



2008 Annual Meeting Report

Pascal Heus / Arofan Gregory / Rob Grim
February 2009

Table of Contents

Overview	2
Agenda	2
Financial Report.....	2
Budget 2009	3
2008 Activity Report.....	3
Events.....	3
Projects.....	4
Publications	5
Review of 2007 resolutions	5
ODaF North America Meeting.....	6
SDMX	7
GLOSS / Web 2.0 Technologies for Social Science.....	8
ISO/IEC 11179	9
Beyond 20/20	10
Space-Time Research.....	10
DDI North America Users' Group	11
Open Source for DDI/SDMX.....	11
Other projects.....	12
Collaboration with Carleton College	12
Standards Mapping Project	12
Personnel changes	12
Directors	12
Advisors.....	12
Management Team	12
2009 Activities and issues.....	14
Management issues	14
2009 Resolutions	16
Annex A: ODaF Overview	18

Overview

The 2008 Open Data Foundation Annual meeting was held at the National Opinion Research Center office in Bethesda, MD, USA on Friday December 5th and attended in person or via conference call by the following participants:

Director: Ernie Boyko

Advisors: San Cannon, Dan Gillman, Wendy Watkins, Wendy Thomas

Managers: Arofan Gregory, Pascal Heus

Observers: Tim Mulcahy (NORC)

Agenda

The agenda was finalized as follows at the meeting as follows

- I. Financial Report
- II. Report on Ongoing Projects and Activities
 - a. Events
 - b. Projects
 - c. Publications
 - d. Review of 2007 resolutions
- III. Report on ODaF North America meeting
 - a. Presentations / discussions
 - b. Potential projects and activities
- IV. Personal change
- V. Activities and management issues for 2009

Financial Report

Income	
Services	60,661.60
Total Income	\$ 60,661.60
Expenses	
Bank Charges	773.01
Dues & Subscriptions	2,500.00
Equipment Rental	2,639.76
Legal & Professional Fees	40,051.60
Meals and Entertainment	2,276.82
Office Expenses	86.90
Promotional	1,550.00
Travel	6,255.58
Utilities	146.11
Total Expenses	\$56,279.78
Net Operating Income	\$4,381.82
Net Income	\$4,381.82

In addition to the above income, we estimate that about \$75K have been contributed in kind from Metadata Technology, Ideas2evidence and other supporting organizations (based on a \$1K/day).

Expenses incurred total about \$56K and primarily consist in conference participation, event organizations, IT infrastructure, membership in the DDI Alliance, and subcontracting on project work (after 15% overhead going to ODaF).

The following income is anticipated early 2009: NORC DDI/SDMX workshop (\$1.5K after travel expenses and fees), NORC Enclave 4Q (\$1.5K), Canada RDC expertise (\$4K after travel expenses, no fee charged).

Our current cash balance is about \$8.5K.

Project / activity income

DIW Berlin	2,466.20
NORC Data Enclave	38,000.00
UKDA DExT	20,195.40
TOTAL	\$60,661.60

Supplier expenses

AplusNet (2 servers)	2,639.76
Arofan Gregory (event refunds)	2,445.91
DDI Alliance Membership	2,500.00
DExT Consultant (Joachim Wackerow)	2,937.42
Intuit (Accounting)	249.50
Metadata Technology Ltd.	36,929.53
Pascal Heus (event refunds)	2,034.27
TOTAL	\$49,736.39

Budget 2009

As anticipated last year, the Foundation continues to operate on a fairly small budget (\$60K) and in kind contributions from the management team. 2009 however will present new challenges as the NORC project combined with small contracts will not be sufficient to cover the operating expenses that are also expected to grow as our regional events grow in size. We will therefore need to identify new sources of funding through projects or contributions.

2008 Activity Report

Events

We started the year with presentations in March at the FedCASIC workshops (Washington, DC) as well as the METIS 2008 session in Luxembourg¹. We then had a well attended and very successful first ODaF Europe meeting hosted at the UK Data Archive in Colchester in April². We moved on to sponsor the IASSIST 2008 conference held in Palo Alto last May where we made several presentations. In June, a workshop

¹ <http://www.unece.org/stats/documents/2008.04.metis.htm>

² http://opendatafoundation.org/events/odaf_europe_2008.php

was organized in Wiesbaden where we presented to research data centers and the statistical office on the differences and synergies between DDI and SDMX (a major debate in Germany). In August, Julia Lane and Tim Mulcahy made a number of presentations related to the NORC data enclave at the Joint Statistical Meetings in Denver that included contributions from ODaF³.

In September we attended a 2-day meeting organized by the Canadian Research Data Center in Winnipeg, as technical experts, to discuss metadata mining and data comparability and prepare a report for planning phase 2 of the RDC metadata integration project. The event was attended by other RDCs from Canada, US and Europe. In October, we made a short presentation on Metadata, DDI, SDMX and ODaF at the monthly FedStats meeting⁴.

In November, ODaF provided assistance for a 5-day DDI 3 workshop⁵ as well as assistance and participation in a 5-day DDI Expert Working Meeting⁶ to develop best practices the following week. Both events were held at Schloss Dagstuhl, a Leibniz Center for Informatics in Germany.

In December, we presented at a 1-day NORC sponsored metadata workshop on DDI / SDMX⁷ held at their Bethesda, MD office on Dec 2nd. Intended as a general introduction for data producers and archivists, it was attended by over 60 participants from various federal and private agencies. This event was a huge success and will likely be repeated as we had to turn down participants. The same week, we also held at NORC Bethesda the first ODaF North America⁸ meeting on Dec 3rd and 4th (more on this below) along with our annual meeting⁹ on Dec 5th.

Finally, we have secured a spot for the ODaF Europe 2009 conference that will be hosted by the Institute for the Study of Labor (IZA) in Bonn the first week of April. This will be a 2-day event, exact dates to be determined.

Projects

Early in the year we successfully completed the DEXT project¹⁰ with the UK Data Archive who donated the resulting products to the DDI Foundation Tools Program (DDI-FTP). Our other main ongoing project is our contribution to the NORC the Data Enclave¹¹, which is very successful and quickly growing in importance and visibility. We also continue to support and contribute to the DDI Foundation Tools Program¹², which is expanding to include the collaborative development of an open-source DDI 3.0 editing suite. Work on the SDMX Browser project has been slower than hoped, but we have a working demo and hope to release a beta in the near future. The release has been

³ <http://www.amstat.org/meetings/jsm/2008/>

⁴ <http://www.fedstats.gov/>

⁵ <http://www.dagstuhl.de/en/program/calendar/evhp/?semnr=08452>

⁶ <http://www.dagstuhl.de/en/program/calendar/evhp/?semnr=08462>

⁷ <http://www.norc.org/news/metadata+workshop+-+the+next+frontier+in+documenting+survey+data.htm>

⁸ http://www.opendatafoundation.org/events/odaf_northamerica_2008.php

⁹ http://www.opendatafoundation.org/events/odaf_annual_meetings_2008.php

¹⁰ <http://tools.ddialliance.org/?lvl1=product&lvl2=dext>

¹¹ <http://dataenclave.norc.org>

¹² <http://tools.ddialliance.org/?lvl1=ftp>

delayed in order to integrate an open-source visualization package provided by the ECB and the Federal Reserve Bank of New York.

Publications

On the publication side, besides the various conference presentations, our joint paper on "Data Documentation Initiative: Toward a Standard for the Social Sciences (Mary Vardigan, Pascal Heus, Wendy Thomas)" was presented last December at the DCC conference in Washington DC was recently published in the International Journal of Digital Curation (<http://www.ijdc.net/ijdc/article/view/66>). Thanks to Julia Lane, we have co-authored a paper titled "Data Access in a Cyber World: Making Use of Cyberinfrastructure (Julia Lane, Tim Mulcahy, Pascal Heus)"¹³ that was recently published in the first issue of Transactions on Data Privacy¹⁴.

In addition to papers, we released in February a first version of the "Managing social, behavioral and economic data and metadata: Guidelines for Tools Development and Recommendations for Operating Environment"¹⁵, a collection of recommendations and resources for the implementers of open source metadata management software products. This document is used as a reference for the DDI Foundation Tools Program, the initial phase of the Canada Research Data Center metadata project and possibly others.

We have also been contracted to contribute a section on "Metadata" for a new book titled "Developing the Research Infrastructure for the Social Sciences (including the Behavioral Sciences) in Germany and Beyond: Progress since 2001, Current State, and Future Demands" to be published in Germany next year. The publication will be a collection of short contributions from several authors.

Additionally, we are working on the standards-mapping projects, which are moving forward nicely given the soon-to-come submission of SDMX v. 2.0 to ISO (the alignment of ISO standards with ISO/IEC 11179 and others is always a point of interest here.)

As discussed at the annual meeting last year, we have drafted a high level position paper for the Open Data Foundation intended for a general public / non-technical audience (available in annex of this document).

Review of 2007 resolutions

1. Brian Fitzpatrick should become an executive manager for ODaF to replace Chris Nelson.
→ See 2008 Personnel changes on (p. 12)
2. A secure area on the ODaF web servers should be established to share internal documents and presentations amongst directors, management, advisors and members.
→ This was not completed but is not seen as necessary as we seldom have this need

¹³ <http://www.opendatafoundation.org/?lvl1=resources&lvl2=papers>

¹⁴ <http://www.tdp.cat/issues/abs.a002a08.php>

¹⁵ http://www.opendatafoundation.org/ODaF_tools_developer_guidelines.pdf

3. ODaF managers should be authorized to contractually engage the organization in any activity related to its mission as long as such activity is of limited scope and duration and does not fall in the category of a Project. This is primarily to cover advocacy, training or short-term technical advisory work.
→ *Completed*
4. A proposal should be submitted to the directors for a collaborative project with the Joint Research Centre of the European Commission for the use of SDMX for the dissemination of data on protected areas.
→ *This project was cancelled by lack of interest from JRC*
5. A project for the development of an open source ISO 11179 registry implementation if sufficient resources can be identified, and assuming that the ISO standard XML bindings are in a sufficiently mature state.
→ *This was further discussed at the North America meeting and will be turned into a full project proposal*
6. A proposal for the establishment of a virtual community to support SDMX Users should be submitted.
→ *still pending, based on discussions within SDMX, possibly a 2009 project*
7. A proposal for a project to promote DDI 3.0 amongst communities of questionnaire design and question management tools developers should be submitted.
→ *This was not completed and should be passed along to the DDI User and Outreach Group*
8. ODaF should be required to produce and publish at least 5 white papers in 2008. → *Completed*
9. A regional European ODaF meeting should be organized in Germany in the Spring of 2008, at the IZA offices if possible.
→ *Completed*
10. A regional ODaF meeting in the US should be organized in the Fall of 2008.
→ *Completed*

ODaF North America Meeting

The first ODaF North America meeting was held at the NORC Bethesda Office on Dec 3rd and 4th. It was attended by the following participants:

- Dan Gillman (Bureau of Labor Statistics)
- Nikos Askitas (Institute for the Study of Labor, IZA, Germany)
- San Cannon (Federal Reserve Board)
- Kenneth Gross (Federal Reserve Board)
- Kiet Bang (Penn State - Population Research Institute)
- Brian Garrett (Space-Time Research)
- Bill Block (Cornell Institute for Social and Economic Research)
- Jeremy Williams (Cornell Institute for Social and Economic Research)
- Liang Long (US Dept. of Transportation)
- Wendy Thomas (Minnesota Population Center)
- Svetlozar Nestorov (U. of Chicago, Computational Institute)
- Arofan Gregory (ODaF)
- Pascal Heus (ODaF)

The agenda was adjusted based on the participants' interest. The following topics were discussed:

- SDMX (San Cannon / Kenneth Gross)
- GLOSS and Web 2.0 technologies (Svetlozar Nestorov, U Chicago)

- ISO/IEC 11179 (Dan Gillman)
- Beyond 20/20 (Liang Loon)
- Space-Time Research (Brian Garret)
- DDI North America Users' Groups (Bill Block)
- Open Source products for DDI/SDMX (Bill Block, Jeremy Williams)

SDMX

San Cannon and Kenneth Gross from the Federal Reserve Board made a short presentation on SDMX and ongoing implementation work ongoing at the FRB. Follow-up discussions focused on how ODaF could be involved or contribute to the adoption of the SDMX standard in general. The following ideas emerged as potential project proposals:

Host SDMX registry for the sponsors

To foster the adoption of the SDMX and the visibility of existing data, we would be to propose to the SDMX sponsors to setup and maintain on their behalf an SDMX registry for their public data/metadata. Several of them do publish SDMX data available on the web but have not actually deployed a registry based infrastructure as this is often challenging from the institutional and technological perspectives. ODaF has more flexibility when it comes to hosting such a solution that could be based on available open source products.

If we move forward on this, several aspects will need to be addressed including:

- Funding mechanisms: implementation, maintenance
- Branding of the product (ODaF branding is not required)
- Do we only register structural metadata or also the data?
- Liability issues: insurance, service level agreements, etc.

It would also be important to clearly explain the benefit of such a service to the sponsors.

We feel the right time to extend such an offer to the SDMX sponsor community would be after the vendor meeting that will take place in Paris in January 2009. This will be followed-up upon in February by Pascal, Arofan and San.

SDMX Startup Guides

Documentation on the SDMX standard has mainly been produced with the ISO certification in mind and, while complete, is not geared towards newcomers or potential adopters. We feel there is a strong need for a suite of User's Guides, targeted towards different audiences such as SDMX data producer, consumer or solution provider. Such effort could be undertaken by ODaF and other partners

Our initial effort would focus on fostering the adoption of SDMX within the United States whose federated statistical system and research communities would greatly benefit from the adoption of metadata standards such as SDMX. Such initiative also seem to fit particularly well in the context on the recently published Science of "Science Policy: A Federal Research Roadmap"¹⁶ by the Office of Science and Technology Policy. We do believe funding for such project could be supported by the National Science Foundation or the Office for Management and Budget, but could also move forward initially as a volunteer effort.

¹⁶ <http://scienceofsciencepolicy.net>

Our next steps will therefore be to identify potential contributors, partners and funding sources.

SDMX Users Forums

The idea to have a public SDMX User Forum has been discussed before and remains an attractive project. The only available space for SDMX users at this time is the SDMX Tools Forum hosted by Metadata Technology. A more global public forum that focuses on a broader set of issues would be beneficial to the community.

To fit this in a framework, we could use two approaches:

- Focus on the United States and deploy such forum in the context of FedStats, the Federal Government portal for official statistics
- Wait for the official approval of SDMX 2.0 as ISO standard (expected Fall 2009) and launch the forum

In order to turn this idea into an ODaF project proposal, we would also first need to identify project sponsors and the right technologies and tools that meet the community needs. One consideration is the willingness of the SDMX sponsors to lend their support to such a forum. Investigations into this possibility will be undertaken by Arofan and San.

GLOSS / Web 2.0 Technologies for Social Science

The NORC sponsored DDI / SDMX workshop gave us the opportunity to meet with Svetlozar Nestorov from the Computational Institute of the University of Chicago¹⁷ where he leads the GLOSS project¹⁸, an initiative that leverages upon collaborative tagging¹⁹ to collect informal researcher metadata at the survey variable level. GLOSS is an interesting and innovating approach that leverages on Web 2.0²⁰ technology to capture community knowledge for social science at the microdata variable level.

As we familiarized ourselves with the tool through conversations with Svetlozar and during the workshop presentations, we saw an opportunity for bringing together GLOSS with structured DDI metadata that would benefit both initiatives: it would help the GLOSS product to provide better metadata to its end user and the tool in turn could be used to collect information from users in Web-based collaborative spaces and add to the survey metadata knowledge. We therefore invited Svetlozar to attend the ODaF regional meeting to further discuss potential collaborative efforts.

During the meeting, we recognized the win-win aspect of such effort and the following ideas were put forward:

- Use DDI to feed metadata to GLOSS and generate the user interface with high quality metadata
- Populate DDI through the Web 2.0 tagging interface to collect user-generated metadata for concepts (based on controlled vocabularies) or keywords (free entry) at various levels (variable, file, study, etc.)
- Use similar interfaces for communal metadata translation (multilingual)

¹⁷ <http://www.ci.uchicago.edu/>

¹⁸ <http://gloss.ci.uchicago.edu/demo>

¹⁹ <http://en.wikipedia.org/wiki/Folksonomy>

²⁰ http://en.wikipedia.org/wiki/Web_2.0

Several challenges were also identified such as:

- The need for a critical mass of user for such system to become practically useful
- Metadata quality control and review processes
- Meaning and definition of terms

At our annual meeting, we therefore decided to further explore the idea of initiating a collaborative research effort with the University of Chicago Computational Institute. Other potential project partners are the National Opinion Research Center (NORC) and the Bureau of Labor Statistics (BLS). Funding for such initiative could be provided through NSF or ICPSR. We will follow-up on this with Svetlozar and other partners with the idea to turn this into a project proposal the 2 quarter of the year.

ISO/IEC 11179

Dan Gillman from the Bureau of Labor Statistics provided us with an update on the ISO/IEC 11179 standard²¹. ISO/IEC 11179 is a high-visibility standard, titled Metadata Registries, which focuses on the semantics and representation of data. It is composed of 6 different parts (Framework, Classification, Basic attributes and metamodel, Definitions, Naming and identification, Registration) whose purposes are the management of semantics and representations of data, the registration and administration of descriptions, and the classification of items based on pre-defined schemes. The 2nd Edition of the standard was published in 2005 and makes strong use of UML²². While a good model, it suffers from a few weaknesses, a major one being the entire model must be implemented in order to conform to the standard.

The 3rd edition of the standard is underway, and one of the two approaches under consideration aims at a significant refactoring of the specification. Edition 2 of the standard presents one complete UML model in Part 3. This new approach would split the model among all 6 of the parts, so each part would contain its own model targeted to the scope of the part. These models would be able to stand alone, any set of them being able to work together, and conformance would be defined separately for each of the parts.

We do believe such an approach to be more appropriate and would like to prototype a proof of concept registry that implements the proposed 3rd edition modifications. We feel a limited amount of resource would be required to do this as we can actually implement this using the metadata report feature of the existing SDMX registry freeware. Demonstrating an actual live implementation of ISO/IEC11179 3rd edition would provide a significant boost for the standard and the proposal as well as high visibility for ODaF. Initial partners for this project would be ODaF and the BLS that could provide development resources. Other interested agencies could include NORC and Statistics Canada.

Our next steps will be to formalize this as a project proposal to the directors. This would be implemented on a simple SDMX based registry with the option later on to implement the standard ISO interfaces now under development.

²¹ http://en.wikipedia.org/wiki/ISO/IEC_11179

²² http://en.wikipedia.org/wiki/Unified_Modeling_Language

Beyond 20/20

The issue of working with Beyond 20/20 cubes was initially raised at the ODaF Europe regional meeting as many data archives have to work with data in this proprietary format. This was brought up during the North America meeting as well and briefly discussed. The main question regards the conversion of Beyond 20/20 cubes into archive neutral formats such as SDMX for preservation, exchange and further processing. We have not at this time determined a concrete proposal to address this or decided if ODaF should play a role in this. We will for now continue to look for interested parties and if a sufficient number of agencies/individuals come together around this, we could turn it into a project proposal.

Space-Time Research

Space-Time Research²³ (STR) is an Australian software company specializing in self-service business intelligence for government agencies. STR provides a powerful suite of software products for organizations needing easier, faster, and safer analytics, business intelligence and reporting capabilities. Their software is used by a number of National Statistical Organizations for tabulation and dissemination of Census data as well as survey information. We initially met company representatives through the NORC Data Enclave project and were overall quite impressed by the functionality. Compared to other business intelligence packages or Web-based data delivery products, STR has a deep understanding of common microdata issues faced by producers or providers (such as anonymization, high-performance on-the-fly analysis, complex data manipulation options, data quality assurance, safe guarding users against inappropriate uses of the data, etc.).

To achieve the impressive performance, STR creates their own analytical data store tools to persist data and metadata. The process could be enhanced by using DDI as it could contain much of the information they need to load the microdata and related metadata into their software. Producers or archives using DDI, who have already invested efforts in documenting their survey, could benefit greatly by some type of integration. If STR software would be compliant and understand the DDI standard, the ETL²⁴ process could be significantly automated. The company was receptive to the idea and we followed up with them through an invitation to the NORC sponsored DDI/SDMX workshop as well as the ODaF meeting for further discussion. This was motivated by the fact that the NORC Data Enclave²⁵ project is considering STR products for use within the enclave as an add-on tool for researcher and to facilitate the disclosure review process. Furthermore, we strongly support the adoption of standards by software vendors as we see this as a key component for their broad adoption.

After further review of the STR products and discussion with technical representative, we see the software as not only a tool to ingest microdata using DDI but one that could also be used at various stages of the production process and for the delivery of tabulated and aggregated data. For these functions, the tool could then also produce data and metadata in DDI and SDMX formats. The process of tabulating microdata and publishing aggregates is a very common one within national statistical organizations, which is a primary market for STR.

²³ <http://www.spacetime-research.com/>

²⁴ http://en.wikipedia.org/wiki/Extract,_transform,_load

²⁵ <http://www.norc.org/DataEnclave>

While ODaF has no interest in promoting specific commercial products, the recognition of standards such as DDI and SDMX by software vendors is an important to foster their adoption and we are interested in supporting such efforts by vendors such as STR willing to invest in the standards. The role of the Foundation in this case would be to provide technical expertise and assistance to the company for the understanding of the DDI and SDMX standards and their integration in the STR products. This may involve signing a non-disclosure agreement with the company. Targeted functionalities would include facilitating the ETL process using DDI, potential use of STR tools as cube builders (DDI and/or SDMX), supporting SDMX as an output format, and provide better support for time-series data. Our next steps will be to further discuss options with STR and make a formal joint research proposal to our Directors.

DDI North America Users' Group

Following in the footsteps of Europe, Bill Block from the Cornell Institute for Social and Economic Research is coordinating efforts towards the establishment of a DDI North America Users Group. As Nikos Askitas from IZA who is coordinating the European initiative was in the Washington DC area during our regional meeting, we got together to discuss lessons learned from the initial stages in Europe, ensure coordination of activities between both groups, and examine if ODaF should play a role in this initiative.

This lead to fruitful discussions and one of the ideas would be to take advantage of the DDI workshop that will be held this summer for the launch of the DDI user group. This could be sponsored by various organizations such as ICPSR, DDI Alliance, NORC, NSF, and ODaF, who could provide the facilities and resources for organizing the meetings. Some potential presenters to be invited to the event would include representative(s) from the DDI Alliance, member(s) from the DDI technical Implementation Committee, the Canada RDC Network, and NORC.

From the ODaF perspective, no action is required at this time. We will remain in contact with Bill and provide our support as needed to ensure the success of the initiative.

Open Source for DDI/SDMX

Bill Block and Jeremy Williams from Cornell both were interested to know which open source tools or initiatives are available for the development of DDI or SDMX tools. This was simply addressed by pointing out to resources such as the ODaF document on "Managing social, behavioral and economic data and metadata: Guidelines for Tools Development and Recommendations for Operating Environment"²⁶, the DDI Foundation Tools Program website²⁷ and road map²⁸, and the SDMX Tools website²⁹ and the SDMX Tools User Forum³⁰.

²⁶ http://www.opendatafoundation.org/ODaF_tools_developer_guidelines.pdf

²⁷ <http://tools.ddialliance.org/?lvl1=ftp&lvl2=overview>

²⁸ <http://tools.ddialliance.org/?lvl1=ftp&lvl2=roadmap>

²⁹ http://www.sdmx.org/index.php?page_id=13

³⁰ <http://www.metadatatechnology.com/userforum/>

Other projects

Collaboration with Carleton College

We have been discussing with Paula Lackie at Carleton College³¹ the possibility for undergraduate IT / Statistics student to contribute to the development of DDI-based tools for data and metadata management. The idea would be to draft requirements for simple tools that could be implemented as a class or semester projects. ODaF would also provide the basic technical support and expertise to support the activity under the supervision of Carleton teachers. The outputs will be donated to ODaF and made available through a free public website under an open source license. This would benefit the students by familiarizing them with statistical data and metadata standards and technologies. It would also benefit the DDI community through the availability of new free tools.

Standards Mapping Project

There has been much recent interest recently in the use of SDMX, DDI, ISO/IEC 11179, XBRL, and some other standards in combination within statistical agencies and some other organizations, to facilitate end-to-end metadata management. The Australian Bureau of Statistics, Statistics Sweden, and some other organizations have been very interested in this work. It is anticipated that a project will be organized to produce not only recommended mappings but also possibly tools focused on the combined implementation early in 2009.

Personnel changes

Directors

In January 2008, Julia Lane joined the National Science Foundation as Program Director Science of Science and Innovation Policy (SciSIP)³². Julia's new position at NSF makes it challenging for her to remain officially involved with ODaF and she will unfortunately have to step down from our Board of Directors. We are immensely grateful to Julia for her support and contribution to the Foundation in the past two years. We would not be where we are today without her assistance.

Advisors

No change was made to the board advisors in 2008. We agreed to discuss with the current Board members their interest with the Foundation as well as expand the size and scope and size of our expert group in 2009.

Management Team

We have reviewed the level of involvement of individual members of the management team and realized that most of the day-to-day operations and activities of the Foundation were conducted by Arofan and Pascal. The other members' professional activities have left them with very little time to contribute to the organization.

³¹ <http://www.carleton.edu>

³² http://www.nsf.gov/staff/staff_bio.jsp?lan=jlane&org=NSF

To be effective and grow the Foundation, we feel that members of the team should at least be able to dedicate a few hours every week to management. We therefore proposed to our Directors last September to open two of the positions to new individuals. This was approved with the agreement that managers stepping down will have the option to join the Board of Advisors.

During the DDI expert meeting in Dagstuhl in November, we approached Rob Grim and invited him to join the ODaF management team. Rob is currently Research Data Specialist & Coordinator e-Science Library and IT Services at Tilburg University in the Netherlands. He represents a very valuable asset for ODaF as he is not only well connected in Europe in domains that we do not currently have access to but also has a broad knowledge of data and metadata issues from the producer, archive and researcher perspectives. We were pleased to hear that Rob accepted our invitation. He quickly cleared his affiliation with his management and we received in December the unanimous approval from our Board of Director. A short biography of Rob is provided below.

This leaves us with one open manager position and we will continue to seek a suitable candidate. As we feel we cover quite well the administrative, technical and metadata expertise aspects, we will be focusing on an individual with good organizational, marketing and fund raising skills that can help us sustain and grow the organization. In the interim, Jostein Ryssevik will remain on the management team.

Biography: Rob Grim

From 1999-2007 I was employed as a data analyst at the Institute for Labour Studies (OSA - <http://www.tilburguniversity.nl/osa/>). I was very much involved in panel data management, data analysis, survey methodology and the production of comparative labour market statistics.

Since May 2007 I am employed as a Research Data Specialist at the Library and IT Services department at Tilburg University (TU - <http://www.tilburguniversity.nl/>); from September 2008 I have been appointed as a e-Science coordinator. As part of the Research Data Specialist position we are aiming to increase and extend the support for the (re)use of research data among research groups, PhD and other students. As part of exploring the role of a e-Science coordinator in the academic library of TU, we are exploring ways to innovate library services specifically relating to research data and semantically rich materials in general.

I have been involved in several projects where statistical -, library – and other automation standards play a significant role (DDI, DIDL/METS, MODS, ISO/IEC 11179, ORE, SDMX, XBRL). I am currently involved in the European project NEEO4 (<http://www.nereus4economics.info/neo.html>) that aims at creating open access content – both publications and data sets – within the economics domain. Furthermore, as part of my position as a Research Data Specialist I am responsible for our Financial Data Support Service located at the library. I am also a consultant to SURF (<http://www.surf.nl/en/Pages/home.aspx>) and the Knowledge Exchange Group (<http://www.knowledge-exchange.info>) on primary research data stake holder demands. I am interested in new scholarly publishing and business models and leveraging open access content.

2009 Activities and issues

Management issues

Funding

We currently do not have sufficient funding secured to sustain the organization in 2009. While we are confident that a sufficient number of new project or activities will materialize in the near future, this is a concern and we will closely monitor the issue.

Growth

It was pointed out during our meeting that the organization does not currently have a formal document that outlines its strategy for the coming years. We therefore agreed that we should draft a 3-5 year strategic plan that describes where the organization hopes to be a few years down the road along with approaches on how to get there.

Contracting ODaF Projects

In the United States, it is commonly accepted - both informally and legally - for a non-profit organization to establish and maintain relationships with for-profit companies. This is also often the model used to support open source initiatives and products or for standard setting organizations (such as the W3C, OASIS, and others), and is governed legally under laws based on the principle of “congruence of interest”. Among other countries, this principle is not always recognized, and thus issues of conflict of interest can arise when a not-for-profit and for-profit organizations collaborate.

As ODaF is only aware of a small number of agencies, companies or consultants with the relevant expertise in the social science data/metadata domain and with the capacity to deliver on our projects, we have not seen a need to establish a formal process to award project work up to this point. This has allowed us to rapidly respond to demand from the community or agencies. We have also always been very careful to clearly differentiate between unpaid contribution to the Foundation by managers, advisors or directors and compensated project work. Indeed, this clarity is required under the US laws governing not-for-profit organizations.

To ensure full transparency of the organization however, we feel it might be necessary to have a more formal process in place for awarding approved project work and we are considering taking the following steps:

- Establish and maintain a list of vendors (agencies or individuals) with relevant IT and data expertise interested in being involved in product development or providing services in domains related to ODaF activities.
- This list will serve several purposes:
 - It will establish a domain-specific database of vendors with specialized skills and expertise
 - It can be used by the community to advertise business opportunities to the vendor community and
 - It will be used by ODaF to potentially recruit resources to support its projects
 - It will foster communication and partnerships between vendors

Open Data Foundation – 2008 Annual Meeting Report

- A light vendor registration process will be put in place on the web site. Applicants will need to provide background information, relevant experience/ expertise related to statistical data, standards and technologies, typical consulting fees, etc. This information will be used to determine eligibility and made available as a public profile.
- When a project or activity of interest is identified, vendors on the list will be notified by email which will give them the opportunity to respond directly

If this work is for an ODaF project:

- Vendors will have the option to express their interest in participating in the ODaF proposal. This would not preclude the vendor from directly bidding on the project as well (if applicable).
- One Important condition for this to be efficient is a rapid response from vendors. RFPs often have tight deadlines and any delay may impede the process. When a call for expression of interest (EOI) is issued on the mailing list for an ODaF project, vendors are expected to respond within 2 business days. A longer grace period may be granted based on the specific project proposal process
- An award review committee will be established to, if necessary, decide upon the distribution of work amongst interested vendors.
- For project under \$100K, vendors are expected to self-coordinate and agree on the distribution of work. The committee will only be called upon if a conflict arises. The initial process will be moderated by the ODaF management team
- For project over or equal to \$100K, the distribution of work will always require review and approval from the award review committee.
- An important rule is that any member of the committee who is also directly involved as a vendor for the project list will not be allowed to participate in the review and award decision process.
- The committee is expected to respond in a timely fashion. In case of a tie or if the board takes more that 5 business days to reach a decision, the management team will have the discretion to award the work.
- Any complain about the committee processes or decision can be brought up to the Board of Directors.

This process will be formalized in the near future based on member, advisors and directors feedback and likely be introduce as a pilot for the first 12 months.

Note that some of the issues and questions we currently facing in this process are:

- Is this a good idea?
- Is \$100K a reasonable threshold for review?
- Can we identify a sufficient number of individuals for the committee?
- Are there other approaches to this issue that we should consider?

501.3c filing

The process of filling for 501c.3 tax exemption status for ODaF is in its closing stage and the accountants assisting ODaF with the filing tell us that it should be retroactive for 2008. We will follow-up rapidly on this as it is very important for how we engage with other organizations on projects.

Liability Insurance

Liability insurance remains an issue as we have struggled to find an affordable plan that would provide the organization world wide. We have agreed in the meantime to at least get insurance for the United States (or North America) and are currently filling out the relevant forms. An update will be provided shortly.

Currently, all liability for work performed is passed on to those supplying services to ODaF, but this model is not optimal for all anticipated activities.

2009 Resolutions

1. As of January 1st 2009, Julia Lane is no longer member of the Board of Director [this does not require a vote]
2. As of January 1st 2009, Rob Grim is Executive Manager and Chris Nelson / Bryan Fitzpatrick are no longer members of the Management Team. Jostein Ryssevick remains member of the management until a suitable replacement is identified at which time he will be converted into an Advisor. [this has already been approved] It has been recommended that Bryan Fitzpatrick be converted into an Advisor as well. [This has not yet been approved.]
3. Actively seek a new Director for ODaF
4. Actively seek a new member of the Management Team whose contributions will focus on marketing, fund raising, and growth
5. Contact members of the Board of Advisors to determine their role and interest in the organization
6. Actively seek to expand the range of expertise of the Board of Advisors
7. Make a proposal to our Directors for the establishment of a formal process to award project work to commercial vendors
8. Organize the ODaF Europe 2009 regional meeting in April 2009 at IZA in Bonn, Germany
9. Organize the ODaF North America 2009 regional meeting in the Fall of 2009 (dates and locations to be determined, could consider Canada)
10. Sponsor and participate in the IASSIST conference in Tempe, Finland [approval not required under standing resolution regarding DDI support]
11. Renew our membership in the DDI Alliance
12. Publish at least 5 papers
13. Draft a 3-5 year strategic plan for the organization
14. Host an SDMX registry for the sponsors of the standard

15. Produce startup guides for SDMX producers and users with a focus on adoption in the United States
16. Support the establishment of SDMX Users Forums
17. Engage with the University of Chicago Computational Institute in research activities for the use of Web 2.0 technologies in social science for the improvement of users metadata
18. Support the establishment of a pilot metadata registry that support ISO 11179 3rd edition (refactoring version)
19. Support the integration of the DDI and SDMX specifications in Space-Time Research product line
20. Support the establishment of a European and North America DDI Users' Groups
21. Collaborate with Carleton College toward the implementation of basic DDI tools by undergraduate students
22. Support mapping efforts between SDMX, DDI, ISO/IEC 11179, XBRL, and other standards

Annex A: ODaF Overview



Our Vision

To measure the well being of their population, the state of their economy, technological progress and many other key indicators, nations rely heavily on socio-economic statistics. These data are collected directly from individuals, households, or enterprises using surveys and through the administrative system. How the actual information is compiled by statistical agencies, preserved and made accessible by data archives and analyzed by researchers or economists is a long and complex process. However, the resulting outputs and indicators are crucial for decision makers to formulate national policies or to monitor progress on the ground. At the international level, they are needed to measure the state of the global economy or steer aid efforts toward developing countries.

Producing socio-economic official statistics is the mandate of the national statistical system whose structure varies greatly from country to country (national statistical office, line ministries, or complex federated systems). Data production activities are, for many countries, a part of day to day government operation. In others, particularly in developing countries, political instabilities, conflicts, or lack of capacity have made this a challenging process that often leads to limited availability of data. Nonetheless, data are nearly always collected one way or the other.

The mere existence of data is far from being sufficient to guarantee that the needs of decision makers are met, however. Furthermore, the overall quality of data is an issue – the data must be sufficient for researchers to turn into information relevant to its intended end users. Data quality is a multi-faceted concept that has been extensively documented¹ and it is commonly agreed that *usefulness* is a good generic indicator. To be useful, data must first be *accessible*: the best data in the world has no purpose if it cannot be reached by users. It must also be usable, which means that *comprehensive documentation* describing the data collection and its production processes must be available to its users.

These issues of accessibility and documentation rarely receive the attention they deserve. Despite the availability of new technologies and standards, data are often inadequately preserved and documented. Discovery and access also remain common obstacles. This situation does not arise from lack of efforts; producers and data archives have expressed their commitment to improving access and documentation. Rather, the challenges seem more to come from the complexity of the data production chain that spans across several domains of expertise, a lack of familiarity with the relevant IT technology, and the need for appropriate tools, standards, and guidelines.

The Open Data Foundation (ODaF) has been established as a platform to address these issues by providing an umbrella for experts from various domains to come together, promoting the adoption of standards and best practices, and coordinating the development of tools. Through its activities, ODaF aims to *foster the global adoption of metadata specifications* that will not only facilitate accessibility to and understanding of socio-economic data, but also support the harmonization and exchange of information

between agencies and with the general public. ODaF also intends to enhance the IT capacity and understanding of relevant technologies by agencies involved in data production, preservation, and analysis.

Rather than competing with existing efforts, ODaF strategy is to complement ongoing initiatives, working to provide support and fill in the gaps where they appear. Activities such as producing, preserving or disseminating data, providing access to statistics, defining metadata standards, or funding projects are already the mandates of statistical agencies, data archives, research centers, foundations or international organizations.

ODaF's objectives are to:

- foster the understanding of *data and metadata management standards and technologies*;
- provide *technological expertise* to agencies in their respective area of activity;
- fill the information technology gaps by supporting the *development of open source tools*
- play a central role in *networking* together individuals and organizations that often work in isolation rather than try to solve common challenges together; and
- promote the production of *open data*.

By *open data* we mean well-documented statistical information that can be easily discovered, accessed for the right purpose, and used in accordance with statistical principles and legislation. *Open data* is a term which is used in many ways, so it is important to be clear about how we define it. It is naïve to think that the simple existence of data in the public domain will make it practically available to researchers for use. There are many impediments to the use of data, and the legal right to use the data is only one of them.

This is not a simple issue – data confidentiality is critical in an age when the Internet and other technologies raise significant concerns about privacy. There are guidelines and legislation aimed at protecting respondents, and we feel that these are necessary protections for respondents.

ODaF therefore focuses as much on secure data centers or virtual enclaves as on public data and metadata, so that data which has historically been kept out of researchers' hands can be provided to those who have the right to access it. We accept that such data is not completely open; our goal is to make it as open as possible, while being responsible in terms of the protection of respondent identity.

Through its efforts, the Open Data Foundation hopes to foster a global harmonized framework that will facilitate the flow of statistical information and promote the dialog between all stakeholders. This will in turn help bring high-quality socio-economic data to the researcher and timely, relevant information to the decision maker to allow for evidence-based policy and monitoring. The adoption of common metadata standards at the global level will also provide for the exchange of information across borders. In a world of economical crisis and where the population in less-developed countries will outnumber that of the developed world by a 6 to 1 ratio by 2050ⁱⁱ, sound statistical data are crucial to assess and monitor the state of our societies.

Open Data Foundation – 2008 Annual Meeting Report

The Open Data Foundation is a US based non-profit organization that was established in 2006. Membership is open and free to interested individuals. Its activities are coordinated by the management team under the supervision of its board or directors. Specialized expertise is provided through its board of advisors composed of individuals from a wide range of agencies. For further information, visit <http://www.opendatafoundation.org> or contact info@opendatafoundation.org

ⁱ <http://www.statcan.ca/english/freepub/12-539-XIE/index.htm>,

<http://stats.oecd.org/glossary/detail.asp?ID=541>, ,

<http://dsbb.imf.org/Applications/web/getpage/?pagename=dqrshomea>

ⁱⁱ 2004 Revision of the official United Nations population estimates and projection