

Geoinformatika *Geoinformatics*

Geographic information infrastructure in Lithuania – components and data themes

Giedrė Beconytė

Vilnius University, e-mail: giedre.beconyte@gf.vu.lt

Cindy Pubellier

GeoLines France, e-mail: cindypubellier@fastem.com

INTRODUCTION

Building geographic information infrastructure (GII) in Lithuania is the major task of national geographic information management. Several studies have been performed (Assessment..., 2004; Framework Report, 2004; Feasibility Study, 2004), which resulted in decisions about the principal architecture of the infrastructure. The problems expected to be solved by the GII, main target groups and necessary strategic interventions have been discussed in our previous paper (Beconytė, 2006). This article describes some implementation issues, such as structural components, different aspects of development and priority data themes to be integrated in the geographic information infrastructure.

THE STRUCTURE OF GII

The main structural parts of the Lithuanian GII are shown in Fig. 1. Physically, the Lithuanian Geographic Infrastructure consists of a network of central and peripheral *nodes* which provide geographic data services and are interconnected by regular flows of data and knowledge, and *clients* who use these services. A partly centralized data model was chosen in order to achieve an uninterrupted availability of metadata and to guarantee that the geographic data come from their primary sources.

The platform chosen for this network has to be the most popular and widely used, i.e. the Internet. In ad-

dition, impacts on the organisations, as a result of the implementation of GII, must be reduced to a minimum. The participating organisations should continue to use the same information environment as before without a major re-engineering effort. Data should be able to be retrieved from an operating system to another environment. Users need not care about the change of format, projection, and other compatibility issues.

The network must be capable of retrieving data, upon request, from the provider and converting it to a universal standard transportable by the network. The data are then transferred to the user's node, converted to the user's proprietary data format, and placed to the user's system. This, no doubt, demands a universal interoperable data exchange standard, protocols for transport, data enveloping and data body over the Internet, as well as tools of Web Map services to be implemented.

Creation of the *national metadata system (register of metadata)* and a system of *metadata services* is the first priority as it serves as the basis for harmonization of information access (Onsrud and Rushton, 1995). Legal leverage over collection and provision of metadata is necessary to ensure that all public sector information is documented in the catalogue. It must be initiated by the coordinating governmental body. For a consistent metadata register a single standard for geographic metadata must be adopted, the database created and tools developed for easy transfer of metadata online, geographic and textual search for metadata and their updating.

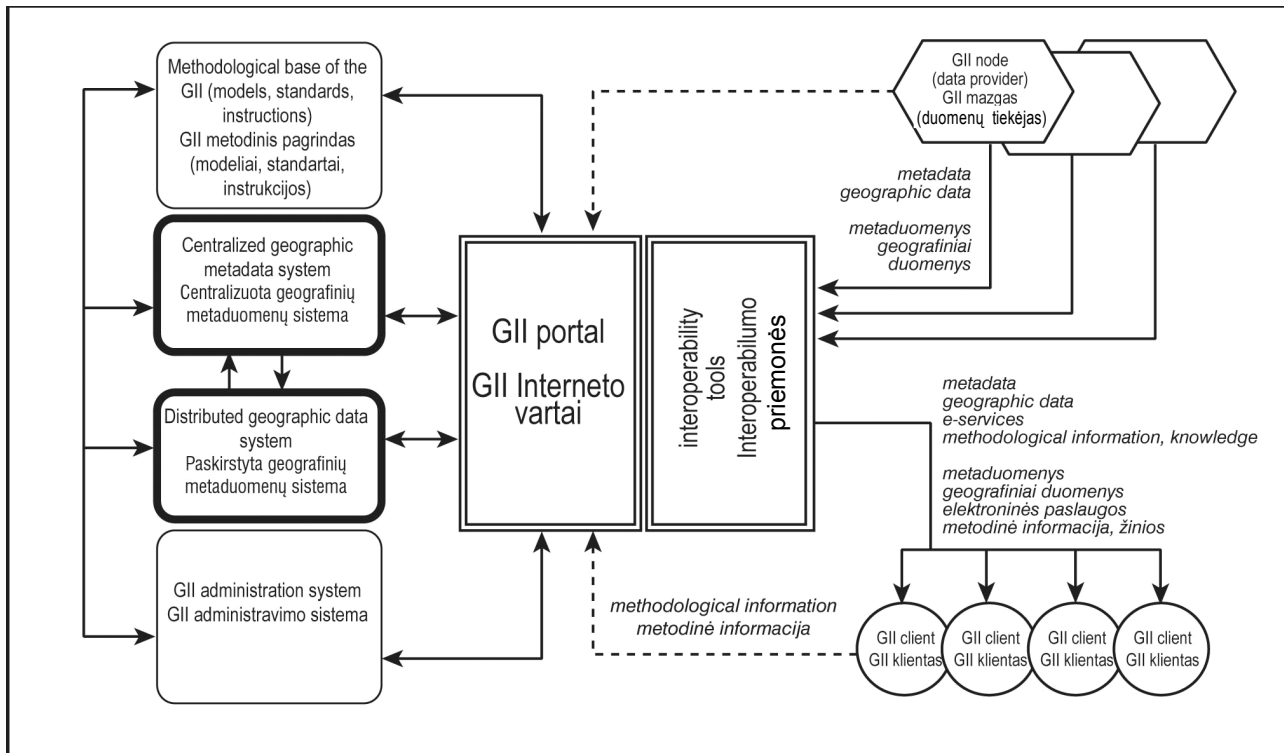


Fig. 1. Main components of Lithuanian GII
1 pav. Pagrindiniai Lietuvos GII komponentai

A single *national geographic reference system* (databank) is the second core component to be created to ensure a proper functioning of the GII. It requires a single geographic data scheme, geographic DBMS and services to support vendor neutral data exchange formats, common industry standards and data conversion on demand (both ways: for data providers and users), extract-transform-load procedures, Internet mapping, geographic and textual search and different access privileges.

These two major components form a partly centralized public data system that provides reference data for all users of the infrastructure and serves as a basis for geographic integration of all State data banks in the future. This system is responsible for publication and exchange of geographic data. It provides services for metadata search and data download. It will be regulated by the National Land Service as the authoritative coordinating body at governmental level and maintained by its state enterprise “GIS-Centras”.

Some of reference data themes like roads, forestry, and hydrography are maintained by other institutions and at scales different from that of the national reference database. One of the most important tasks of the GII is to ensure a proper linking of such data instead of replicating them in the central node and to keep the system transparent to the users. The National Land Service is responsible for political actions such as legal regulations necessary to ensure the robustness and sustainability of the infrastructure, and the necessary institutional arrangements.

All producers and owners of geographic data of national or regional importance are encouraged to connect to the infrastructure as its peripheral *nodes*. Provision of their data (different thematic layers) to be accessible through the Lithuanian GII portal is subject to negotiations and mainly based on the free will of the institutions. Technologically, the infrastructure is planned as to require minimal interventions into existing activity models and data schemes.

There are two types of peripheral nodes. *Type 1* nodes refer to nodes owned by providers of reference data and national geographic data, which are capable of providing geographic information services. They are responsible for maintaining their own information. *Type 2* nodes refer to nodes owned by providers of reference data and national geographic data, which are not yet capable of providing geographic information services on their own.

The processes of electronic data exploitation can be provided for open access through the main web services, based on common specification for all nodes:

- discovery services (search tools);
- view services (map navigation tools);
- download services (data extract tools);
- transformation services (optional for Lithuanian market because of a common system of coordinates accepted, but mandatory for international data access);
- other generic geoprocessing services relevant to the business of the data provider.

Interoperability tools assure an efficient sharing of geographic data in diverse semantic and technological

formats. A uniform object model and a classifier of geographic objects are prerequisites to make the spatial data objects semantically and geometrically consistent, which is also very important for interoperability.

THEMATIC DATA

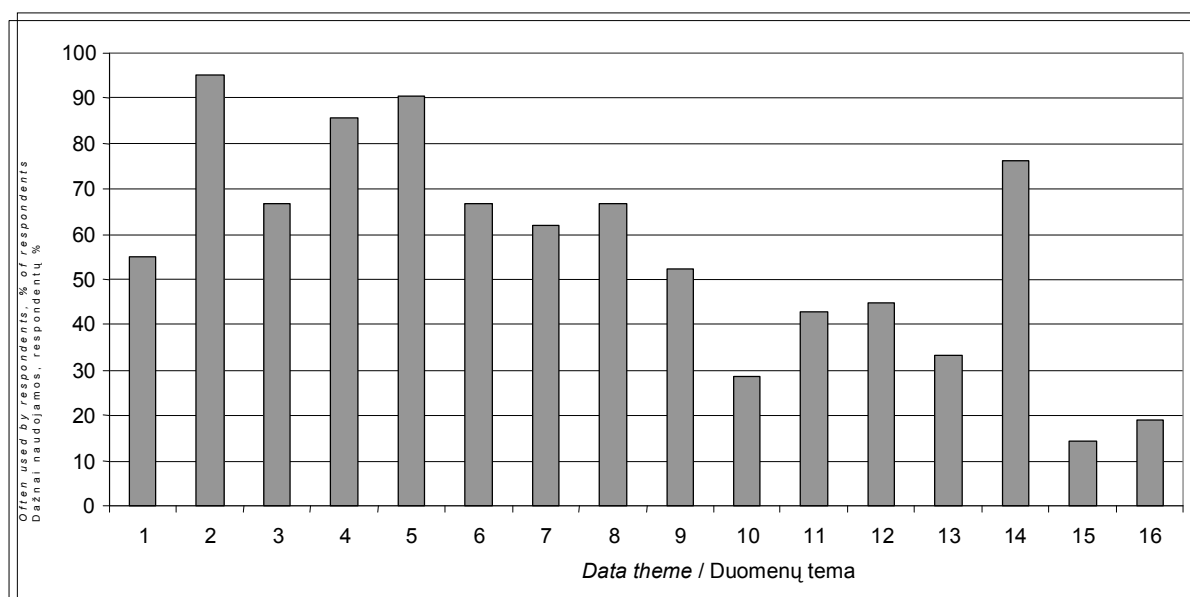
Key data themes

Not all the datasets can be incorporated and into shared through the infrastructure immediately. The project team communicated with over 50 organizations of different types approximately representing the structure of the active players in the field of geographic information in Lithuania by sectors. The following chart shows which data themes are currently most intensively used according to the responses to the distributed questionnaires. As it could have been expected, the seven “Common Reference Data” themes (namely Geodetic reference, Units of administration, Units of property rights, Ad-

resses, Topography, Orthoimagery and Geographic names) defined by INSPIRE also fall within the set of data themes often used by more than 50% of respondents in Lithuania.

It is evident that the mentioned themes include the *reference data* – a limited set of most widely used data themes such as transportation, hydrography, cadastral and administrative boundaries, elevation, digital imagery, and geodetic control, which provide a base for collecting, registering, integrating and analyzing statistical data. Reference layers are (or should be) publicly available, maintained for the common good, useful for many purposes, and each is likely to comprise at least a subset of that data layer for any particular Information Community.

The most significant benefit from all reference data themes is that they will be commonly approbated, compliant with European reference systems and will allow easy visualization or spatial connectivity between different data components. Besides the above mentioned



- 1 – Basic geodetic control data / *Geodezinio pagrindo duomenys*
- 2 – Administrative units / *Administracinės ribos*
- 3 – Land cadastre / *Nekilnojamojo turto kadastro informacija*
- 4 – Transportation / *Transportas*
- 5 – Hydrography / *Hidrografija*
- 6 – Addresses / *Adresai*
- 7 – Height, DTM / *Aukščių taškai, skaitmeninis reljefo modelis*
- 8 – Orthoimagery / *Ortofotovaizdai*
- 9 – Toponymy / *Vietovardžiai*
- 10 – Geological data / *Geologiniai duomenys*
- 11 – Network of electric power lines / *Elektros perdavimo linijos*
- 12 – Pipe lines / *Vamzdynai*
- 13 – Telecommunication network / *Telekomunikacijų tinklai*
- 14 – Forests / *Mišakai*
- 15 – Civil protection / *Civilinės saugos duomenys*
- 16 – Various other themes / *Įvairios kitos temos*

Fig. 2. Use of different data themes by enterprises, % of enterprises

2 pav. Skirtingų duomenų temų naudojimas įmonėse ir įstaigose

benefits from the GII as an e-service, there are quite many agencies and organizations which will benefit from standardisation and completion of particular reference data themes.

Interventions

It can be stated that striving for the interoperability of national geographic database, state registers and other databanks of national importance, some interventions are essential for all the reference data themes. With the exception of some specific shortcomings, there are several common characteristics of all the reference themes in Lithuania. A general summary of the type of activities and the extent to what they must be planned to improve the interoperability of each of reference data themes is as follows.

1. Implementation of GII policy in respective fields:

1.1. Legal regulations (liability of institutions for different datasets, clear financing, copyright and pricing policies, data exchange regulations);

1.2. Adopting standards (state classifier system, meta-data and data standards, and standards for data quality).

2. Developing an information system in a corresponding field ready to be integrated in the LGII (technological interventions):

2.1. Developing the database management system;

2.2. Harmonization of data flows among institutions;

2.3. Creating an infrastructure for publishing meta-data (data) in the LGII Internet portal and effective exchange of data among the institutions.

3. Data collection:

3.1. Collection and digitizing of missing referenced data;

3.2. Updating of outdated datasets.

These activities have different priorities in the context of LGII short-term goals. The following table shows the types of intervention necessary for the themes to meet the short-term goals of the project.

All these data-related interventions undergo some risks, especially when it is very difficult to fully rely on state subsidiaries while producing geographic information.

INTEROPERABILITY

Local pieces of geospatial information could be created once in many places and tied together into a multi-layered base-map on which value-added layers and services will be built. The mortar required to hold that mosaic together consists of a variety of standards, including those for data exchange and documentation, data content and interoperability.

The high level international standards to be adopted for LGII are:

- standards for metadata,
- standards for geographic data classification,

Table. Priority interventions for harmonization of reference data themes (priority: high (+++), medium (++) or low (+))
Lentelė. Poveikio priemonės, būtinos referencinių duomenų temų harmonizavimui (prioritetas: didelis (+++), vidutinis (++) arba žemas (+))

Reference theme <i>Referencinių duomenų tema</i>	Type of intervention / <i>Poveikio priemonės</i>						
	1. Policy building <i>Politinės</i>		2. Technological <i>Technologinės</i>		3. Data collection <i>Duomenų kaupimo</i>		
	1.1	1.2	2.1	2.2	2.3	3.1	3.2
Geodetic control information <i>Geodezinio pagrindo duomenys</i>	+	+++	+++	+++	+++	+++	+
Administrative units <i>Administracinės ribos</i>	+	++	+	+	+++	+++	+
Land and real estate cadastre <i>Nekilnojamojo turto kadastras</i>	++	++	+	+++	+++	+	+
Addresses <i>Adresai</i>	+	+++	+++	+++	+++	+++	+
Hydrography and coastline <i>Hidrografija ir kranto linija</i>	+	+++	++	+++	+++	+	+
Transportation networks <i>Transporto tinklai</i>	++	+++	+	++	+++	+	+
Elevation data <i>Reljefo duomenys</i>	+	+++	+	+++	+++	+++	+++
Toponymy <i>Vietovardžiai</i>	+++	+	+++	++	+++	+	+++
Orthoimagery <i>Ortofotovaizdai</i>	+	++	+	+	+++	+++	+
Forests <i>Miškai</i>	+	+	++	+++	+++	+	+

- standards for data transfer,
- quality standards.

Standardisation should be widely promoted amongst all geographic information users, focusing on easy conversion among the formats, data translation without any loss of information and protecting users' investments (Dueker and Vrana, 1995). The users should be also given an incentive to adapt to new *standards* in order to sooner overcome such common in Lithuania problems as duplication of data, variety of data sources, incomplete, inconsistent or outdated reference data, lack of documentation (metadata), incompatible and inconsistent datasets and limited access to a wider geoinformation market.

As described in the previous chapter, some actions at different levels are needed to overcome the problems of low interoperability among the existing databanks. The task of *data harmonization* is to ensure that different reference data themes are combined, used and analysed together:

- all data themes must be validated and harmonized with the other themes;
- each data theme must be validated and harmonized across its sub-themes and different scales;
- all relevant non-geographic databases must be linked to the reference data through their spatial component.

The levels of harmonization of common data should cover:

- semantic interoperability;
- topological consistency (edge matching, shared topology);
- access harmonization (all public information databases must function on interoperable technical platforms).

There are two semantic interoperability issues to be settled by the GII:

- semantic consistency – a common object model, classifier of all geographic objects and unique identifiers (when every geographic object is assigned a unique code for referencing within the country);
- cross-scale compatibility (generalisation levels specified in a way that ensures possibility of combining and analysing together the data at different scales).

Summarizing the previous experiences of relevant developments of components for geographic information infrastructures in the other countries (EuroSpec activity, 2003, GINIE 2000, Masser, 2005), it should be stressed that important success factors for data harmonization are:

- active governmental support;
- interaction and coordination of the involved institutions, strong multi-sector coordination;
- overall strategy and early adopted top-down approach;
- decentralization and involvement of the private sector from the outset;
- adopted common standards and regulations;
- clear data policy, licensing framework and access procedures;
- flexibility of the system;
- technical support as well as leadership timely provided;
- implemented quality control procedures;

- transparency implementing the strategy and proper communication to the public.

The Geographic Information Infrastructure also provides a scientific understanding of the environment around us. Therefore, it is essential to set up a *unified data model* meeting the following requirements:

- a multipurpose infrastructure. The rules apply on data modelling. The data structuring or the data model should be based on the principle that data should be independent of the functions or actions taken. So, the GII data model should represent generalised view of the real world and existing objects;
- scalability: the need for continuous system's expansion with new data sets. That means that the system and data storages should be extendable with new elements (objects and features) without any hindrance of normal working flow or major data restructuring;
- the conceptual model should be able to be implemented as a whole, or partially by components, depending on the requirements of a particular system or application.

To achieve this objective, it is necessary to establish appropriate coordination mechanisms and common rules for data management over the Lithuanian GII. The most important is the establishment and maintenance of a common data model for all GII objects. This is especially important in case of modelling the objects linked to the same kind of real world objects. This will allow starting a harmonised spatial data infrastructure that will facilitate a combination of information of various sources and a more advanced analysis.

CONCLUSIONS

A partly centralized system model has been chosen for the GII in Lithuania. A centralized national metadata system allows a fast and easy search and access to existing metadata. The central server is responsible for collecting metadata together into a single database and making them available in one format over the Internet. Decentralization of the geographic data system reflects the actual distribution of different geographic datasets – their owners become the so-called nodes of the LGII, which transfer their data to the GII system on demand using unified gateways. Due to such structure, the impact on existing internal business models of geographic data producers and managers can be minimized and organizations of different levels of preparedness can participate in GII as its functional nodes. Some nodes of the future LGII are of special importance (municipalities, registers and cadastres, environmental and cartographic information systems). They form potential local/thematic spatial data infrastructures integrating different data themes. It is expected that in the future all GII nodes will be able to provide a full set of electronic geographic information services.

Ten reference data themes are of special importance and form a vision of the Lithuanian GII informational component. But seeking for a seamless and flawless

system of basic geographic information in Lithuania, some interventions at different levels must be made. Among them, terms and conditions of access to and use of geographic information should be designed to facilitate the use of the corresponding data by individuals and organisations.

ACKNOWLEDGMENTS

The documents and studies provided by the LGII development team in 2004 were largely used to prepare this paper. More information on the ongoing project is available at the LGII Internet site maintained by State Enterprise "GIS-Centras" (<http://www.gis-centras.lt/gisweb/index.php?pageid=137>).

Received 29 09 2006

Accepted 20 10 2006

References

1. Assessment for priority intervention needs. *Development of Lithuanian Geographic Information Infrastructure (LGII) Project Preparation* (2004). ASTEC Global Consultancy. Vilnius.
2. Beconytė G., Pubellier C. 2006. Geographic information infrastructure – towards interoperability of the public data. *Geography*. 42(1). 23–27.
3. Framework Report. *Development of Lithuanian Geographic Information Infrastructure (LGII) Project Preparation* (2004). ASTEC Global Consultancy. Vilnius.
4. Feasibility Study. *Development of Lithuanian Geographic Information Infrastructure (LGII) Project Preparation* (2004). ASTEC Global Consultancy. Vilnius.
5. GINIE D 2.7.2 (2000). *Geographic Information Policies in Europe: Recommendations for Action*.
6. EuroSpec activity (2003). EuroGeographics. URL: http://www.eurogeographics.org/eng/03_projects_eurospec.asp (last visited February 2006).
7. INSPIRE Data Policy and Legal Issues Working Group Position Paper (available at <http://www.inspire.jrc.it>)
8. Masser I. (2005). GIS Worlds. *Creating Spatial Data Infrastructures*. Redlands, CA: ESRI Press. 312 p.
9. Onsrud H., Rushton G. (1995). "Metadata issues in sharing geographic information". *Sharing Geographic Information*. New Brunswick: Rutgers, The State University of New Jersey. 499–501.
10. Dueker, K., Vrana, R. (1995). "Systems integration: a reason and a means for data sharing". *Sharing Geographic Information*. New Brunswick: Rutgers. 149–71.

Giedrė Beconytė, Cindy Pubellier

GEOGRAFINĖS INFORMACIJOS INFRASTRUKTŪRA LIETUVOJE – KOMPONENTAI IR DUOMENYS

Santrauka

Lietuvos geografinės informacijos infrastruktūros (GII) kūrimo tikslas – sukurti skaidrias ir palankias geografinių duomenų

naudojimo įvairiems tikslams sąlygas, įskaitant sprendimų priėmimą ir pridėtinę vertę turinčių paslaugų bei produktų kūrimą valdymo, verslo, socialinėje, kultūros ir kitose srityse, taip didinant geografinės informacijos naudojimo apimtį ir efektyvumą. Straipsnyje aprašoma principinė šios infrastruktūros struktūra ir jos informacinį branduolį sudarančios duomenų temos.

Du pagrindiniai infrastruktūros komponentai yra nacionalinė geografinės informacijos metaduomenų ir nacionalinė geografinės informacijos sistemos (1 pav.).

Nacionalinė geografinių metaduomenų sistema turi tapti bendra ir visiems prieinama geografinės informacijos dokumentavimo sistema, pagrįsta ISO 19115 geografinių metaduomenų standartu. Pasirinktas centralizuotas metaduomenų sistemos modelis garantuoja, kad metaduomenys bus greitai ir lengvai pasiekiami bet kuriuo metu. Geografiniai duomenys, kurių rinkinių yra nedaug, tačiau labai didelės jų apimtys, bus saugomi ir tvarkomi ten, kur tai yra efektyviausia – jų gamintojų ar valdytojų institucijose. Taigi nacionalinė geografinių duomenų sistema turi būti decentralizuota, tačiau pagrįsta vienu duomenų modeliu, suderinamu su pagrindiniais tarptautiniais informacinių sistemų kūrimo standartais. Šios sistemos pagrindą sudaro GII mazgai institucijose, teikiančiose ar administruojančiose svarbius geografinius duomenis. Tarp visų infrastruktūros mazgų turi egzistuoti interoperabilios sąsajos (techninės, programinės ir standartizavimo priemonės), kad jų teikiama geografinė informacija ir paslaugos būtų integruotos į nacionalinę geografinių duomenų sistemą, o iš ten Internetu pasiekiamos naudotojams jiems patogiais, tarpusavyje suderinamais formatais. Numatoma teikti šias elektronines paslaugas, pagrįstas bendromis specifikacijomis: paieškos, peržiūros (žemėlapių navigacijos), duomenų atsisiuntimo, koordinatų transformavimo bei kitas paslaugas, susietas su specifine duomenų teikėjo veikla.

Išanalizavus dabartinį nacionalinių duomenų rinkinių naudojimą ir Europos INSPIRE iniciatyvos rekomendacijas, matyti, kad ypač svarbios yra šios dešimt duomenų temų (daugiau kaip 50%; 2 pav.): geodezinio ir kartografinio pagrindo duomenys, administracinės ribos, hidrografija, transporto tinklai (ypač automobilių kelių duomenys), nekilnojamojo turto kadastro duomenys, adresai, ortofotovaizdai, reljefo duomenys, vietovardžiai ir miškai. Jų integravimui į GII turi būti skiriamas ypatingas dėmesys. Siekiant pagerinti duomenų temų naudojimą ir kokybę, būtinos kelių lygmenų poveikio priemonės: GII politikos įgyvendinimas atitinkamose srityse (teisinės ir standartizavimo priemonės), informacinės sistemos duomenų teikėjo institucijoje plėtra (duomenų bazių valdymo sistemos išplėtojimas, duomenų srautų tarp institucijų harmonizavimas, technologinės infrastruktūros kūrimas) bei duomenų rinkimas (trūkstančių duomenų gavimas, pasenusių duomenų atnaujinimas). Šių priemonių aktualumas kiekvienai temai parodytas lentelėje.

Viena svarbiausių sėkmingo GII įgyvendinimo sąlygų yra bendras valstybės geografinių duomenų modelis. Siekiant įdiegti tokį modelį, būtinas strateginis geografinės informacijos valdymo koordinavimas valstybės mastu.