UN-GGIM EUROPE

UNITED NATIONS
COMMITEE OF EXPERTS ON
GLOBAL GEOSPATIAL
INFORMATION MANAGEMENT

Report of the Work Group “Data Integration” about how to manage side-effects induced by data combinations

Content

1 Executive Summary

This report of the UN-GGIM:Europe Work Group B on Data Integration focuses task 3 of the work plan on “managing side-effects induced by data combinations”. This overall assignment resulted in the identification of three objectives:

1. Identification of side-effects (section 5 and annex I),
2. Evaluation of the pragmatic dimension of interoperability (section 4, 6, 8) and
3. Possibilities to reduce, influence and steer side-effects (section 7 and annex II).

The main deliverable was to define selected side-effects that could be influenced in an easy way, leading to simple and usable recommendations regarding the combination of geospatial and statistical data. However, in practice it appeared that realising this deliverable has been much more complicated than expected beforehand. Identification of side-effects lead to a long list of notions, which were agreed to be side-effects. The list of agreed side-effects has not been completed. Nevertheless, the working group put some effort on the categorization of side-effects and delivered the three occurrence levels of side-effects.

Evaluation regarding the pragmatic dimension of interoperability lead to the identification of practical use in data combination. Main questions focused on what kind of data could be combined in an interoperability framework, what categories of geospatial data and tables could be distinguished and what principles of statistical geospatial frameworks should be taken into account? Regarding this last question first results on initial requirements for data combination and the main definition for the structuring of core data are given.

The focus on possibilities to reduce, influence and steer side-effects guided the discussion along design thinking methodology, serendipity effects and main management methods of information technology, as these are applied in decentralized service-oriented architectures (the fundament of modern interoperability frameworks).

This report results in three highly recommended actions to be considered for further planning:

- Issue the observed occurrence levels of side-effects.
- Develop a practical guide for management methods in geospatial interoperability frameworks.
- Establish communication maturity evaluation for cross-domain interoperability framework stakeholders.

For all discussion items the group searched for examples and references. All findings led to the conclusion that decentralized interoperability frameworks are not mature enough at the moment in Europe and very rare examples could be found. Many examples are in development, but are not productive/operational.

This report describes the outcomes of Work Group B (WG B), elaborated by the subgroup for task 3 “Managing side-effects induced by data combinations” and the main results as they are relevant for the proposed actions and a successive derivation of actions. It is designed in a way that main questions on this topic are answered. The content follows the structure: background, data combination, side-effect, side-effect grouping, governing methodology and examples.

2 Introduction

This report describes the discussion of the Work Group B on task 3 “Managing side-effects induced by data combinations” and the main results as they are relevant for the proposed actions and a successive derivation of recommended actions. It is designed in a manner that the main recognised questions on this topic are answered and follow the structure: background, data combination, side-effect, side-effect grouping, governing methodology and examples.

For all discussion items the group tried to create a generic and common understandable view. For many reasons, like the members’ background and experiences, some aspects had to be reduced to geospatial and statistical examples.

Organisational setting and mission

The regional committee UN-GGIM: Europe focuses on two issues:

- increasing data interoperability and harmonization by proposing core geospatial data, and
- enabling the integration of geospatial data with other information/data (statistical, environmental, etc.) in order to foster further usage.

Issues related to the integration of geospatial data, including cadastral parcels, with other information are tackled by the UN-GGIM: Europe Work Group B on “Data Integration”.

This Work Group B report is the deliverable for task 3 “Managing side-effects induced by data combinations” (B3). The report is on a strategic, non-technical level reflecting the UN and the European goals.

Understanding a side-effect is quite difficult. The work group often came to situations when we discussed about a side-effect, but we meant something else and in the end the topic was not about the identification of a side-effect induced by data combination. In these terms the following definitions should help to understand the characteristics of a side-effect more clearly.

For example the accessibility effect comes up with access restrictions, because one or more of the data sources, which were used in a data combination, does not own appropriate licenses. Therefore the data combination becomes unusable.

The obvious occurrence of side-effects in data combinations leads to the main question: Could side-effects being managed, governed and used?

In order to govern and manage side-effects, one needs to know when side-effects occur, if steering processes to govern side-effects exist, if dependencies influence a side-effect behaviour or if prediction possibilities exist. In the end, management measures will be needed to evaluate side-effects in their impact and importance.

Introduction to the chapters of the report

The “background, acknowledgements and disclaimers” section (Section 3) explains the setting of subgroup B3 and the requirements given. It connects this report to the overall assignment of the UN-GGIM:Europe Work Group B.

The section “the essence of data combination” (Section 4) describes how modern map production frameworks work and may have an impact on the creation of side-effects. It explains the overall understanding of data combination and its critical action of linking information together.

The section “side-effects from data combinations” (Section 5) defines the notion “side-effect”. This section enhances the characteristic of side-effects. By establishing various occurrence levels for side-effects a more precise fundament for the evaluation of side-effects is given and links to interoperability frameworks.

The section “governing and managing side-effects” (Section 6) gives an insight in management methods to govern side-effects. It elaborates on the question if side-effects could be predicted or even provoked on purpose.

In the section “mining examples for side-effects” (Section 7) the most important observations for side-effect examples in interoperability frameworks are described. As the number of active interoperability frameworks is limited, but new frameworks are rapidly evolving, more diverse examples of side-effects induced by data combination should be observable in a short future.

The section on “actions” (Section 8) highlights the most important actions that should be investigated as next steps. The work group is convinced that the listed actions could lead to solid recommendations for governing side-effects as soon as its content and environment is explored and documented in depth.

Annex I lists more detailed descriptions of selected side-effects. This list is an excerpt. The list as well as its content should be completed in order to receive a comprehensive overview on influencing factors for data combinations. In the end, a complete inventory of side-effects cannot be obtained due to the nature of these effects; they can pop up instantly, due to the fact that side-effects occur unintended in the production environment by interacting circumstances.

Annex II suggests a common model for geospatial statistical management processes: the general statistical business process model.

Annex III lists all contributors of this report.

3 Background, acknowledgements and disclaimers

The United Nations initiative on Global Geospatial Information Management (UN-GGIM) aims at playing a leading role in setting the agenda for the development of global geospatial information and to promote its use to address key global challenges. It provides a forum to liaise and coordinate among Member States, and between Member States and international organizations.

The regional committee UN-GGIM: Europe was established on 1 October 2014. Its work plan mainly focuses on two issues: increasing data interoperability and harmonization by proposing core geospatial data and enabling the integration of geospatial data with other information/data (statistical, environmental, etc.) in order to foster further usage. One of the several Work Groups focuses on integration of geospatial data, including cadastral parcels, with other information.

Germany chairs Work Group B “Data Integration”. It is common understanding that Work Group B envisages a global vision with the focus on Europe for all tasks / deliverables. Strategic and political papers for “evidence based decision making” are needed rather than technical ones.

Work plan

Following its work plan, Work Group B will supply three deliverables for three main tasks 1-3:

1. Definition of the priority user needs for combinations of data
2. Recommendation for methods implementing the prioritised combinations of data
3. Recommendation about how to manage side-effects induced by data combinations

WG B decided to distribute the work to three subgroups B1, B2 and B3, one for each task.

Subgroup B3 started its activity in September 2015 and concludes its activity by submitting this report to the UN-GGIM: Europe Executive Committee.

According to the work plan task B3.1 focused on the identification of side-effects. What kind of side-effects may occur with data combinations? Can these side-effects be categorized and how? The identification and categorization of side-effects induced by data combinations will accordingly help to evaluate its impact and risk on geospatial information management. Task B3.2 followed the pragmatic dimension of interoperability on how side-effects influence interoperability and usability. All effort and governance are needed to enhance technical-, semantic- and policy interoperability. Task B3.3 looked at the reduction, manipulation and steering of side-effects. It will identify if there is a role for stewardship and if any other agreements are useful to govern side-effects on an international level. This task enhances governance methodology for dealing with side-effects in multi-source spatial data (data combinations) and applications.

More focus in B3 was given to statistical principles of the Statistical Geospatial Framework (SGF), e.g. ethic issues above data quality which are not obvious for many stakeholders in the geospatial domain. The impact of disaggregation (add information) was elaborated as well, e.g. if statistical principles can be kept in an open geospatial framework.

Work Group B collaboration

Subgroup B3 has held several teleconferences with the subgroup leader B2 and the acting WG B chair. The proposed outline and structure of the tasks were discussed and agreed. The document for deliverable B3 was compiled and edited by subgroup B3 within a collaborative platform.

Since its constitution, 17 countries comprising 20 organizations are committed to Work Group B. These organizations are either mapping or statistical ones. There is no representation from other thematic domains yet. Unfortunately, there are only weak responses or commitments and contributions from the Baltic, Balkan and Eastern European countries.

Work Group B has taken into account the global recommendations from UN-GGIM such as of the UN Expert Group on the Integration of Statistical and Geospatial Information (UN-ISGI) as well as of other relevant UN and global initiatives. Work Group B considered the requirements of National Statistical Institutions (NSIs) for the integration of statistical and geospatial information.

Considerations of Work Group A

Work Group A on "Core Data" aims at identifying essential data for sustainable development i.e. the core data needed by UN, European and national activities related to sustainable development, in order to get political and financial support to fulfil this need. The coordination between WG A and B is deemed crucial and has been followed regularly. The topics remained at a conceptual level. For example the pre-selected core data list and the scoping paper for core data (provided in mid-April 2016) has been completely considered. In terms of SDG indicator requirements, data combination between statistical and geospatial content has been observed to gain further importance. This is due to SDG indicator characteristics that will call for a more extensive data combination.

Generally, the purpose to share the work between as many participants as possible is important in order to ensure both that individual contributions accommodate participant's availability, and that the collective work, benefiting from as many contributors as possible, will be able to deliver relevant outcome and findings.

A full list of those who have contributed can be found at the end of this report. We are grateful to every person and organization for giving their time; either to provide written contributions or to attend the meetings and telephone conferences for allowing us to include their contributions in this report.

The UN-GGIM: Europe Executive Committee approved this report in **October 2016** and uploaded it to the UN-GGIM website.

4 The essence of data combination

Data combinations are crucial actions in spatial data preparation, for spatial data analysis and visualization. Data combinations can enhance the value of single datasets due to information densification, adaptation to defined use cases and responsiveness to specific questions. The influence of information quality and occurring side-effects on a further usage and geo-communication, call for specific management procedures and may need restrictions.

4.1 What should we understand by data combination?

"Data combination" suggests a merging, processing or creation of datasets. It does not generally restrict sources, characteristics, qualities, extension or specification. Therefore, data can be combined from various sources, with different characteristics, qualities or just varying extension. This means that even an aggregation of a specific data theme with itself at different precision levels delivers a combined dataset. For example, a combined point cloud resulting from different image matching procedures and airborne laser scanning will cover points at different qualities. Of course these aggregated data could be used for further combinations with statistical sources or any other available and linkable information (nested data combinations).

4.2 What kind of data could we combine?

In general, the combination of (geospatial) data is not restricted. The data variety encompasses lists and tables as well as simple features and complex thematic geospatial objects. Data combination perspectives exceed simple data table joining. It covers thematic as well as regional extensions of data. For example a mosaicking process extends the regional extent of a national dataset in order to make a supranational one. In general all source data will consist of more or less the same schema/structure. A further example elaborates on the geocoding process: any administrative, business, etc. dataset (table) can be transformed to geospatial information by the use of geospatial references or core geographies. As soon as the link (persistent identifier) between the table content and the geospatial anchor exists, geospatial information can be created.

4.3 How to combine data?

The most important requirement is the existence of unambiguous "linking keys" (persistent identifiers) between datasets and objects in order to make a join possible. In many cases thematic name conventions or codes, e.g. the ISO abbreviations for country names, deliver a usable linking key as long as it is persistent during time and semantic.

The ''linked data'' approach adds a further perspective to data combinations. It directly uses geospatial features and objects for the data combination, like linking text and images on the Internet. One main requirement for all linking data is the availability of consistent identifiers and their technical and semantic persistency. If identifiers get lost, data cannot be referred to anymore and will be lost for the use case of data combination. Also if the meaning for data change, their reference may be wrong or misleading.

Beyond the thematic linking key, a location anchor could be used as main link between features and objects. As soon as information is bound to space, this space could be used to link embedded data, even if their ontology, meaning or quality do not match. It is obvious that a thoughtless data combination could bring up false interpretations, quality mismatchings and even ethical issues within a further processing.

4.4 Are there specific data classes to be combined?

As the combination of data is not restricted, it may be useful to structure (geospatial) data sources in order to receive guidance for the quality and dimension of occurring side-effects, data and service quality estimations or data combination restrictions. Possible occurring side-effects concern an a priori classification of side-effects and identification of their relations. The guidance for data and service quality estimations focus on expected quality restrictions for data, e.g. the best expected quality after a data combination or keeping confidentiality of the information used, and services, e.g. changing response times for a service after or during the combination of different data classes.

The data classes could differ between three different states. They specify basic classes, called core reference geography, which is used as spatial reference, a core thematic geography that comes along with geometry and theme specific content, and spatial related information which enriches spatial features and objects.

- Core reference geographies form common spatial information frameworks that are defined and used by different stakeholders and in various location strategies. These geospatial infrastructure data are used to perform geospatially enabled statistics. Examples are point clouds consisting of georeference and ID (e.g. surface points, address points,…), trace collections or polygons (e.g. geographical grids, statistical units, etc.). The main link attribute and persistent identifier could be formed by a code generated from the spatial reference.
    - Georeferenced address- and building registers,
    - Cadastral Parcels,
    - Administrative units,
    - Statistical and grid geographies,
    - Topographic geometries.

- Core thematic geographies cover theme specific features. They consist of geometries including theme specific attributes. For example administrative units will have to embed the hierarchy level, codes or geographical naming of administrative polygons. These core thematic geographies are used
    - to enhance (i.e. add variables to) existing statistical and administrative datasets (e.g. distance to green space: road network, parklands, etc.) or
    - to elaborate production of new statistical content (e.g. land accounts: cadastral information - land use and value, gridded land cover from Earth Observation, etc.).

- Spatially related information are documents, tables or other features that consist of a relation to space or a relation to core geographies in order to enrich the combined dataset, enhance completeness and therefore support broader investigations. Therefore these geospatially enabled data or statistics describe statistical or administrative data that could be easily linked to a geospatial object.
    - on a fine level: unit record data linked to a coordinate, small linear object or small area geography building block

- on a mid level: unit record data linked to a large geographic unit or aggregate data linked to a small area geography building block
- on a course level: aggregated data linked to any medium or large geographic unit.

All three data distinctions are covered by a basic core data definition. The definition of "core data" has been part of several reports of UN-GGIM: Europe WG A. For example, it is included in the report "Contribution of regional committees and thematic groups to the global geospatial information agenda" (E/C.20/2016/3/Add.1) for UN-GGIM-6 (August 2016). In that report it is stated:

"…In terms of a concept, core data can be seen as the authoritative, harmonised and homogeneous framework data which both national and international users need to either fulfil their requirements or to geo-reference and locate their own thematic geospatial data. Besides, core data should follow a bottom-up approach from authoritative data of member states..."

4.5 Do accessibility specifics exist for data combinations?

In general, data combinations form a main source for further analysis or dissemination. Therefore their accessibility has to fulfil the same requirements as for the source datasets. Effectively the product of data combinations comes along with some specifics, which concern licensing and the combination on the bases of linked data.

The licensing issue for data combinations concerns the mixture and possible variety of licenses from the source datasets. A resulting license for the data combination product could be more demanding than the most demanding license of the source datasets, but it must not be less demanding because of license violation against the data sources. Furthermore it could be difficult to fulfil all the requirements of data source licenses within the data combination.

The data combination on the basis of linked data may perform real-time data combination in terms of attribute extension or regional extension or complement. The linked data approach is one specific way to access data and datasets. Appropriate and common accepted business models are required to establish the linked data combination. Commonly accepted means that all stakeholders for the linked data approach support its specific data access. The main concepts of open data could be understood as one specific business model. As soon as the stakeholders establish the open data framework business model, they could also implement the linked data approach. On the other hand, the business model "open data" is not an essential requirement for linked data. Also other business models could support linked data, e.g. with the embedding of authentication and authorisation mechanisms. Examples for non-open-data approaches can be found for the product portfolio of EuroGeographics, which defines a purpose bound license.

Example: Challenges with data sharing and data quality of the consolidated address data set for Germany.

In Germany, the responsibility for topographic and cadastral reference data is assigned to the federal states. The official German address dataset is compiled from contributions by the mapping authorities of the 16 states. The mapping authorities extract the information from their cadastral repositories. Thus, the quality of the address data depends on the completeness and up-to-dateness of the cadastre. The address data does not always reflect the real situation as new buildings are not introduced to the cadastre until construction is completed and the footprint has been measured by licensed surveyors. In some instances the official survey gets delayed for years. The mapping authorities have recognized the problem and discuss on solutions.

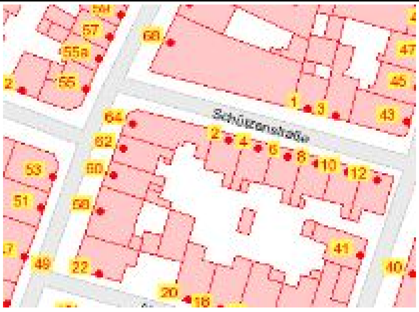


Figure 1 – Visualized address data on cadastral map.
If data is not up-to-date, this may cause inconsistencies (e.g. different time stamps)

The Federal Agency for Cartography and Geodesy (BKG) matches the address dataset of the federal states with address data from a commercial source that has a better up-to-dateness, and combines both sources into a consolidated dataset. The consolidated dataset is provided to users within the federal government, including the Federal Statistical Office of Germany. However, the license agreement between BKG and the provider of the commercial address data does not allow BKG to forward the consolidated dataset to users outside the federal government. For instance, the European Commission and Statistical Agencies of neighbour countries are not allowed to use the consolidated dataset. A mandate for cross-border applications would support BKG in funding the additional license fees for extended usage. Additionally, neither BKG nor the Federal Statistical Office are allowed to forward the consolidated dataset to the Statistical Offices of the German federal states in order to establish uniform national reference data at all levels of administration. Yet the Federal Government is negotiating with the states on this problem.

The Federal Statistical Office of Germany used the consolidated dataset for georeferencing with the Population-Census 2011. In the process the address data was further improved by matching it with the information from official registers. This resulted in a dataset of very high quality with respect to consistency, completeness and up-to-datedness, the register of addresses and buildings that was used for the organization and coordination of the Census 2011. But due to the German confidentiality laws the Statistical offices were not allowed to feed the improved data back into the datasets of the Mapping Authorities. Moreover the data must be deleted 6 years after the census at latest. Therefore all efforts have to be repeated with the next census. However, at the moment a new law is in preparation in order to keep address-data permanently for statistical purposes.

Analysing the given example it can be recognised that side-effects were not considered before the production start. Therefore any occurring negative effect had to be reduced in post-processing procedures.

4.6 Are there main principles for the combination of geospatial and statistical information - a statistical geospatial framework?

The development of a Global Statistical Geospatial Framework, which is proposed by the United Nations Expert Group on the Integration of Statistical and Geospatial Information (UN EG-ISGI), consists of five main principles that are considered for integrating geospatial and statistical information. These five principles are:

- Principle 1 - Use of fundamental geospatial infrastructure and geocoding
- Principle 2 - Geocoded unit record data in a data management environment
- Principle 3 - Common geographies for the dissemination of statistics
- Principle 4 - Interoperable data and metadata standards
- Principle 5 - Accessible and usable geospatially enabled statistics.

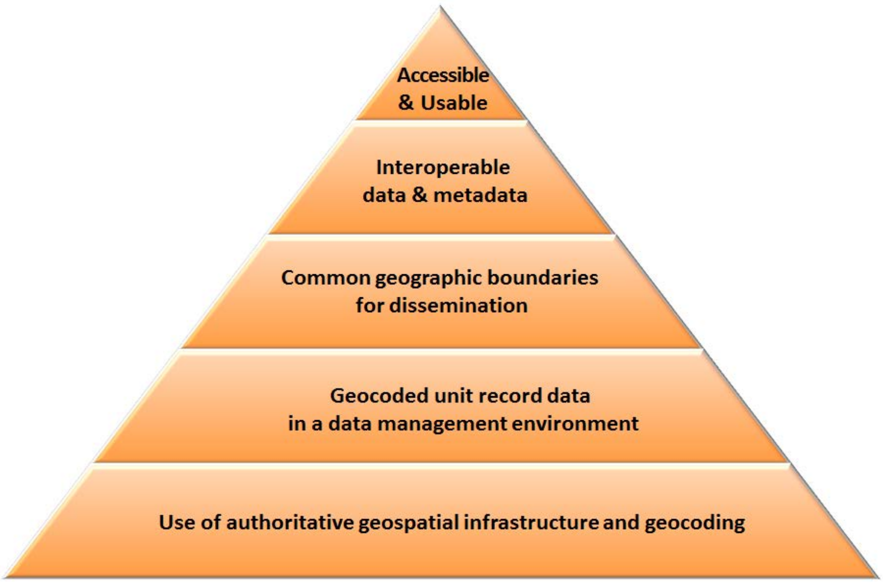


Figure 2 - the 5 principles of the Statistical Geospatial Framework (SGF)

The Global Statistical Geospatial Framework is envisaged as a high-level framework which shall facilitate consistent production and integration approaches for geo-statistical information. It is generic and permits application of the framework principles to the local circumstance of individual countries[1].

4.7 Is there a difference of geospatial management in modern production environments? (monolithic/central production processes versus decentralised production environments)

Main characteristics for well-functioning/successful production environments are reliability and productivity, which means that involved procedures run from the beginning to the end in a controlled and quality assured environment. The risk for side-effects are diminished. Therefore every outcome is controlled by one single producer/process. This control is being embodied in monolithic/central processes, which can be characterised by a one step by step approach. Decentralised geospatial infrastructures can be characterised as being up to date, agile and run by distributed data producers. Infrastructure content (geospatial data and services) is provided at the most actual state. One drawback is that changes of infrastructure components, as it is done for system maintenance, may result in a loss of sources if main interfaces or points of contact to the data or services change.

A comparison between centralised and decentralized production environment shows that monolithic production environments generally consider business as well as technical side-effects as part of the production process. Unplanned side-effects may occur in the first stages of the decentralized usage. Side-effects in production environments arise especially when new (business) processes are established and/or new sources are embedded or modified for a data combination.

[1] http://www.efgs.info/wp-content/uploads/informationbase/framework/BG-2016-31-proposal-for-a-global-statistical-geospatial-framework-E.pdf, visited 2016-05-12

5 Side-effects from data combinations

The following chapter provides a definition for "side-effect" and clarifications about how Work Group B "Data Integration" has tackled the issues associated with it.

5.1 What could we understand by side-effect? (Definition)

As short definition for side-effect that is used for the coming descriptions, we could define "*side-effect as something that occurs unintended with the data combination and needs further effort to be exploited, removed or steered."*

From another viewpoint and in other words, ''side-effect'' is an effect that is secondary to the one intended; it can also occur to be beneficial, although unintended. This means that in order to recognize the side-effect the whole production process from the beginning to the final product has to be reflected. There is the need to describe and control every step in the production process and product dissemination. If unintended outcomes are observed, their benefit or drawback has to be clarified. Repetition is a key aspect at that stage. If beneficial outcomes can be reproduced, then they can be exploited. If drawbacks can be retraced, then they could be avoided. The outcomes or drawbacks should be considered in the production management.

Side-effects do not only occur during the production, but also for the dissemination and usage stage. This may lead to two different viewpoints: what has to be done to manage and predict side-effects in the production process? What has to be done to inform and train users about data combination and their side-effects? In general, a data combination product designed for specific activities could also have a downside. For instance, a comprehensive data combination that supports the management of refugee camps could also be misused.

5.2 Is a side-effect something good or bad?

A side-effect is neither good nor bad. Positive as well as negative effects may occur in central as well as distributed production processes. It should be the aim of the management to maximize the positive effects and minimize the negative ones.

Example: Semantic interoperability in INSPIRE is achieved by means of UML models, that are supported by code examples, thesauri, codelists, registers and semantic descriptions.

Positive side-effect: Data validation and consistency checks can be automated because the UML models are machine readable.

Negative side-effect: Human experts with no or rare knowledge of UML have difficulties in understanding the data content and the data requirements of INSPIRE. Therefore additional activities to support the semantic understanding are needed.

5.3 Is a side-effect predictable?

In general a side-effect occurs by accident. If an effect is repeating itself, it will not be a side-effect anymore. Section 6 elaborates more on the management of side-effects.

Based on the working group appointment "managing side-effects induced by data combinations" we noticed, that an observed effect will stay a side-effect as long as it will be used for production purpose and therefore the weight and impact on the product or dissemination are known. As soon as a side-effect impact is knowingly employed it becomes a designated part of the production lane or product definition and is therefore manageable.

Side-effects could be predicted by statistical evaluations or technical assessments. Forecasting can help to derive quality effects that are induced by various and inhomogeneous data source qualities. In the area of risk management those evaluation and assessment tools are used to predict risk zones and occurrences.

5.4 What kind of side-effects exist?

Side-effects have different causes. The attached list is not exhaustive (Annex I: detailed side-effect descriptions) and new side-effects could be added. The main problem with the consideration of side-effects is their weighting: not all effects do have the same impact on availability, accessibility and acceptance. Therefore this report demonstrates possible classifications and tries to derive an important weighting by using different levels of side-effect occurrences.

Based on the aspect of serendipity, which brings up new unplanned outcomes, the list of side-effects cannot be completed. From a very pragmatic point of view side-effects will originate, have impact and depend on various topics. These topics are technical frameworks, data and service quality, thematic geospatial competency, economic structures, legal aspects, organizational models, interoperability densification and data source capabilities.

Technical frameworks deal with standardization and the technical infrastructure. Data and service quality defines quality and service level agreements. Thematic geospatial competency makes the importance of literacy and education in terms of geospatial data creation and usage clear. Economic structures consider mechanisms of business models and their requirements. Legal aspects concern legal frameworks that have to be respected, modified or created. Organisational models describe structures that support or prevent collaboration. Interoperability densification focuses on procedures to enable and enhance interoperability. Data source capabilities try to formulate the provenance/lineage of data and services, which clarify their possible usage, restrictions and quality expectations.

5.5 Can differences between side-effects be observed?

In general, side-effects occur unintended when different datasets are combined. Therefore the occurrence of side-effects could be divided in primary-level occurrence, secondary-level occurrence and possibly even a third-level occurrence.

A *primary-level occurrence* of side-effects directly impacts the technical or semantic dimension of data combination. This means that the side-effect is directly assigned to the data or services combined, thus the resulting product.

A *secondary-level occurrence* of a side-effect concerns organizational or legal areas that are not obsessively related to the resulting product of data combination (but is invoked by it). For example organizational actions in terms of stewardship programs or the establishment of legislation are derived from the observation of secondary-level side-effects.

A *third-level occurrence* of side-effects is related to observations in societal areas. The considering and keeping of ethical restrictions, their widening and its discussion are based on observations of third-level occurrences. For example, a minimum amount of education is needed to understand the characteristics of interoperability frameworks. Appropriate educational programs are main requirements for the success of geospatial interoperability implementation. If this educational activities are missing, then we could observe third-level side-effects as consequence.

All three occurrence levels on their own may have different impact on interoperability and the pragmatic dimension of data combinations. For instance, it is useful to have geospatial data for refugee routes and camps available for assistance and care. At the same time side-effects for misuse may occur and result in security issues for the refugee stakeholders. From this simple example we observe the importance of disclosure and its legal determination. It is less a primary-level occurrence of side-effects (technical) than a secondary- and third-level occurrence of side-effect with enormous impact if the misuse of the data prevents actors from providing assistance for refugees.

6 Governing and managing side-effects

The evaluation for governance of side-effects in their three occurrencies leads to the fundamental idea that ethical issues, legal frameworks and confidentiality are the most important issues that will guide management activities. In the statistical domain these issues are embedded in the statistical principles[2], which may be less known in the geospatial domain or others. Keeping these principles prevents from ethical, legal and privacy problems. In this viewpoint the impact of mathematical set theory, disaggregation and statistical recalculation has to be further investigated for open interoperable frameworks.

6.1 When and how do side-effects occur?

As long as the weight and impact on the production and dissemination are not known, the side-effect stays an unintended effect that is not considered, steered or exploited within the interoperability framework or production procedures.

6.2 What kind of processes can be established to govern side-effects?

The governance of side-effects bases on continuous processes that evaluate, compare and measure data combinations and their identified (thus known) side-effects. Actual methodology follows resource as well as flow management. Depending on the maturity of the interoperability framework, its occurring side-effects and the maintenance procedures in place, appropriate management tools have to be applied. For a large extent these processes come from controlling (balanced scorecards), quality management, change management, supply chain and even LEAN management[3].

For example, quality management ensures that a side-effect is consistent perceivable and therefore becomes usable or preventable. It consists of four main stages and the achieving of those:

- quality planning
- quality assurance
- quality control and

[2] Statistical principles such as the UN Fundamental Principles of National Official Statistics http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx or the European Statistics Code of Practice http://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice

[3] LEAN management is an approach to running an organization that supports the concept of continuous improvement, a long-term approach to work that systematically seeks to achieve small, incremental changes in processes in order to improve efficiency and quality.

- quality improvement[4]

Achieving more consistent quality for side-effects allows for evaluating their continuous impact, observing changes of the side-effect impact and identifying new or similar side-effects.

Primary level occurrence of side-effects with their impact on the technical and semantic dimension of data combination mainly concern resource based management methods, which allow for resource evaluation, modification and creation.

Secondary level occurrence of side-effects with their impact on organisational and legal areas take effect on process based management methods, which steer the modification and creation of processes. Within the statistical community the GSBPM (generic statistical business process model)[5] should be used by producers of official statistics, at both the national and the international levels, to describe and assess the quality of processes based on surveys, censuses, administrative records and other non-statistical or mixed sources. It is a generic business process model that covers all activities in the statistical production and dissemination.

6.3 Are there possibilities to predict side-effects?

As a side-effect is something unintended any prediction seems to be impossible. In fact specific methodologies, like design thinking, exist that provoke new combinations, solutions and observations with a mixture of divergent and convergent thinking[6].

Design thinking is a formal method for practical, creative resolution of problems and creation of solutions, with the intent of an improved future result. Design thinking employs divergent thinking as a way to ensure that many possible solutions are explored in the first instance, and then convergent thinking as a way to narrow these down to a final solution. Divergent thinking is the ability to offer different, unique or variant ideas adherent to one theme while convergent thinking is the ability to find the "correct" solution to the given problem[7].
In case of side-effects based on data combinations the given method of design thinking could provoke any identification of new side-effects.

6.4 How to make use of side-effects?

Knowing about an overall reaction (side-effect) of data combination allows for impact planning and exploiting preparation. Being conscious about side-effects helps to establish steering activities - avoiding negative ones and enhancing positive ones. Generally, diverse management methods help to find peculiarities and define appropriate actions for specific occurring effects. Some examples are given in section 7.

---

[4] Rose, Kenneth H. (July 2005). Project Quality Management: Why, What and How. Fort Lauderdale, Florida: J. Ross Publishing. p. 41. ISBN 1-932159-48-7.

[5] UNECEpaperonapplyingGSBPMtoregistermaintenance, http://www1.unece.org/stat/platform/download/attachments/57835551/BR%20meeting%20paper2.doc?version=1&modificationDate=1312874708483&api=v2, visited 2016-05-27

[6] Keeley, Larry, Helen Walter, Ryan Pikkel, and Brian Quinn. Ten Types of Innovation: The Discipline of Building Breakthroughs. Hoboken, NJ: Wiley, 2013.

[7] Plattner, Hasso et al. Design Thinking: Understand, Improve, Apply. Berlin; Heidelberg: Springer, 2010.

The proposed value-adding chain of decentralized production networks bases on occurring side-effects. The main argument is that most reproducible effects, that were not planned beforehand, could be used as positive input and therefore will add value. This value could be added to the existing production lane, the product itself or evolves to an independent product. The unplanned occurrence of new outcomes is entitled serendipity effect. Serendipity effects are provoked by design thinking methodologies.

6.5 Are there existing examples for managing and using side-effects?

Generally, existing examples for managing side-effects seem to be rarely described. Either an extensive engagement is missing or side-effects are exploited or avoided in established production lanes due to their resource and flow management.

There exists one extensive action on the management of side-effects concerning confidentiality of census information. It is an activity of the European Statistical System. One main result is the handbook on statistical disclosure control, which means that side-effects concerning privacy, or confidentiality in general, are governed in a way that the ethical and legal obligation to prevent sensitive information from dissemination are fully respected[8]. The topic of confidentiality is applicable to digital information in general and not only to census data.

One side-effect that may occur with data combination is called disclosure. It occurs when the final product (data combination) reveals knowledge about another person or organisation that was not observable before and concerns identity or attributes.

“..*Identity disclosure occurs with the association of a respondent´s identity with a disseminated data record that contains confidential information. Attribute disclosure occurs with the association of either an attribute value in the disseminated data or an estimated attribute value based on the disseminated data with the respondent.*”[9]

7 Finding examples for side-effects

This section should focus on how side-effects are found. How obvious are side-effects and their impact? Therefore the question arises if the work group observed side-effects for existing geospatial frameworks in its surveys? Derived from these observations a generic description for side-effects could be defined.

7.1 What kind of side-effects could be prominently observed in existing statistical geospatial frameworks?

From a pragmatic point of view (the daily work of statistician and geoinformatics) some generic and prominent examples could be derived, which cover the variety of data providers, matching reference maps, regional statistical units, cross-border grid based mapping, confidentiality issues, quality dimensions and even missing interoperability.

---

[8] ESSNet SDC (2010) Handbook on Statistical Disclosure Control, https://ec.europa.eu/eurostat/cros/sites/crosportal/files/SDC_Handbook.pdf

[9] Duncan, et al (2001) Disclosure Risk vs. Data Utility: the R-U Confidentiality Map, Technical Report LA-UR-01-6428, Statistical Sciences Group, Los Alamos, N.M.: Los Alamos National Laboratory.

The **“variety of data providers”** is one main characteristic for distributed interoperability frameworks. Stakeholders/data providers span a lot of domains which follow their own requirements. If data combinations sourced by cross-domain interoperability frameworks should be successful, their stakeholders need to be covered by common requirements, organizational agreements and/or a shared legislation. The missing of these factors leads to difficulties/side-effects that will call for specific solving actions. All three occurrence levels are affected by an unregulated variety of data providers.

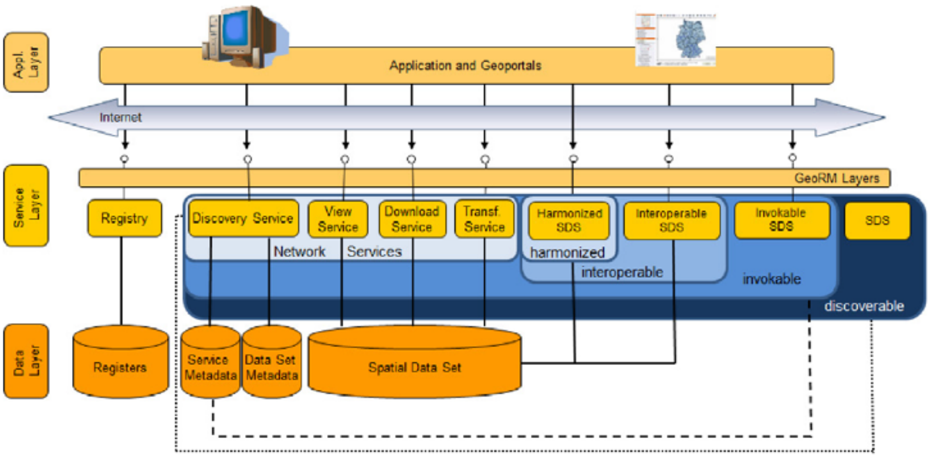


Figure 3 – Updated INSPIRE architecture in INSPIRE Spatial Data Services and services allowing spatial data services to be invoked – Draft implementing rules (01/08/2012)

Matching reference maps/data are crucial for data combination (see also section 4.3). Matching concerns semantic usability. For example the administrative structure may change due to administrative reforms during time and the size, name and aggregation of administrative units will change. Matching thematic data (tables) from a given timestamp to the administrative units will call for exactly the same lifetime in order to enable a correct matching. Historical states and the tracking of changes in reference datasets are crucial for data combinations. Ignoring this situation leads to negative side-effects with wrong or none-matching results. One very basic example of a first level side-effect occurrence is the reference-time-mismatching of data sources and reference geographies that will result in empty data, because administrative unit aggregations change or towns change due to administrative reforms.

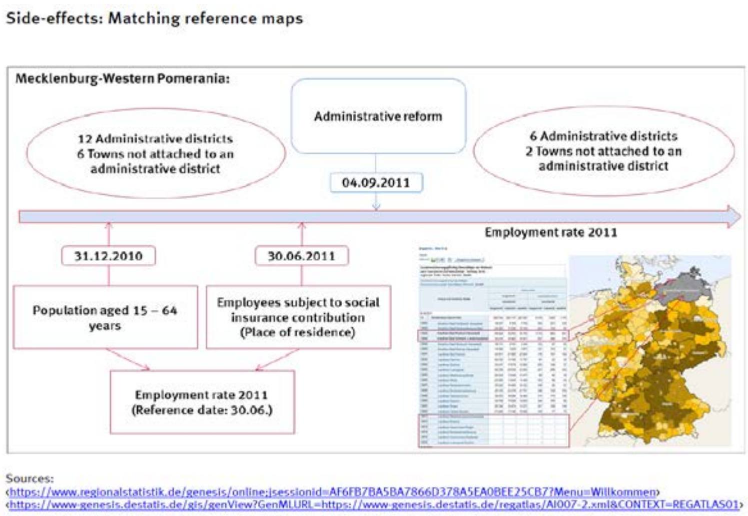
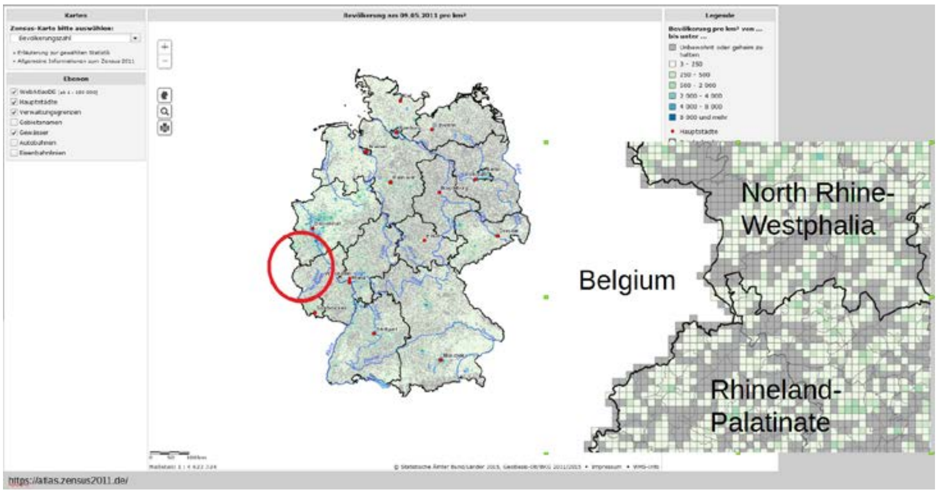


Figure 4 - Matching reference maps after administrative reform in Germany

Regional statistical units are the fundament for statistical geospatial data combinations. The availability of these **core reference geographies**, which are used as main reference to space, is a requirement for the data combinations and its use cases. These core reference geographies vary in their relevance, accessibility, licensing and cost, which may result in different side-effects if the technical, organisational, legal and semantic interfaces are not defined.

Cross-border grid based mapping comes up with several side-effects concerning a grid based data exchange, the standardization of addresses, a uniform address allocation (e.g. no addresses for churches or farms) and the tracking of address changes in municipalities. Whereas the grid based data evaluation aims at comparability and semantic independence through time, the mentioned side-effects cause discrepancies in cross-border grid cells and result in errors.



Source: https://atlas.de/

Confidentiality issues occur with increasing granularity and the amount of combined quantities. Following the statistical principles at least three units have to be aggregated and any modifications have to be controlled. Different geometries and grid sizes reduce the risk of disclosure. In addition, large size-classes for statistical data should be preferably defined.

7.2 Are there "more important" side-effects?

The importance or weight of a side-effect is definitely not the classification to first-level-, secondary-level- or third-level-occurrence. In all three levels very important side-effects occur. The previous given example of disclosure risk describes a side-effect that has secondary-level- and third-level-occurrence.

8 Actions

The following actions are proposed to the UN-GGIM: Europe Executive Committee to be considered for future planning. The actions are on a high level. Concrete activities on research and development were out of scope for the WG B3.

8.1 Action 1: Issue the observed occurrence levels

The Report B3 describes side-effects at a general and theoretical level. The Executive Committee is invited to consider further elaboration of side-effect examples in their different occurrence levels from current European interoperability frameworks. Such frameworks are evaluated in Report B2 which is compiled in parallel with report B3.

Action: Elaborate examples for side-effects taking into account one or more of the interoperability frameworks detailed in report B2

8.2 Action 2: Management methods for geospatial interoperability frameworks

Proper management tools and procedures are crucial for reducing the negative impacts of side-effects.

Action: The existing interoperability frameworks shall be analysed for their management procedures. The analysis results in a practical guide of good practice.

8.3 Action 3: Communication maturity for cross-domain interoperability framework stakeholders

Stakeholders from the Mapping domain and the Statistical domain have a different understanding of terminology and its organisational, legal and technical framework. Therefore, particular attention should be paid to well-accepted evaluation scores in cross domain frameworks. Some methodology for using a GIS competency model and maturity scores is described by URISA´s GIS management institute (http://www.urisa.org/main/gis-management-institute/).

Action: Implement cross-domain maturity checks at primary milestones of the statistical geospatial framework. But also focus on bringing worlds together. Communication and understanding is key!

Annex I: Detailed side-effect description

The following list of side-effects is incomplete, heterogeneous and conceptual. It gives only an impression on the discussion and its level of detail within subgroup B3.

1. Accessibility Effects

Each source dataset in a compound dataset consists of characteristics and errors. Whereas for the data sources accessible metadata exist, the compound dataset may miss documentation of lineage, errors, quality and composition. There exists the risk to propagate errors. In order to avoid additional errors, quality descriptions and metadata for compound datasets are needed.

The combined result may consist of restricted accessibility, if one or more of the sources does not own appropriate licenses.

Production procedures that rely on specific sources may suffer from inaccessibility of single data sources, if their access point or license changes.

2. Aggregation Effects

Aggregation of data generally aims at completeness gain for datasets to support selected tasks. A dataset will be of high quality if all planned actions can be fulfilled satisfactorily. Therefore aggregation and building of combined datasets are core actions in data preparation. They are needed for any initial procedure that extends well-established workflows and analysis.

Aggregation effects concern consistent data completeness as well as transparency in data quality for aggregated data and their lineage/provenance.

Consistent data completeness faces the extension of geometries and attributes for planned actions, at which the aggregation procedure is done consistently and successfully. For example the join of attribute tables uses persistent identifiers that will semantically and syntactically match and deliver valid results.

Transparency in data quality for aggregated datasets should follow guidelines for data quality description that clearly documents origin, aggregation and processing of datasets. This transparency and documentation is needed for further combinations of aggregated/combined datasets.

3. Analysis Effects

Analysis itself may not have effect on a combined dataset, but the requirements, which an analysis presupposes, influence the combination of datasets. The requirements of analysis actions may call for specific data schema, precision, resolution or even formatting.

On the other hand data combination done on purpose for specific analysis enhances a given dataset with additional information, structure or resolution.

(also see aggregation effects)

4. Availability Effects

Data combinations may have an impact on availability of datasets due to license aggregation or missing service-level agreements. If a given license of one source product excludes third party use or embeds strict regulations for any further use, then the combined dataset will generally include these regulations.

In case of real-time combinations, as they are produced by service orchestration, the combined dataset will depend on a common service level agreement (SLA): what is the availability, capability and performance of the service? A missing SLA may lead to missing data sources and result in non-availability of combined datasets.

5. Business (Model) Effects

Business models may be triggered by combined datasets: any enhancement with additional information will lead to more complicated business processes, communication/stewardship structures and license combinations.

The enhanced usage perspective of combined datasets may lead to additional business models and new communication structures. Additional created business models (due to combined datasets) are often described in the "value adding" characteristic of Service-Oriented architectures.

6. Communication Effects

In terms of reliability, security and actuality, combined datasets with their references to data origins (data providers) may call for communication efforts to clarify security issues, organisational issues or further usage maturity (if not clearly/precisely stated in the metadata).

7. Data Security Effects

Data combinations may come up with specific security issues if e.g. the thematic dimension introduces more detail that allowed for a georeferenced dataset.

Enabled data security arrangements for single datasets could be overruled by orchestration.

Therefore data security tasks for single datasets have to be adopted and aggregated in combined datasets.

It has to be stated that misuse of data can hardly be prevented if data are accessible and publicly open for specific tasks. Removing anonymous access to the datasets could have a positive effect. An urgent example is the need of information on the European refugee stream for volunteered support activities. This information could also be negatively used by extremist groups. The unsolved question of this data security effect is one of the reasons for the missing of precise and official information on the refugee stream.

8. Deciding Effects

The combined dataset and its communication preparation takes influence on decisions to be made. Missing or wrongly defined data or combinations of data will effect deciding processes, which then could be made on false premises.

In many cases this effect occurs for semantic heterogeneity, especially in cross-border situations. Cross-border situations describe geographical and/or thematical neighbouring areas/domains. For example at national borders or neighbouring domains like geological survey and land use (soil).

9. Economic Value Effects

Economic value could increase with data combinations, their conditioning and contextual widening.

In general, economic value is a measure of the benefit provided by a dataset or -service to a client (stakeholder, user, …).

Note that economic value is not the same as market price, nor is economic value the same thing as market value. If a consumer is willing to buy a good, it implies that the customer places a higher value on the good than the market price. The difference between the value to the consumer and the market price is called "consumer surplus". It is easy to see situations where the actual value is considerably larger than the market price: purchase of conditioned OpenStreetMap data is one example [Steve Keen (2001) Debunking Economics, New York, Zed Books, ISBN 1-86403-070-4].

10. Education Effects

Combining and conditioning datasets could enhance educational purposes/application fields. The main assumption is that the meaning, quality and cross-relations are transparent and understandable for a cross-domain user group.

11. Governance Effects

Combined data and orchestrated services could be part of a decentralized geoinformation architecture, which allows to collect, combine and use its components in a stable way. All of these components undergo their periodical update processes and changes. In order to keep the architectures functionality (availability, accessibility and use) a overall governance (coordination) process is needed in order to take the right "development" decisions.

Governance of a decentralized geoinformation architecture (as basis for data combinations) is supported by stewardship conventions which ensure that data and services exist for the agreed time period.

12. Homogenisation Effects

The process of data combination itself could also cover data homogenisation efforts, although homogenisation is triggered by specified use cases and broad agreements.

If the point in time for data interface modifications due to data combination overlap with homogenisation requirements, then additional benefit and reduced effort could be generated.

13. Infrastructure Effects

An IT infrastructure is designed for specific use cases, applications, organisational (business) models which allow to estimate hit rates, any requirements for quality of services and the user/application needs.

If data combination is done as real-time service, then specific impact on the IT infrastructure and -architecture occurs that has to be considered. If not, these services and their datasets might be unavailable and inaccessible until the IT infrastructure is corrected.

Appropriate organisational and legal conventions are needed to allow "on-the-fly" combinations or the dissemination of combined datasets coming from decentralized sources via spatial data services.

14. Judicial Effects

The effects occurring by data combinations for legal issues are various and span from data security to licensing. Judicial effects can be seen as upper-level notion from a more general point of view: a combination of datasets also means a "combination" of legal frameworks. Therefore agreements are needed to respect and follow all involved legal frameworks.

One pragmatic approach to establish legal reliability are stewardship programs which also cover judicial effects.

15. Lean Effects

Data combinations could trigger constant changes in terms of modifying data- and service interfaces, application schemes or implementing new standards. There are various examples, for instance the MIG (maintenance and implementing group) of INSPIRE continuously deals with changes of the INSPIRE infrastructure. If this change process is accepted, then traditional production lanes can be re-evaluated and undergo process redesign. This redesign procedure should follow the idea of lean management (making processes more effective).

16. License Effects

License effects are one specific occurrence within judicial effects. Data combinations may use different data sources and therefore different licenses. In order to come up with one license for the data combination, a license homogenisation is needed, which is a common agreement between all stakeholders.

License effects become observable when the licenses for single data sources are more "demanding" than the others ("demanding licenses"). For example "Creative Commons - Share Alike" licenses may distribute derivative works only under a license identical ("not more restrictive") to the license that governs the original work. Without share-alike derivative works might be sub-licensed with compatible but more restrictive license clauses, e.g. CC BY to CC BY-NC ["Baseline Rights". Creative Commons. June 12, 2008. Retrieved February 22, 2010].

17. Linked Data Effects

Data combinations suffer from "linked data effects" when the quality, structure, timestamp, etc of sources does not match. This means that provenance/lineage of the different data sources is not compatible. Neither the data nor the metadata will prevent from combining incompatible sources. Linked data effects can only be identified if this incompatibility is observable in the resulting data combination.

Data combinations on the basis of "linked data" presuppose persistent, reliable and common agreed identifiers for a consistent "join". Wrong joins destroy a resulting combination. In addition these "wrong joins" are generally hard to observe.

When users are presented with geospatial information that has been integrated from different sources they need to understand the provenance and metadata in order to trust it. Trust is a term with many definitions and uses, but in many cases it can be seen as a judgment that a user makes based on the context of information they see. Provenance refers to the sources of information, such as entities and processes, involved in producing or delivering a product. The provenance of information is crucial to making determinations about whether information is trusted, how to integrate diverse information sources, and how to give credit to originators when reusing information. To simplify we can define provenance as a grouping of all the information the user is interested to know about who/what/when/how/why the information was generated.

Descriptive metadata only becomes part of provenance when one also specifies its relationship to deriving an object. For example, a file can have a metadata property that states its size, which is not considered provenance information since it does not relate to how it was created. The same file can have metadata regarding creation date, which would be considered provenance-relevant metadata. So even though a lot of metadata has to do with provenance, both terms are not equivalent. In summary, provenance is often represented as metadata, but not all metadata is necessarily provenance.

18. Logistic / Supply Chain Effects

Data combinations may be constantly and automatically established (for periodical updates). Therefore the quality of services of participating data sources have to be guaranteed. In detail this aspect of logistics calls for appropriate architectures: where and what to cache? How to ensure appropriate network access (broadband)? Where to embed SaaS (Software as a Service)? .... These and other questions occur in the logistic domain and can be directly adopted to geospatial processing.

Logistic/supply chain effects for data combinations are generally applicable to (or will have their main impact on) real-time data combinations, like embedding sensor systems, or big data flows, especially when the amount of data cannot be processed in an adequate timespan.

Annex II: Suggestion for the management processes

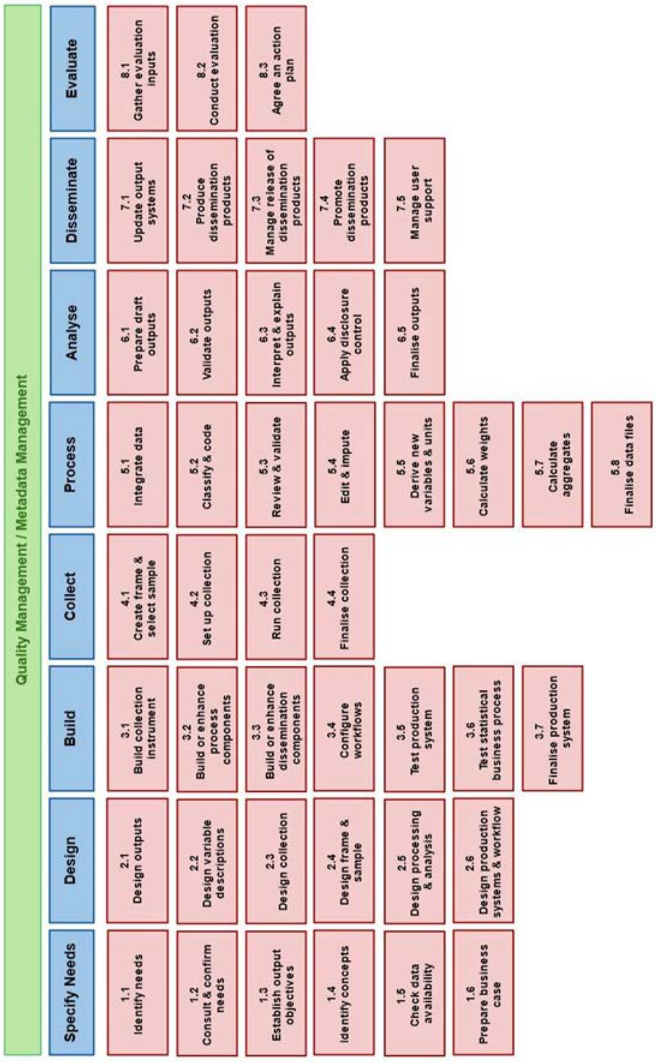
A management process for handling side-effect should set the starting-point in the GSBPM (General Statistic Business Process Model). The GSBPM set processes for quality and metadata but not for management and risk assessment. A proposal is that some chapter should be put in the process documentation that handle side-effects. According to the discussion in the work group some considerations have to be made in the different stages of the GSBPM process.
At the first stage “specify needs” a risk analysis should focus on licenses, ethics and political consequences. The second stage “design” should focus on linked data and semantic.
The next phases ”Build and Collect” should have the focus on quality issues.
Of course all aspects should have being considered in all stages of the process. All cross border initiative should have a steering committee that have skills in the separate countries law and culture to make this risk analysis process effective. The purpose of the steering committee is to handle the risk and side-effects that come up during the risk analysis work. The actions can be to handle or eliminate the risk and side-effects or stop the ongoing process according to the risk. This also concerns positive side-effects that can be promoted.
To get a good effect in the work with side-effects a cross functional and cross border management committee has to be appointed. The group has to contain skills in statistic, legal issue, business models and geospatial data to predict and estimate both positive and negative side-effects.

Risk analysis methods is used in every authority today and is used for handling information security.
A simple risk analysis is a good start when combining information from different authorities but the side-effect that occurs is hard predict in a general view. This is due to different licensing model, different legislation and other complication that occurs when data is combined cross border.
A risk matrix is also necessary, the matrix has to be adopted to the purpose. In this case to the subgroups described in chapter 5. I think that the important part is that the group doing the risk analysis can use a method that they are familiar with but that the risk matrix is centralized.
A process to react on possible negative side-effects can be hard to detail on an operational level. A centralized process including a communication plan on how to communicate and report possible side-effect must be developed and documented.
There are two international standards that can give guidance in risk management ISO 28000 and ISO 31000 and there is a number of methods used in risk analytics that can be useful for predicting side-effects[10].

---

[10] Lerche, Ian; Glaesser, Walter (2006), *Environmental risk assessment : quantitative measures, anthropogenic influences, human impact.*, Berlin: Springer, ISBN 3-540-26249-0, retrieved 27 September 2010; O’Brien, Mary (2002), *Making better environmental decisions: an alternative to risk assessment*, Cambridge, Massachusetts: MIT Press, ISBN 0-262-15051-4, retrieved 27 September 2010 Paperback ISBN 0-262-65053-3;
Shrader-Frechette, Kristin; Westra, Laura, eds. (1997), *Technology and values*, Lanham, Maryland: Rowman & Littlefield, ISBN 0-8476-8631-0, retrieved 27 September 2010;
ISO 28000:2007 Specification for security managment system for the supply chain
http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=44641
ISO 31000:2009 Risk management - principle and guidelines
http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=43170

GSBPM Process:



GSBPM - UNECE statistics Wikis
http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0

Annex III: List of contributors

Active Members of the UN-GGIM: Europe Work Group B “Data Integration”

Name	First name	Institution	Country (code)
Abad	Paloma	CNIG (Geographic Information Unit)	ES
Afflerbach-Thom	Sabine	Federal Agency for Cartography and Geodesy (BKG)	DE
Agius	Carol	UN-GGIM: Europe Secretariat	BE
Dygaszewicz	Janusz	Statistics Poland (CSO)	PL
Etienne	Malte	DESTATIS / Federal Statistical Office of Germany	DE
Eggers	Olav	National Survey & Cadastre Denmark	DK
Fohgrub	Bianka	Federal Agency for Cartography and Geodesy (BKG)	DE
Hagg	Lars	Lantmäteriet Sweden	SE
Haldorson	Marie	Regions and Environment Department Statistics Sweden, Subgroup B.1 leader	SE
Hopfstock	Anja	Federal Agency for Cartography and Geodesy (BKG)	DE
Humbel	Rainer	Swiss Federal Statistical Office (FSO; Bundesamt für Statistik)	CH
Illert	Andreas	Federal Agency for Cartography and Geodesy (BKG)	DE
Jobst	Markus	BEV - Bundesamt für Eich- und Vermessungswesen	AT
Kersten	Berny	Dutch Kadaster	NL
Kutterer	Hansjörg	Federal Agency for Cartography and Geodesy (BKG), Member of the UN-GGIM-Europe Executive Committee, Chair Work Group B	DE
Kuzma	Igor	Statistics Slovenia	SL
Lilje	Mikael	Lantmäteriet	SE
Migacz	Mirosław	Statistics Poland (CSO)	PL
Norris	James	Ordnance Survey	UK
Nunes	Càtia	INE Portugal	PT
Özcan	Şeyma	Turkish Statistical Institute	TR
Petri	Ekkehard	European Commission (Eurostat)	EU

Name	First name	Institution	Country (code)
Radenkovic	Milutin	Statistical Office of the Republic of Serbia	RS
Rainer	Norbert	Statistics Austria	AT
Saarikoski	Antti	National Land Survey of Finland	FI
Santos	Ana Maria	INE Portugal	PT
Schnorr-Bäcker	Susanne	DESTATIS / Federal Statistical Office of Germany	DE
Sekular	Monica	Statistics Poland (CSO)	PL
Smits	Paul	European Commission (JRC)	EU
Tiainen	Esa	National Land Survey of Finland	FI
Ticca	Pina Grazia	GIS Unit - Territorial information and Geographic Information System Division	IT
Vala	Francisco	INE Portugal	PT
van der Vegt	Haico	Dutch Kadaster	NL
Velasco Martín-Varés	Amalia	Spanish Directorate General for Cadastre	ES
Wardzińska-Sharif	Amelia	Statistics Poland (CSO)	PL
Zaccheddu	Pier-Giorgio	Federal Agency for Cartography and Geodesy (BKG), Acting chair Work Group B	DE