meetings; data repository interviews, inc. those identified by GSF.
- **Q2 July-Sept 2016:** Complete income streams landscape survey; focus group on innovative income streams; develop economic analysis of business models.
- **Q3 Oct-Dec 2016:** Stakeholder Workshop (inc. GSF) on Business Models.
- **Q4 Jan-March 2017:** Iterate draft report and recommendations with Expert Group.
- **March/April 2017:** final report with recommendations on sustainable business models presented to GSF for final approval.

- **Expert Group:** comprising nominees from GSF delegates and from CODATA, RDA and WDS.
- **Consultant:** draft texts, facilitate workshop.
- **Economics Consultant:** key role in preparing analysis of business models.
- **Workshops** 1) on innovative income streams and 2) to test business models.

OECD BETTER POLICIES FOR BETTER LIVES

ICSU WORLD DATA SYSTEM

RDA RESEARCH DATA ALLIANCE

Thank you for your attention!

Credit for contributions to slides: Ingrid Dillo, Anita de Waard.

Simon Hodson
Executive Director CODATA
www.codata.org
http://lists.codata.org/mailman/listinfo/codata-international_lists.codata.org
Email: simon@codata.org
Twitter: @simonhodson99
Tel (Office): +33 1 45 25 04 96 | Tel (Cell): +33 6 86 30 42 59

CODATA (ICSU Committee on Data for Science and Technology), 5 rue Auguste Vacquerie, 75016 Paris, FRANCE