Building an SDI as a Community Project – Challenges in Emerging Economies

Josef STROBL[1]. Mariana BELGIU[2] and Ainura NAZARKULOVA[3]

[1]Centre for Geoinformatics, University of Salzburg – josef.strobl@sbg.ac.at
[2]Institute for Geographic Information Science, Austrian Academy of Sciences – mariana.belgiu@oeaw.ac.at
[3]ACA*GIScience at KSUCTA, Bishkek, Kyrgyzstan – ainura.nazarkulova@aca-giscience.org

Abstract

This paper is focused on the entire framework of establishing national and regional Spatial Data Infrastructures in regions undergoing fundamental administrative, economic and societal restructuring. Based on experience from a GSDI small grants project aimed at developing prototype demonstrator SDI components for the Kyrgyz Republic and Central Asia, recommendations and guidelines are proposed and complemented with observations of organizational and other success factors to be considered.

Lessons learned are based on the premise that in environments where state institutions like National Mapping Agencies or Cadastral Offices are not in a position to establish an inclusive and open SDI, there still is a (maybe even bigger) need for Spatial Data Infrastructures. GIS user organizations like NGOs, United Nations Organizations, international development and funding organizations, grassroots initiatives and academic institutions all are dependent on a broad range of geospatial data, and multiple acquisition / redundant data holdings are rampant.

Cooperating across institutional boundaries is the key to reaping the benefits of an integrated infrastructure, giving all participants the chance to focus more on their core business and less on base maps in a narrow as well as wider sense. The authors have established a demonstrator geoportal intended as the hub for fostering such cooperation, and as a communication and education resource to build a broad understanding of potential SDI benefits.

Keywords: *Spatial Data Infrastructures, SDI, Organization, Communication, Geoportal, Community based SDI*

1. INTRODUCTION

Particularly in 'emerging economies' the role and importance of Geographic Information as a core infrastructure for administration, development and businesses cannot be underestimated. At the same time, discontinuities in administration and disruptive developments in societies and economies are a substantial obstacle in building the general understanding of open and shared infrastructures. Obvious stakeholders and key actors frequently are not ready to step on a common and transparent platform, the protection of 'turfs', the perception of information as power and lack of integrity are critical obstacles.

While there are technical obstacles in the regional power, ICT and human resources infrastructures, organizational, qualification and societal factors are considered the main hurdles to overcome when aiming at the development of open SDIs reaching beyond purely re-labeling the activities of e.g. a national mapping or cadastral agency. Interestingly, international organizations and donor institutions, development projects and NGOs currently are the most active and open potential partners and contributors towards a truly collaborative regional SDI.

The above mentioned joint project chose the strategy of developing and implementing demonstrator SDI components, in order to effectively convey the overall idea of an SDI, to communicate its value for a broad range of activities and to offer a platform for teaching and learning. This demonstrator is centered around a catalog service implemented with a proposed regional metadata profile. Accessed via a geoportal, this discovery environment offers access to a range of local and regional data sets hosted on a dedicated services infrastructure, as well as links to other catalog services.

In the second phase of this initiative additional institutions or 'orphaned data sets' are planned to be integrated, and the main emphasis will shift towards educational activities in order to demonstrate characteristics and components of an SDI architecture, help with understanding potential benefits, and motivate to join this kind of shared, community-based activity where the core infrastructure elements are contributed by academic institutions aiming at supporting all actors ranging from government through NGOs to business.

The authors are cooperating under the umbrella of the Austria-Central Asia Centre for GIScience (ACA*GIScience), a joint academic initiative with a vision to contribute to the development of a regional SDI through technical infrastructures and education. ACA*GIScience in collaboration with the Kyrgyz GIS Association are currently serving as the focal point for implementing and promoting a regional SDI geoportal.

2. WHOSE STAKES?

In most countries and regions SDI initiatives (MASSER, 2005) are being established through national government-led initiatives or as transnational frameworks (MASSER, 2007). Across much of the developing world and in emerging / threshold economies this kind of government leadership frequently is missing. Clearly this fact is less rooted in economic or technical limitations, but rather in the perceived loss of control by traditional government or state institutions, and the impression of giving up potential sources of revenue.

Governmental institutions of course are not the only stakeholders in an SDI, and depending on a particular situation might even take a back seat when considering de facto activities and practical interests. Taking the case of Central Asian countries as an example, a survey of recent regional conference proceedings (ABDYKALYKOV et al 2009, 2010), project websites and newsletters and tenders demonstrates a broad range of institutions, initiatives/programs and projects having very substantial interests in, or being entirely based on geospatial data sets:

- United Nations actors
- International development and cooperation agencies (GTZ, Helvetas)
- Geospatial research institutions and programs (GFZ Potsdam with CAIAG, CDE Bern)
- Research at academic institutions in the region (University of Central Asia)
- Development funding institutions (ADB, IDB)
- Projects implemented through various programs (EU INTAS, US CRDF)

All these stakeholders do in principle have a common need for much of their geospatial base data sets as well as large overlaps in the demand for topical data, but current practice shows very little cooperation. Data sets are exchanged, sold and bartered, leading to a hodgepodge of usually undocumented versions. This is compounded by a pervasive and persistent lack of personal and institutional integrity, data sets are frequently considered sources of personal income, and thus any move towards a documented infrastructure would run counter the personal interests of many such ‘stakeholders’.

Only few attempts have been made to move beyond this jumble of data sets ‘somehow’ supporting one-off project tasks, the UN Spatial Data Infrastructure initiative of the UN Geographic Information Working Group (UNGIWG, 2010) being one promising exception from an otherwise bleak outlook. As UN SDI is primarily aiming at supporting UN organizations, the huge majority of the above listed actors will only be able to benefit to a limited extent.

Global datasets available through geospatial browser platforms like Google Map/Earth, bing or ArcGIS Explorer Online currently are establishing de facto base maps and reference imagery, making some legacy data sets obsolete. Their impact currently is still limited, though, due to the implied demands on a network infrastructure and the lack of an underlying open services architecture – with ArcGIS Explorer Online offering a partial solution to the latter issue.

Since there obviously is a clear demand for, or at least would be a clear benefit from moving from the current, largely uncoordinated, practice of desktop-centric application of GIS toolsets towards a distributed SDI, the questions of leadership, coordination and implementation have to be answered.

With the aim of establishing a platform for Geoinformatics professionals the Kyrgyz GIS Association (KgGISA) has been founded in 2009, as it turned out to be too difficult to implement a single organization across Central Asian countries. This association is now attempting to bring together regional stakeholders from all different types of institutions with the aim of working towards an SDI. This grassroots approach is a community-building effort towards jointly reaching common goals more effectively.

3. STRATEGIES FOR ESTABLISHING AN SDI

Ways towards national and regional SDIs are well documented in the pertinent literature (e.g. Singh, 2010; Akinyemi *et al*, 2010, Fonseca *et al*, 2009, Harvey *et al*, 2007), but virtually all of these instances succeeded through a top-down approach, some of them with multiple stakeholder involvement. Very little experience exists, though, with community-based initiatives in the absence of governmental or other institutional leadership.

Therefore approaches will have to differ from proven strategies, and necessarily will have to be tied into international cooperation frameworks. With regional research funding and development support not existing, the stakeholders interested in establishing a common-use geospatial platform will have to look towards themselves. As very substantial savings can be realized from the potential synergies and reduction of redundant efforts, such a self-supporting view towards existing project budgets is quite realistic.

There still are two major issues, namely leadership and technical platform. Virtually all stakeholders do not have a core interest in geospatial information and technologies, but consider themselves users with the aim of supporting their respective applications and topical interests. This requires one institution to take the lead, in the particular example presented here this is a collaboration of KgGISA with ACA*GIScience.

Secondly, even though an SDI today typically is a fully distributed set of services where each partner maintains and hosts 'their own' data sets, like every infrastructure an SDI needs several components tying data services together, creating value added beyond the sum of various components – like the switches, load balancers, control rooms, meters and other equipment facilitating our electrical energy infrastructure.

In the case of an SDI, the components integrating distributed data services are data catalogs holding metadata and portals as user interfaces and access points. In the present case, these critical building blocks are provided by researchers in a European research center, guaranteeing higher availability and stability than would be feasible at the sites of regional partner institutions.

Infrastructures -and thus SDIs- in a true sense are built by one or many actors based on established standards, and are made available to multiple users for different purposes. These users and uses might not all be known from the outset, and might even operate under different business models. Initiating an SDI therefore has to anticipate the yet-unknown, has to be open and flexible – which is more easily said than done.

Looking at recent developments in standards, software technologies and overall connectivity, it can be safely stated that developing an SDI is more than ever not so much a technical as it is a people and thus organizational issue (de Man, 2006). Even when a technical question like adjustments to a metadata standard are at stake, it is the people who have to agree reconciling the interests of their institutions.

In a region where institutions currently face difficulties either establishing themselves or adjusting to changing frameworks, the challenge of establishing a community-based peer organization which can only rely on volunteer contributions supporting common interest is a major one. Chances for ultimate success only can be judged in hindsight, but the major steps to be taken, and currently being taken, are actually well known:

1. Create awareness for the benefits of an SDI, make sure this concept is recognized not only as a new name for a geospatial database.
2. Offer educational opportunities for stakeholders to learn about the components, protocols, standards and tools to build an SDI.
3. Move actors from a desktop mindset towards network thinking and on to an infrastructure philosophy by providing best practice examples.
4. Build a community of practice (WENGER, 1998), any infrastructure is much more likely to be used well when stakeholders could participate in and identify with specifications.

5. Get over the hurdle from 'talking infrastructure' to building, using and living in an infrastructure by 'just doing it' instead of waiting for it to be done.
6. Establish mechanisms to make an SDI sustainable, to establish a platform where common interests can be identified

Clearly, all of the above again needs leadership and organizational frameworks. Where these do not exist in public institutions, a community effort is needed to fill the gap – and the results might easily be more valuable for a majority of users and applications than in the case of highly centralized SDI efforts frequently leading to resources only difficult to access by most end users.

4. GEOPORTAL: REGISTRATION, DISCOVERY AND ACCESS

The above mentioned approach of '*just doing it'* sounds maybe naïve and hard to achieve at the same time. Still, creating awareness, offer learning opportunities and ultimately convincing stakeholders to 'buy into' the concept and practice of an SDI requires access to concrete implementations, otherwise the idea will remain too abstract for most practitioners.

The initial implementation focus therefore was on offering a geoportal as an easily accessible interface sitting on top of an increasingly complex architecture of services. For developing this application, the authors used the ESRI ArcGIS Server Geoportal Extension (http://www.esri.com/software/arcgis/geoportal) providing the technological basis for sharing and reusing resources across applications and communities. The Central Asia geoportal application supports finding, accessing, presenting and maintaining information required by a large community under the principle of 'truth-in-labeling'. It can be currently accessed at http://geoportal.aca-giscience.org.

This portal primarily is an interface to a registry database of metadata, pointing to online and offline data sets, whether freely available or not. Essentially, this geoportal binds together a dispersed and disparate set of data and services into an initial infrastructure. From there, the portal serves as a focal point for organizational measures and simultaneously for community building.

The Geoportal Application includes the following services:

- Online metadata editor that enables data and services producers to document their respective assets.
- Search and discovery functionality enables users to find the available geographic resources, to assess their fitness for use, and to exploit them.
- Data / services viewing application.
- Metadata harvesting for integration of metadata from other repositories

The geoportal aims at establishing a communication platform between users and data producers, following the 'publish-find-bind' patterns of web services architectures (Ostensen, 2002). Metadata published by data holders using an online metadata editor is gathered in the metadata repository. Authenticated users search for information by querying the portal's metadata catalogue and can retrieve geographic resources from the host content service (if and where available).

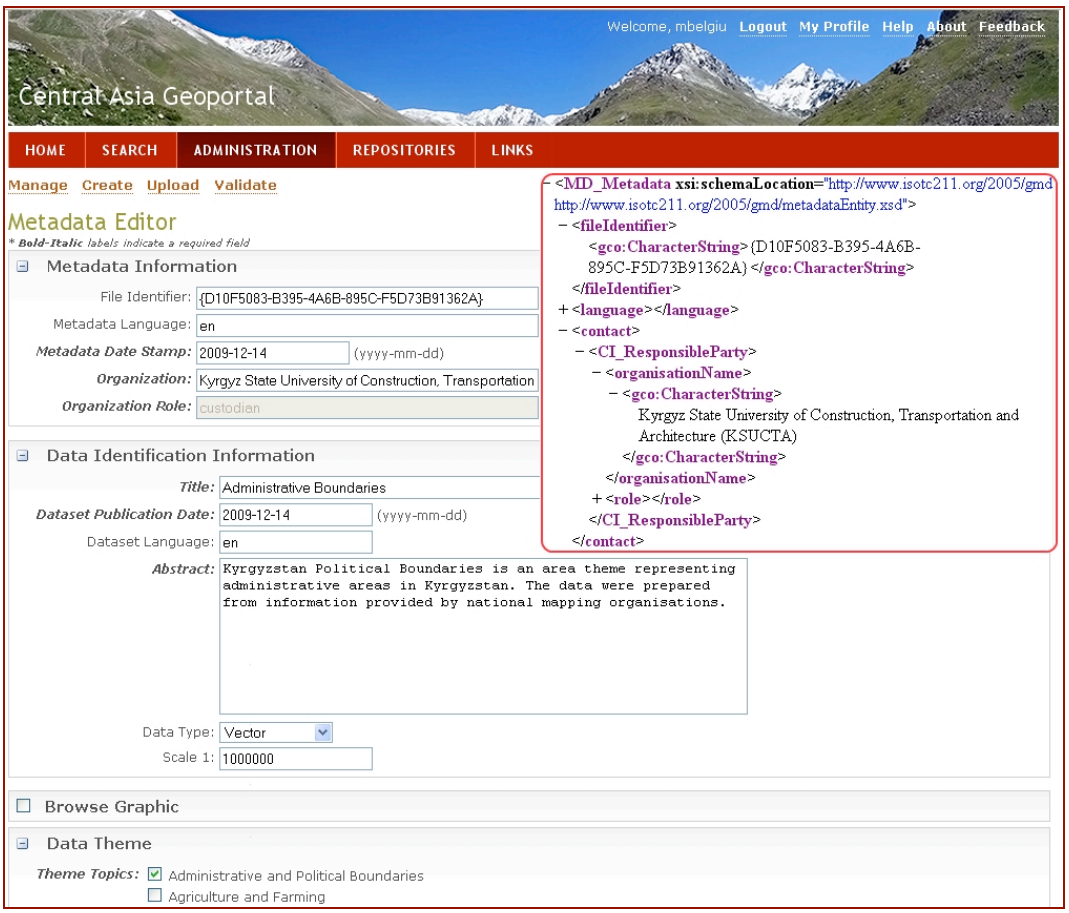
The publishers are able to bookmark or share their metadata records (i.e., through Dig, Google, Facebook or Twitter) using the "AddThis" universal sharing platform. This functionality enables both metadata publishers and users to set up discussions about the content of the described data.

Figure 1: Link to metadata records published on the geoportal application

4.1 Metadata Editor

Metadata is a critical component within Spatial Data Infrastructure frameworks. It guarantees the "long term traceability" (Nogueras-Iso *et al*, 2005) of spatial data and services, supporting data discovery, assessing data fitness for use, data access and data use. The availability of comprehensible documentation generates advantages for both data providers and data users. While the users are provided with adequate information to find and access the data and use it in an appropriate context, data custodians are able to organize and maintain their investment in data in a systematic and structured way. As long as the public or private organizations invest a small amount of time and resources into documenting assets, they stand to receive "dividends in the future" (SDI Cookbook, 2009).

Figure 2: The available online Metadata Editor enables spatial data or services producers to publish structured information about their assets



The geoportal application provides data producers with an online metadata editor that facilitates publishing geographic metadata in conformance with international standards (i.e. ISO 19139/19115 Metadata Standard for Dataset).

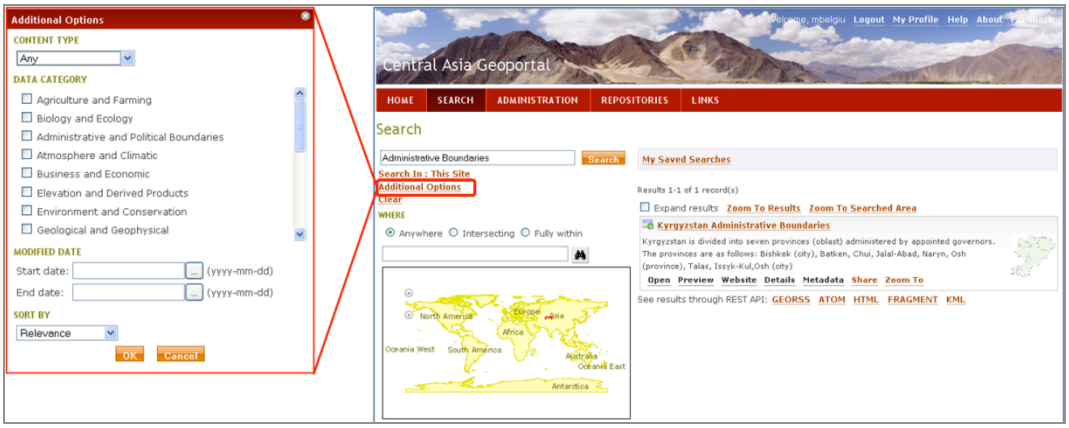
Online metadata contains the metadata elements that are defined as mandatory in the standards and a few optional elements that are considered critical for assessing whether the discovered spatial data achieves users' requirements. Thus, the metadata publishers should provide as much information as possible because insufficient documentation makes it difficult or even impossible for users to discover the needed resources and to assess whether a given geographic resource is useful for their tasks.

4.2 Search and Discovery

The geoportal application supports distributed search through a single user interface to all metadata collections leading to valuable resources, such as maps, data, web services or analytical models (Athanasis *et al*, 2009). Searching for geographic information in the geoportal is based on keywords. The search results can be narrowed using either the spatial filter or additional options filter.

While the spatial filter enables users to define the spatial extent of the records to be returned from the search action, the additional options help users to narrow their queries by content type, data category or modified dates. For instance, the data category filter defines all nine topic categories specified in ISO 19115 metadata standard: ranging from administrative and political boundaries or agriculture and farming to biology and ecology or utilities and communications. Through this search functionality, the users will retrieve the metadata records that have the selected data category defined in the data theme tag.

Figure 3: Users can narrow the search results by selecting spatial or thematic filters

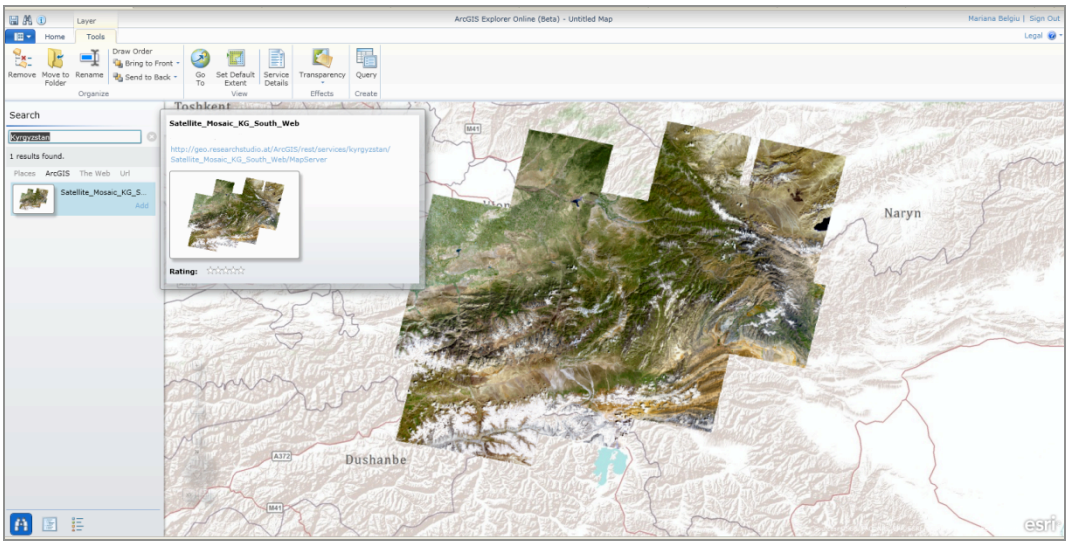


The search results can be evaluated and integrated into different viewing or mapping applications, such as Google Earth (e.g. via the KML format).

4.3 Services Integration into Client

The CSW clients enable searching CSW catalogues (OGC Catalog Service for the Web) directly through ArcGIS Explorer Online. Thus, data identified as “Live data and Maps” can be added as a new layer and visualized in ArcGIS Explorer Online (Figure 4). This GIS Viewer is a free application that helps data producers to provide visual access to their assets within and between larger communities.

Figure 4: Visualization of a spatial dataset in ArcGIS Explorer Online



5. COMMUNITY OUTLOOK AND PERSPECTIVES

As Erik de Man (2010) has stated, SDI frameworks are established through context-dependent practice. This statement is fully supported by the authors experience in Central Asia and in the Kyrgyz Republic, where contexts are definitely different from many other regions.

Still, many observations are applicable beyond this particular context:

- The potential (traditional?) leading role in SDI for national mapping agencies in many countries today is obsolete due to the focus on very diverse and distributed thematic data, restrictive policies controlling 'public' base maps and an increasing availability of alternative base data.
- In particular where high quality survey / cadastral foundations are not functional or non existing, 'natural' lead agencies are not available.
- Remote sensing increasingly takes over a part of the base map / reference base role, made accessible for broad range of users through geospatial browsers now available with global coverage.
- While this development substantially reduces the role of state / government institutions, their peer-level participation is as desirable as it is hard to achieve.

Today 'communities of geospatial practice' are making progress establishing themselves as partnerships setting up, promoting and maintaining SDIs. Lessons learned from the evolution of Volunteered Geographic Information (VGI) and from the emergence of distributed peer-level development of Wikipedia-style knowl-

edge can be applied to community based SDIs. The longer term sustainability and stability of 'Volunteered Geographic Data Infrastructure Elements', though, still needs to be proven.

The principle of '*sharing information is more important than having it'* still has a long way to go to be fully understood and embraced. Education, communication and community building is the only way of getting there.

ACKNOWLEDGEMENTS

Support to the author representing ACA*GIScience by the Global Spatial Data Infrastructure Organization (www.gsdi.org) through their Small Grants Program is gratefully acknowledged.

The Eurasia-Pacific Uninet (www.eurasiapacific.net) supports collaboration between universities from Central Asia and Austria through its Austria-Central Asia Center for GIScience (www.aca-giscience.org).

The GIScience Research Cluster Salzburg (www.giscience-research.org) is currently hosting the demonstrator Geoportal for Central Asia and is providing technical infrastructure support through its SDI working group.

Particular thanks go to Manfred Mittlboeck from the iSPACE Research Studio (ispace.researchstudio.at) for development of the metadata profile used in the above outlined application as well as essential architecture elements.

REFERENCES

Athanasis, N., Kalabokidis, K., Vaitis, M., Soulakellis, N. (2009), Towards a semantic-based approach in the development of geographic portals. In: Computers and Geosciences, Volume 35, Issue 2, pp. 301-308.

Harvey, F., Tulloch, D. (2006), Local government data sharing: Evaluating the foundations of spatial data infrastructures. In: International Journal of Geographic Information Science, Volume 20, Issue 7, pp. 743-768.

Fonseca, F., Davis Jr. A. C., Camara, G. (2009), Spatial Data Infrastructure for Amazon: a first step towards a global forest information system. In: Earth Science Informatics, Volume 2, issue 4, pp. 189-192.

Giff, G., B. van Leonen, Crompvoets, J., Zevenbergen, J. (2008). Geoportals in Selected European States: a non-technical Comparative Analysis. In: Int. Journal of Spatial Data Infrastructures Research, Special Issue GSDI-10.

GSDI Cookbook, 2009: http://www.gsdi.org/gsdicookbookindex, accessed on 19-th of June, 2010.

Man, E. de (2006). Understanding SDI; complexity and institutionalization. In: International Journal of Geographical Information Science, Volume 20, Issue March 2006 , pp.329 – 343

Man, E. de (2010). Spatial Data Infrastructuring: praxis between dilemmas. In: Int. Journal of Spatial Data Infrastructures Research, article under review

Masser, I. (2007). Building European Spatial Data Infrastructures. ESRI Press,91p.

Masser, I. (2005). GIS Worlds: Creating Spatial Data Infrastructures. ESRI Press, 336 p.

Masser, I. (1999). All Shapes and Sizes: The First Generation of National Spatial Data Infrastructures, International Journal of Geographical Information Science, Vol. 13 (1), pp. 67- 84.

Ostensen, O. M., Smiths, P. C. (2002), ISO/TC 211: Standardization of geographic information and geo-informatics, Geosciences and remote Sensing Symposium, IGARSS.

Singh, K. P. (2009), Spatial Data Infrastructure in India: Status, Governance Challenges and Strategies for effective functioning. In: International Journal of Spatial Data Infrastructures Research, Vol. 4, pp.359-388.

Strobl, J. and A. Nazarkulova (2009). Places Along the Information Highway: the Long Path towards Spatial Data Infrastructures. In: Proceedings Third Central Asia GIS Conference - GISCA'09 "GIScience for Environmental and Emergency Management in Central Asia". KSUCTA, Bishkek: pp. 16-21.

Abdykalykov, A.A., Strobl, J. and S. Jorobekova (2010): Proceedings of the Fourth Central Asia GIS Conference - GISCA'09, "Water: Life, Risk, Energy and Landuse". 122 pp. Austria - Central Asia Centre for GIScience. Bishkek.

Abdykalykov, A.A. and Strobl, J. (2009): Proceedings of the Third Central Asia GIS Conference - GISCA'09, "GIScience for Environmental and Emergency Management in Central Asia". Austria - Central Asia Centre for GIScience. Bishkek.

Strobl, J. et al (2009). Proceedings Third Central Asia GIS Conference - GISCA'09 "GIScience for Environmental and Emergency Management in Central Asia". KSUCTA, Bishkek.

UNGIWG (2010): United Nations Geographic Information Working Group – Documents and Meetings at http://www.ungiwg.org/documents.htm [accessed 12 June 2010].

Wenger, Etienne (1998). Communities of Practice: Learning, Meaning, and Identity. Cambridge University Press.