Democratizing Data
The IASSIST Strategic Plan for 2010-2014

by

Members of the IASSIST Strategic Plan Action Group

Joel Herndon (Chair)

Michelle Edwards

Paula Lackie

Rob O'Reilly

Democratizing Data
The IASSIST Strategic Plan for 2010-2014

Executive Summary

In 2008, the International Association of Social Science Information Services and Technology initiated a new strategic planning process to evaluate IASSIST's current operating environment, gather feedback from its membership, and propose new strategies for the future of the organization. Under the leadership of the IASSIST Administrative Committee and with the feedback of the membership, the following strategic plan provides a series of strategies that will help IASSIST address new challenges while deepening the organization's commitment to supporting research data, data infrastructure, and the communities surrounding these data.

Strategic Directions for 2010-2014

To fulfill our vision, IASSIST will focus on the following strategic directions:

Strategic Direction 1: ADVOCACY - Advocate for policy and practices that increase responsible access to data and promote responsible data curation.

Strategic Direction 2: COMMUNITY - Cultivate a community that supports the full data life-cycle; creation, use, and preservation of data.

Strategic Direction 3: PROFESSIONAL DEVELOPMENT - Provide and develop resources that foster the career development of data professionals.

I. Democratizing Data: The Planning Process

This document represents the culmination of a two year effort to update IASSIST's 2004-2009 strategic plan. Initially launched at the IASSIST 2008 annual conference, "Democratizing Data" builds on the themes of the 2004-2009 IASSIST strategic plan, addresses changes in IASSIST's operating environment, and extends the role of IASSIST to better serve its members needs. With a focus on the three key themes of **Advocacy, Community, and Professional Development**, this plan examines the current state of the organization, its operating environment, and suggests strategies for moving forward in alignment with members' stated goals.

In 2008, the Administrative Committee charged the Strategic Planning Action Group (SPAG) to examine the current environment in which IASSIST operates, to survey the membership about their thoughts on the organization, and to identify themes to guide the organization through the next five years.

In the fall of 2008, members of the SPAG surveyed the IASSIST membership on their impressions of the previous strategic plan, their comments on the current state of the organization, and their aspirations for the organization's future. The IASSIST 2009 conference provided several fora for sharing results of the survey including a special presentation to the Administrative Committee, an open discussion of the results with the membership, and several meetings for members interested in joining the planning process. As a result of these discussions, the Administrative Committee charged the Strategic Planning Action Group to draft a plan for the 2010 conference.

SPAG is grateful to both the IASSIST membership and the Administrative Committee for their thoughtful comments over the last two years. This document is designed to reflect the comments of the membership, the goals of the larger organization, and provide a base of ideas for IASSIST as it grows over the next five years.

II. The Planning Environment

Organizational Context

Membership

In 2010, IASSIST is an international organization of roughly three hundred professionals sharing the common goals of improving data services, advancing infrastructure in the social sciences, and fostering a community of social science data professionals. Members work in diverse institutional settings and geographic locations. IASSIST has members working in universities, research libraries, data archives, national statistical agencies, research groups, private companies, international organizations, and non-profit organizations among other settings. Geographically, the current membership is global in scope, although the majority of the membership remains based in the United States, northern Europe, and Canada.

While IASSIST members represent an extremely diverse set of institutional environments, the organization serves three primary communities sharing concerns with social science data. Social science researchers and scientists who produce and use social science data comprise the first community. Information specialists who preserve, manage, advise users in the selection of, and distribute social science data form the second community. Finally, computing specialists who develop the infrastructure and methods to share, manipulate, and analyze social data form the third community. IASSIST programming, training, and conferences are largely designed to appeal to members of these communities. The communities define and provide structure for both the organization's structure and planning.

Organizational Goals and Structure

The IASSIST Constitution defines the organization's objectives as:

- To encourage and support the establishment of local and national information centers for social science machine-readable data.

- To foster international exchange and dissemination of information regarding substantive and technical developments related to social science machine-readable data.

- To coordinate international programs, projects, and general efforts that provide a forum for discussion of issues relating to social science machine-readable data.

- To promote the development of standards for social science machine-readable data.

- To encourage educational experiences for personnel engaged in work related to these objectives.

In order to pursue these goals, IASSIST relies on a combination of formal structures and informal activities. The organization's officers and committee structure provide a formal structure for implementing the business of the organization such as planning the annual conference, workshops/training sessions, communications/web presence, finances, and cooperation with other organizations. Informally, IASSIST members represent the organization in a variety of settings ranging from local training sessions on data issues to international programs to build statistical capacity around the world. IASSIST's high level of organizational capacity remains notable for an all-volunteer organization.

Membership Input on the 2010-2014 Plan

The process for the 2010-2014 plan employed a number of different methods to solicit input from the IASSIST membership on their concerns, goals, and aspirations for the organization. Members provided a wide range of informal feedback ranging from

conversations at IASSIST conferences, open sessions devoted to sharing opinions about the process, emails, and phone calls. More formally, the SPAG employed two formal surveys designed to elicit ideas and comments from the membership, including those who were unable to attend the conferences. The first survey targeted SPAG members and members of the Administrative Committee and asked them to list up to ten themes that they believed were the primary issues for the 2010-2014 Strategic Plan. The second survey[i] (https://survey.cla.umn.edu/38717/) targeted the general membership and focused on a series of questions about the previous Strategic Plan and the IASSIST current landscape.

From the two surveys, the following themes emerged:

- **Scope of Data Services:** Members note that the scope of data and the interests in this area have grown exponentially since the last strategic plan. Members called for the organization to reevaluate its traditional focus on social science data and to consider partnerships and training opportunities with other organizations. As one member noted: "Data are larger than the social sciences."

- **Improving Communication:** Members expressed satisfaction with the existing IASSIST email list, but offered a number of suggestions on modifying other forms of communication both within the organization and beyond the organization. Some members suggested the need for a new suite of online tools to facilitate communication in the organization and to appeal to new members.

- **Training:** Members expressed support for extending IASSIST's current training programs to reach a larger audience through a variety of strategies including building communities of best practice globally, digital curation initiatives, and regional training programs led by individual members.

- **Advocacy**: Many members noted IASSIST's success at having individual members engaged in advocacy. Members expressed support for building resources within IASSIST designed to encourage advocacy and outreach of individual members on behalf of IASSIST.

External Environment

IASSIST has evolved to operate in an environment substantially different from the environment described in the previous plan. While it was clear during the last planning process that the importance of data was growing and crossing disciplinary boundaries, few could have predicted the rapid expansion of concerns surrounding data produced by a confluence of new open data initiatives, the rise of new repositories for research data, a reinvigorated set of data citation initiatives, and recent data management requirements by major research funding agencies around the world.

IASSIST currently finds itself in a situation where it is one of many groups addressing the challenges of supporting the full range of the data life-cycle. At present, IASSIST also faces an environment where the very term "data" remains contested by different

groups seeking to influence data curation policy and practice. With the very nature of "data" under discussion, IASSIST must reevaluate its core mission of primarily supporting social science data concerns and assert its extensive collective expertise in data curation and data management.

Other prevalent trends in the external environment include:

- The rise of data curation concerns in the natural sciences (cyberinfrastructure, eScience, and/or eResearch)

- An expansion in the number of government and intergovernmental open data initiatives

- A growing demand for data curation training in a variety of disciplines (but especially the information sciences)

- An expansion of both institutional repositories and domain specific repositories accepting research data

- An increasing number of complex research questions requiring access to primary and secondary data sources

- Widespread usage of GIS tools by the general public and an expectation for researchers to employ spatial data in their research

- An increasing realization of the importance of data citation and several new initiatives to address these concerns

- Increasing number of research concerns that utilize multiple types (quantitative, qualitative, GIS) of data

- Increased use of social media sources to communicate about issues of data resources, tools, and practices

III. Strategic Directions over the Next Five Years

Strategic Direction 1: Advocacy - *Advocate for policy and practices that increase responsible access to data and promote responsible data curation*

Over the last few years, both governmental organizations and academics are increasingly of the view that research results and government funded data should be more accessible to the public. In the United States, the National Institutes for Health, and the National Science Foundation have required or will soon require grant applicants to include data sharing plans as part of their grant applications.[ii] Canadians are also

talking about the imperative to make publicly funded research data more readily discoverable.[iii] In the United Kingdom, meanwhile, the Labour government of Gordon Brown took various steps to increase the availability of government-gathered data, such as opening up the UK Ordnance Survey to the public and mandating the creation of a modern "Domesday Book" that will list all "non-personal" datasets created by government agencies.[iv] Finally, at a multi-national level, the OECD Committee for Scientific and Technical Policy issued a communiqué in 2004 encouraging increased public access to research data.[v]

Academics have long argued that scholars should make their research data more accessible and that academic journals should have policies that require authors to provide data associated with published articles.[vi] Proponents of data sharing hold that for research to be truly "scientific," the results from that research must be *replicable* – other researchers should be able to reproduce those results and to examine how the data were assembled and what assumptions went into the analysis.[vii]

Despite the initiatives and arguments for open data, there is a tension between the norms and expectations of how research *should* be done versus the day-to-day realities of how research often *is* done. The latter realities often work against data sharing. Adding to the challenges facing open data are a variety of regulatory, ethical, and proprietary reasons why access to certain data, such as data pertaining to sensitive topics of health or inter-personal relationships need to be restricted or tightly-controlled.[viii]

Rules and regulations such as the U.S. Confidential Information Protection and Statistical Efficiency Act, the U.S. Health Insurance Portability and Accountability Act, the EU Data Protection Directive, and the British Data Protection Act all have implications for how research data can be made public.[ix] Microdata that include geographic coordinates or other geographic identifiers introduce additional complexity. The potential utility of such data for researchers is considerable, but the potential for disclosure risk is also considerable, so safeguards must be in place before such data can be made public.[x]

IASSIST can play an advocacy role in bridging the gaps between expectations and practices, as its members are in a unique position to reconcile these differences in a way that can increase the availability of locally-created research data. Furthermore, IASSIST's long experience with metadata standards in the social sciences offers a proven set of tools for increasing responsible access to data.

Strategies to achieve this direction include:

1. **Facilitate linkages among IASSIST members in various institutional settings to enhance responsible access to data**

IASSIST members include people who set the norms and expectations for data management and planning; people who implement these norms to ensure the long-term availability of data; and people who work with the researchers who are meant to comply with the norms and expectations and will need assistance in how to actually do so. All of which makes IASSIST well positioned to advocate in favor of facilitating responsible

data access. Following are some examples activities the organization could pursue to improve data access and analysis while also protecting confidentiality:

    A. Develop guidelines, practices, and recommendations that researchers can reference during the research process so that their data can be shared in a manner that is useful to others
- Create a database or compilation of policies and requirements for funding organizations with regard to data sharing
- Create an annotated database or compilation of relevant laws and regulations that pertain to the protection of personal information
- Provide a venue for researchers to connect with IASSIST members for consultation on their data management needs

    B. Advocate as an organization for the research and social benefits provided by open data (see Community section of plan for possible venues)
- Provide ready-made materials for members to use in data advocacy presentations elsewhere; "IASSIST in a box"

    C. Build a repository of case studies for researchers seeking examples of data management strategies

2. Promote metadata standards that improve access to qualitative and quantitative data

The commitment of IASSIST to metadata standards for social science data continues to remain a central theme throughout the organization's history. In the 2004-2009 strategic plan, the metadata advocacy effort focused on the Data Documentation Initiative (DDI). In 2010, IASSIST members continue to promote the implementation of DDI in an environment in which the number of metadata standards for data collections has expanded considerably. The following are some examples of how IASSIST members could promote metadata standards that improve access to research data:

    A. Continue to collaborate with the DDI Alliance to advocate for an expansion in the usage of DDI.
- Partner with the DDI Alliance to hold local and regional DDI training sessions

    B. Advocate for the usage of DDI in research organizations and libraries
- Work with the Alliance to create a wider set of tools for creating, managing, and converting DDI

    C. Collaborate with other organizations supporting social science data to build crosswalks across metadata standards used for data collections
- Partner with schools of information science to include DDI in the metadata curriculum

    D. Encourage organizations working on open data initiatives to utilize metadata standards
- Partner with other organizations which advocate for Open Data and Open Access standards[xi]

Strategic Direction 2: Community *- Cultivate a community that supports the full data life-cycle; creation, use, and preservation of data*

The landscape of the data community has changed considerably from our last strategic plan. The breadth and depth of this data-focused community has expanded dramatically. Those focused on social science data are now being joined by recognized digital humanities, eScience, and open government initiatives. IASSIST should strive to build relationships among these emerging organizations to encourage the development of a shared understanding of standards for all phases of the data life-cycle. IASSIST should also pursue alliances with organizations that might be mutually beneficial.

Strategies to achieve this direction include:

1. Enhance the IASSIST online community

Building a stronger online presence can improve internal communication while expressing our commitment to the broader data community. Following are some examples of specific activities the organization may consider:

    A. Enhance the IASSIST website with new technologies that encourage both internal and external participation
- Add resources of interest to the broader data community such as guidelines and tools for managing data collections, discussion boards, resources for data curation, and a speakers' bureau
- Help current and potential members recognize their increasingly important role as data professionals and data advocates in many aspects of their careers as well as clarify the role of IASSIST in supporting their career progress.

    B. Strive to market our publication – the IASSIST Quarterly (IQ) to a broader community than IASSIST
- Include IASSIST Quarterly in advocacy efforts surrounding data curation
- Lobby data curation instructors to consider using resources from the IQ in courses on data curation at schools of information

    C. Build a strong external presence by taking advantage of social networking technologies
- Encourage IASSIST members to include IASSIST social networking strategies into their local data strategies

2. Build data communities of awareness and practice with other organizations with a shared data-related mission.

As mentioned in *Strategic Direction 1: Advocacy*, IASSIST exists in an environment of diverse organizations committed to issues surrounding responsible data access, data management, metadata, and preservation. In the last five years, IASSIST has made several notable initiatives to improve connections with related data organizations (for

example, with **Association of College and Research Libraries** and Committee On Data for Science and Technology, CODATA). Additionally, IASSIST members currently serve on a number of organizations with significant data initiatives in a variety of government, non-government, and research environments around the world.

Despite this participation, alliances with IASSIST as an organization remain elusive. Following are some examples of specific activities the organization might want to pursue:

A. Build strategic alliances with associations, groups, and organizations interested in data and data related issues through IASSIST interest groups or member/liaison participation. Groups and organizations might include:

A variety of library-related organizations such as:
- Library and Information Association of South Africa (LIASA)
- Society of Archivists (UK)
- The Society of Finnish Archivists/Arkivföreningen ry
- American Library Association (ALA)
- Association of Research Libraries (ARL)
- Canadian Association of Research Libraries (CARL)
- German Society for Information Science & Praxis (DGI)

A variety of data user-groups:
- Association of Public Data Users (APDU)
- Canadian Association of Public Data Users (CAPDU)
- Center for Research and Teaching Economics, Mexico (CIDE)
- A Danish formal central repository of information on data interchange standards (http://digitaliser.dk/)
- The Open Knowledge Foundation & Open Government Data Organization

Data archives and meta-organizations around the world:
- Council of European Social Science Data Archives CESSDA
- International Federation of Data Organizations for the Social Science (IFDO)
- International Initiative to Facilitate Access to Research Data (DataCite)
- Inter-university Consortium for Political and Social Research (ICPSR)
- National Statistics Offices (NSO) eg: Central Statistical Agency of Ethiopia (CSA)

Other communities/organizations of/for data professionals:
- The Committee on Data for Science and Technology (CODATA)
- Digital Humanities groups
- Educause Center for Applied Resarch (eCAR)
- eScience / eResearch communities
- International Household Survey Network (IHSN) & Accelerated Data Program (ADP)
- Open Geospatial Consortium (OGC)

B. Collaborate with Information Technology communities at a local, national and international levels to promote and test new technologies for data such as cloud

computing and grid computing. Share these new technology processes with the larger IASSIST membership to expand awareness of new data paradigms.

3. **Expand the IASSIST Community to include regions that are not well represented in the IASSIST membership**

As an organization, IASSIST remains committed to working with data organizations and communities around the world engaged in building statistical capacity, providing responsible data access, data curation, and preservation. As an organization, IASSIST's International Outreach Committee has provided funds and logistical support for individuals from emerging economies seeking to attend the IASSIST annual meeting. In all of these efforts, IASSIST faces the logistical challenge of building communities of practices surrounding data with limited resources.

In the next five years, IASSIST should continue to pursue strategies that broaden its presence in regions that are not currently well represented while remaining pragmatic about the resources available for these efforts. Following are some examples of specific activities the organization might want to pursue:

A. Develop web based resources for the IASSIST community with the intention of using these resources to reach individual data-professionals and communities that may not be able to attend the annual conference

B. Refocus IASSIST's community building efforts to include more members from underrepresented areas; support the formation of regional interest groups to sustain and encourage:
   • the dissemination of IASSIST materials in different languages
   • the regional affiliations among data professionals
   • the greater development of the awareness of IASSIST in areas where it has not yet taken hold; the Spanish and Portuguese speaking world, Southeast Asia, Asia and further support for the budding IASSIST community in Africa and Eastern Europe

C. Deepen partnerships with other international data organizations to leverage the strength and resources of each organization.

Strategic Direction 3: Professional Development - Provide and develop resources that foster the career development of data professionals

2004-2009:
The strategic direction for professional development is one that carries forward from the 2004-2009 strategic plan. The following was documented in the last strategic plan:

As an organization, IASSIST has a history of working to educate its members about matters of common professional interest to the social science data community. Traditionally, this education has taken the form of professional development opportunities available in member-initiated and member-taught workshops at the annual IASSIST conference. Yet as the world of social science data grows increasingly complex, staying

abreast of new developments in the profession is likely to present an ever-increasing challenge for IASSIST members. As a result, the educational component of IASSIST should expand to encompass far more than the traditionally inward looking workshop training on new standards, techniques, or technologies. Indeed, a far-reaching and outward-looking educational component to IASSIST can go far in achieving many of the goals outlined in the other areas of this strategic plan. In formulating an expanded educational component to IASSIST, we must be careful to concentrate our efforts on activities that will provide the greatest benefit.

2010-2014:

Education is an aspect of our community that is constantly growing and shifting. In reviewing the 2004-2009 strategic plan many of the directions set out in the previous plan still hold true. For 2010-2014 the education strategy will carry forward from the last plan with new goals.

Strategies to achieve this direction include:

1. **Continue and improve professional development and educational opportunities for IASSIST members (training for membership)**

Professional development can be viewed in many different ways. IASSIST should strive to provide professional development opportunities in many formats and through different venues. Following are some examples of specific activities the organization might want to pursue:

   A. Developing new skills
      - Keep abreast of new data technologies, standards, and tools
      - Keep abreast of new educational methodologies that translate well to data issues
      - Share best practices for metadata, outreach, and instructional methods

   B. Continue to enhance the IASSIST conference by:
      - Establish through the Education Committee a network of trainers based on a model of peer-to-peer instruction;
      - Conduct train-the-trainer workshops in conjunction with the annual conference; and Include half-day discussion sessions at the annual conference
      - Create training opportunities for IASSIST members beyond the IASSIST conference by initiating or partnering with other data related organizations on regional workshops on matters of importance

   C. Create training opportunities for IASSIST members beyond the IASSIST conference:
      - Initiate through the Education Committee and Regional Secretaries special training activities outside the annual conference, such as regional workshops on matters of importance to IASSIST members

2. **Train the next generation of data professionals**

Formal education for data librarians and data professionals is ad hoc and not widely available in member countries. As an association of data professionals, many of whom

have been trained within the association, the organization should strive to work with higher education institutions to formalize data training. Following are some examples of specific activities the organization might want to pursue:

A. Foster relationships with higher education institutions offering data programs and/or courses related to the data professions.
- Collaborate with higher education institutions to provide formal data training opportunities
- Promote data curation as a profession

B. Enable a mentorship program
- Create through the Education and Membership Committees a network of mentors and a method of pairing new colleagues with more experienced mentors
- Develop through the Education and Outreach Committees a coaching/mentoring program for staff members of data archives and related organizations
- Establish through the Education and Outreach Committees a leadership program for the management of data archives and related organizations
- Facilitate through the Membership Committee the exchange of members between established institutions and those new to, or seeking guidance from the social science data community.

3. **Build a repository of training materials created and used by IASSIST members**

One of the greatest strengths of IASSIST is the commitment of its members to the data profession. Education and training for students, staff and/or faculty members at many of our home institutions is often only one aspect of our position. Sharing training resources amongst members will only augment and expand our rich resource, IASSIST members. Following are some examples of specific activities the organization might want to pursue:

A. Create a training repository
- Collect training resources used by members at home institutions to be shared and repurposed amongst members
- Create repository to provide easy access to collected resources
- Encourage data-related grant recipients to submit their successful grant applications to a repository for IASSIST members and provide a forum space for aspiring grant writers and those working to accomplish their grant-related goals.

Appendix A: Members of the Strategic Plan Action Group from 2008-2010

Bill Block
Ernie Boyko
San Cannon
Gordan Chancellor
Michelle Edwards
Jen Green
Joel Herndon (Chair)
Lynda Kellam
Helena Laaksonen
Paula Lackie
Thomas Lindsay
Kristen Partlo
Robin Rice
Bo Wandschneider
Wendy Watkins
Amy West
Melanie Wright

A special thanks to the many members who contributed to the planning process through conversations, emails, phone calls, and contributions! A special thank you to the members of the 2004-2009 Strategic Plan whose work strongly influenced this document.

Appendix B: Cited Resources

Anderson, Richard G., William H. Greene, B. D. McCullough, and H.D. Vinod. 2005. *The Role of Data Program Archives in the Future of Economic Research*. St Louis, MO: Federal Reserve Bank of St. Louis Research Division.

Freese, Jeremy. 2007. "Replication Standards for Quantitative Social Science: Why Not Sociology?" *Sociological Methods & Research* 36:153-172.

King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics* 28:443-99.

Gleditsch, Nils Petter, Claire Metelits. 2003. "The Replication Debate." *International Studies Perspectives* 4:72-79.

Gleditsch, Nils Petter, Claire Metelits, and Havard Strand. 2003. "Posting Your Data: Will You Be Scooped or Will You Be Famous?" *International Studies Perspectives* 4:89-97.

National Research Council. 2002. *Access to Research Data in the 21st Century: An Ongoing Dialogue Among Interested Parties Report of a Workshop*. Washington, DC: National Academy Press. Downloaded from http://www.nap.edu/catalog/10302.html on 17 May 2010.

National Research Council. 2007. *Putting People on the Map: Protecting Confidentiality with Linked Social-Spatial Data*. Washington, DC: National Academy Press.

Endnotes

[i] William Block, San Cannon, Thomas Lindsay planned and implemented the online membership survey.

[ii] For the NIH's data-sharing policies, see http://grants1.nih.gov/grants/policy/data_sharing/index.htm. For the NSF's recent statement on data sharing, see http://www.nsf.gov/news/news_summ.jsp?cntn_id=116928&org=NSF.

[iii] See, for example: "Improving Canada's Digital Advantage: Strategies for Sustainable Prosperity" (Consultation Paper on a Digital Economy Strategy for Canada) (May, 2010, http://de-en.gc.ca/consultation-paper/)

[iv] Brown's proposals generated much commentary. See the following for more information:
http://www2.labour.org.uk/gordon-browns-speech-on-building-britains-digital-future

http://www.ordnancesurvey.co.uk/oswebsite/media/news/2010/April/OpenData.html

http://www.guardian.co.uk/technology/blog/2010/mar/22/digital-britain-digital-economy-bill

http://www.techradar.com/news/digital-home/gordon-brown-announces-digital-domesday-book--678594

[v] See http://www.oecd.org/document/15/0,3343,en_2649_34487_25998799_1_1_1_1,00.html.

[vi] See, for instance, the data availability policy of the *American Economic Review*: http://www.aeaweb.org/aer/data.php.

[vii] Once again, there has been much commentary on this topic. See, for instance, King (1995) and Freese (2007).

[viii] McDermott (2010); Albright and Lyle (2010); and Gleditsch and Metelis (2003).

[ix] National Research Council (2007).

[x] National Research Council (2007).

[xi] Organizations such as the Open Data Foundation (http://www.opendatafoundation.org/), The Comprehensive Knowledge Archive Network (http://www.ckan.net/), The International Houshold Survey Network (http://www.ihsn.org), Citizens for Open Access to Civic Information and Data (http://civicaccess.ca/), Public.Resource.Org (note, for example, their "8 Principles for Open Data" http://resource.org/8_principles.html), and many more.