Proceedings XII GEOINFO, November 27-29, 2011, Campos do Jordão, Brazil. p 13-26.

Core Concepts of Spatial Information: A First Selection

Werner Kuhn

Institute for Geoinformatics (ifgi)
University of Muenster (Germany)
kuhn@uni-muenster.de

Abstract

The work reported here explores the idea of identifying a small set of core concepts of spatial information. These concepts are chosen such that they are communicable to, and applicable by, scientists who are not specialists of spatial information. They help pose and answer questions about spatio-temporal patterns in domains that are not primarily spatial, such as biology, economics, or linguistics. This paper proposes a first selection of such concepts, with the purpose of initiating a discussion of their choice and characterization, rather than presenting a definitive catalog or novel insights on the concepts.

1 Introduction

Research and development concerned with spatial information have been trying for at least two decades to integrate spatial information into mainstream information technology as well as science and society at large. For example, the vision of OGC (the Open Geospatial Consortium) went from, originally, "The complete integration of geospatial data and geoprocessing resources into mainstream computing" to today's "Realization of the full societal, economic and scientific benefits of integrating electronic location resources into commercial and institutional processes worldwide." The expectation behind the many large-scale efforts to create Spatial Data Infrastructures, nationally and internationally, is that more of the purported 80% of decisions in society with a spatial component will eventually become informed by spatial information, thereby improving business, governance, and science.

Progress toward this goal has been modest and often happened outside or even despite the efforts of spatial information experts, rather than through them. Today's most popular sources of spatial information are imagery and map data from Google, Microsoft and other companies, who acquired and developed their technology largely outside the communities of spatial information experts. The problem has been recognized (and in part dealt with) in terms of a need for simpler models and standards than those typically produced by experts. As a result, simply structured spatial data can now be accessed, and elementary analyses performed on them, using non-specialist data formats and services built around them.

Yet, for spatial information to become a cross-cutting enabler of knowledge and analysis, such bottom-up technological solutions alone are not sufficient. Their confusing letter soup of acronyms and their growing plethora of "standards du jour" do not encourage a broader understanding and communication. An effort at the conceptual level is needed, in order to present a coherent and intelligible view of spatial information to those who may not want to dive into the intricacies of standards and data structures. Spatial information is too valuable for society to leave it up to the specialists; but as specialists, we can do a better job explaining it and demonstrating its benefits.

Spatial information answers questions about *themes* in *space* and *time*. All its varieties result from treating these three components as fixed, controlled, or measured **(author?)** [15]. For example, information about objects is produced by fixing time, controlling theme (i.e., choosing the objects of interest), and measuring space. A unified treatment of time and space turns Sinton's structure into the geo-atom <x, Z, z(x)>, which links a point x in space-time to a property-value pair <Z, z(x)>, where z(x) is the value of property Z at x [10]. The geo-atom answers the questions *what is there?* (looking for z, given x) and *where is this?* (looking for x, given z), two questions whose dualism is characteristic for spatial information.

Spatial data as such are not spatial information, but generate it once humans interpret them. For example, *1-5-3 Yaesu* is spatial data and can be interpreted in the right context as the address of a post office in a ward of Tokyo; the same goes for *6p21.3,* the locus of a gene in a chromosome. Concepts are the mental mechanisms needed to interpret data. For example, the concept of location is needed to interpret addresses or gene loci, and the concept of value is needed to interpret copyright regulations for spatial data.

The overall goal of this work is to explain spatial information and its potential for science and society through a set of core concepts. These are concepts of spatial *information,* defined here as concepts used to answer questions about themes in space and time *as well as* being represented by data or services. They include *spatial concepts*, which serve to reason about space [1] and *information concepts*, in the sense of concepts *about* spatial information, which may be spatial or not. An example of the former is *location,* referring to space and being represented digitally. An example of the latter is *value*, which refers to spatial information, but is not spatial. An example of both is *resolution*, which is a spatial measure, but also describes spatial information. The co-existence of such content and meta concepts, and the consequent need to understand both, are characteristic for information sciences in general.

[1] A comprehensive overview with an emphasis on learning is http://www.teachspatial.org.

2 Spatial information in science and society

Why bother with spatial information at all? First of all, because some of the biggest societal and scientific challenges require a better understanding of, and better decisions about, the location and interaction of things in space and time. Consider climate change, biodiversity, financial systems, poverty, security, health, energy or water supply – spatial information is essential in addressing each of these global as well as many regional and local challenges. In fact, key notions in today's scientific and social debates on these challenges are essentially spatio-temporal - consider risk, sustainability, vulnerability, or resilience. Secondly, studying spaces at smaller scales (atoms and subatomic particles, molecules, crystals, biological cells) as well as larger ones (planets and galaxies) remains among the most fascinating - as well as most costly - scientific adventures. Thirdly, solutions to *non-spatial* problems often use spatial analyses in real or metaphorical spaces, the latter ranging from the data cubes used in data mining to the spatializations used for information retrieval or mnemonic techniques.

Dealing with these and other challenges requires approaches transcending those of single disciplines, even if these disciplines have themselves broad scopes, as geography or computer science do. Transdisciplinary research addresses challenges that span multiple disciplines and have direct social relevance. Its goal is to make progress in solving these problems, not just to gain knowledge. Its approaches often benefit from exploiting spatial information, either because the problems are inherently spatial or because space and time act as unifiers and organizers for all phenomena and our knowledge about them. For example, combining satellite imagery of the Amazon rain forest with ground sensor measurements and socio-economic models reveals deforestation patterns, whose presentation to farmers, decision makers, and the general public helps reduce the depletion of this planet's lungs [1]. Recent developments in the social sciences and humanities, referred to as a *spatial turn* [17], together with numerous technological advances (navigation systems, mobile computing, high-resolution satellite imagery, virtual globes, sensors, and crowd-sourced information) further amplify and exemplify the opportunities for spatial information in science and society.

To be able to bring spatial information to transdisciplinary work, scientists of any disciplines need to be supported in understanding and exploiting *spatiality* in their theories and models. This requires an explanation of spatial information in a theoretically sound and technically informed way, maximizing the scope of applications and minimizing technological jargon. The spaces to be considered are those where transdisciplinary challenges arise, i.e., primarily those of human experience in one to three dimensions plus time. They include geographic spaces (such as a neighborhood in a city or a river catchment), indoor spaces (such as a room or a hallway), body spaces (such as a human body or organ), tabletop spaces (such as a desktop or workbench), and images. Smaller and larger spaces (such as cells, atoms or galaxies) as well as higher-dimensional ones (such as those used in statistics or data mining) are typically understood through mappings to these experiential spaces. Among them, geographic spaces are the ones with which we have the richest set of experiences, giving geographic information science a privileged status in dealing with spatial information.

3 Concepts of spatial information

Surprisingly, a comprehensive treatment of concepts of spatial information with a transdisciplinary scope does not yet exist. All explanations of spatial information have a technological bias toward Geographic Information Systems (GIS) or a disciplinary one toward geography, surveying, or computer science. Where they discuss core concepts at all, these are often limited to spatial concepts or even geometric ones. While there is no shortage of calls to improve this situation (for an early GIS-oriented example with a broad view, see [2]), comprehensive results are missing or only starting to emerge[1].

This lack of a conceptual consensus on spatial information across disciplines, spaces, and technologies may be both a reason for, and a result of, the fact that the "science behind the systems" [7] remains largely there - *behind* the systems. Two decades after it started to call itself a science, and despite significant accomplishments [9], efforts to present geographic information science *to outsiders* as an intellectual (rather than just technological) venture lack a coherent conceptual basis[2].

As a consequence, biologists, economists, or linguists interested in testing a hypothesis using spatial information have to dig into GIS text books and standards documents, even just to understand what spatial information and reasoning could possibly do for them. Those interested in smaller or larger spaces are even worse off, since no textbooks or standards address these across disciplines. Geographic information science, thus, has to ask itself where economics would be today if its core concepts were limited to either micro- or macro-economics and hidden in specialized textbooks and manuals for spreadsheets and accounting software.

Taking up this challenge, the work presented here is neither about technologies, nor about particular disciplines or domains. It attempts to cut across their boundaries, targeting a bigger role for spatial information in science and society. It strives for an explanation of *what is special about spatial* for those who are *not* specialists of spatial information.

4 A first selection of concepts

The discussion of the proposed nine core concepts of spatial information in this section begins with location and the dual pair of field and object information. It continues with the spatial concepts of network and process, and ends with the information concepts of resolution, accuracy, semantics, and value. Each

[2]Some examples (with a variety of goals) are http://spatial.ucsb.edu/, http://uspatial.umn.edu/, http://spatial.uni-muenster.de, http://lodum.de/

concept gets only briefly characterized, in order to initiate and guide discussions. Detailed explanations exceed the available space and will be provided in a revised and extended article [13].

4.1 Location

The starting point for a journey through concepts of spatial information has to be location: spatial information is always linked to location in some way - but what exactly is location and how does it play this central role? Location information answers *where* questions: where are you? where is the appendix? where are this morning's traffic jams? Perhaps counter-intuitively, location is a relation, not a property. This is so, because nothing has an intrinsic location, even if it remains always in the same place. The house you live in can be located, for example, by a place name, an address, directions, or various types of coordinates. All of these location descriptions express relations between the *figure* to be located (your house) and a chosen *ground* (a named region, a street network, coordinate axes). How one locates things, i.e., what ground and what relation one chooses, depends on the context in which the location information is produced and used. Spatial reference systems, for example the World Geodetic System 1984 (WGS84), standardize location relations and turn them into easier to handle attributes within such a system. Yet, when data use multiple reference systems (for example, latitude and longitude as well as projected coordinates), locations need to be understood as relations and interpreted with respect to their ground (for example, the Greenwich meridian or a projected meridian).

Relating different phenomena through location is fundamental to spatial analysis. An early and famous application is Dr. Snow's 1854 finding that Cholera is spread by drinking water, after he had observed that many Cholera deaths had taken place around a certain water pump. The great power of such locational analyses results from the fact that nearby things are more related than distant things, which has been dubbed Tobler's First Law, based on its first explicit statement in [16].

4.2 Field

Fields describe phenomena that have a value everywhere in a space of interest, such as temperature. Generalizing the field notion from physics, field-based spatial information can also represent values that are statistically constructed, such as probabilities or population densities. Field information answers the question *what is there?,* where "there" can be anywhere in the space of interest. Fields are one of two fundamental ways of structuring spatial information, the other being objects 4.3. Both fix time, with fields resulting from controlled space and measured theme, and objects resulting from controlled theme and measured space. Time can also be controlled, rather than fixed. Controlling it together with space leads to space-time fields; controlling it together with theme produces object animations. Fields have been shown to be more fundamental than objects, capable of integrating field and object views in the form of General Field models [14].

Since it is not possible to represent the infinitely many values of a field, they need to be discretized for explicit digital storage. There are two ways to achieve this, either through a finite set of samples with interpolation between them or through a finite number of cells with homogeneous values, jointly covering the space of interest. The cells can all have the same shape (forming a regular grid of square, triangular, hexagonal, or cubic cells) as in the so-called *raster model* for spatial data, which is best known from digital images. Or they can have irregular shapes, adaptable to the variation of the field, as in the finite element models used in engineering or the triangulated irregular networks (TIN) used to represent terrains.

An important kind of fields captures values on two-dimensional surfaces such as those of the earth or of the human body. These fields are typically organized into thematic *layers*. The idea of a layer is rooted in traditional paper- or film-based representations of spatial information, such as maps, and the production of models from stacked transparent layers of data about a theme. The main computational use of layers is to *overlay* them in order to relate information about multiple themes or from multiple sources.

4.3 Object

After fields, objects provide the second fundamental way of structuring spatial information. They describe individual things with spatial, temporal, and thematic properties and relations. Object information answers questions about properties and relations of objects, such as *where is this?, how big is it?, what are its parts?*, *which are its neighbors?, how many are there?*.

For many applications, one is interested in things that are *features* in the way that a nose is a feature of a face, i.e. parts of a surface. Features are important siblings of objects, but can be understood as a special case. The simplest way to carve out features from a surface is to name regions on it. Geographic places are the prototypical examples, carved out of the earth's surface by naming regions; but the same idea applies, for example, to models of airplane wings, sails, or teeth. Object and feature models can co-exist, and the general tendency today is to complement two-dimensional feature models with three-dimensional object models and provide more or less seamless transitions between them. For example, your house may be represented as a feature of the earth's surface in a digital map, as a feature of street view images, and as a three-dimensional object. The resulting blended feature-object notion pervades geography, but also exists in biology and medicine (features of cells, organs, or bodies), and generally in imaging (features extracted from images of anything).

Many questions about objects and features can be answered based on simple representations as points with thematic attributes. For example, doing a blood count or determining the density of hospitals in an area require only point representations. On the other hand, some questions do require explicit *boundaries* enclosing or separating objects. For example, determining the neighbors of a land parcel, the extent of a geological formation, or the health of blood cells may require boundary data. The frequent occurrence of boundaries in object information, however, has mainly historical reasons, since analog representations like images or maps were digitized by drawing or following lines on them.

The so-called *vector models* for spatial data capture objects with boundaries at various levels of sophistication. Like surface fields in raster data, collections of features in vector data can be organized into thematic *layers*. Processing vector data exploits the geometry of boundaries to compute sizes, shapes, buffers around the objects, and overlays. Yet, many objects, particularly natural ones, do not have crisp boundaries [3]. Examples are geographic regions or body parts such as the head. It can be harmful to impose boundaries on such objects only for the sake of storing them in vector models. Differences between spatial information from multiple sources are indeed often caused by such more or less arbitrary delimitations. For example, boundaries of climate zones are vague by nature, and the variation in boundaries between different definitions matters much less than the overall extent and location of the zones. Thus, whether modeling objects with explicit boundaries is necessary or even desirable has to be carefully assessed for each application. It is certainly not something that the concept of an object implies.

4.4 Network

Connectivity is central to space and spatial information. The concept of a network captures binary connections among arbitrary numbers of objects, which are called nodes or vertices of the network. The nodes can be connected by any relation of interest. Network information answers questions about connectivity, such as *are nodes m and n connected?*, *what is the shortest path from m to n?*, *how central is m in the network?*, *where are the sources and sinks in the network?*, *how fast will something spread through the network?*, and many others.

The two main kinds of networks encountered in spatial information are transportation and social networks. Transportation networks (in the widest sense) model systems of paths along which matter or energy is transported, such as roads, utilities, communication lines, synapses, blood vessels, or electric circuits. Social networks capture relationships between social agents, such as friendships, business relations, or treaties.

All networks can be spatially embedded, which means that their nodes are located. This is often the case for transportation networks and increasingly for social networks. If the embedding space is a surface, networks can be organized into thematic layers, like the surface fields and feature collections encountered above.

Network applications benefit from the well studied representations of networks as graphs and the correspondingly vast choice of algorithms. Partly due to this sound theoretical and computational basis, networks are the spatial concept that is most broadly recognized and applied across disciplines.

4.5 Process

Processes are of central interest to science and society - consider processes in the environment, in a human body and its cells, and in machines or molecules. Processes that manifest themselves in field, object, or network information are considered spatial. Information about spatial processes primarily answers questions about *motion*, *change*, and *causality*.

Controlling time and measuring space generates information about *motion*; controlling time and measuring theme informs about *change*. Time is typically controlled through time stamps in spatial information. Temporal reasoning on time stamps (and on time intervals formed from them) is the basis for understanding motion and change. Migration or embolism are examples of motion. Growth, such as that of vegetation or social networks, exemplifies the change of objects or networks. Diffusion, for example in the form of climate change, collapsing house prices, or spreading innovation, is an example of a change in fields, objects or networks.

The most complex relation between spatial processes is that of *causality*. Dr. Snow's tracing of cholera to drinking water is a case of determining that one process (drinking some water from a contaminated pump) is the cause of another (contracting cholera), based on the patients being located near the pump.

Real-time spatio-temporal data from sensors and spatio-temporal simulations are the two key sources of process information. In order to makes sense of these dynamic data and models, science needs better *theories of change* [4]. One of the main benefits to be expected from a list of core concepts of spatial information is indeed to establish the conceptual foundations for such theories. If the theories can be formulated in terms of the proposed core concepts, their choice will be corroborated; if not, other concepts will have to join or replace them. For the current proposal, this means that all spatial change needs to be explained in terms of operations on locations, fields, objects, networks, and processes themselves.

4.6 Resolution

Resolution is the first and most spatial *concept of information* on this list. It characterizes the size of the units about which information is reported and applies to all three components of space, time, and theme. For example, satellite images have the spatial resolution of the ground area corresponding to a pixel, the temporal resolution of the frequency at which they are taken, and the thematic resolution of the spectral bands pictured. Vote counts have the spatial resolution of voting districts, the temporal resolution of voting cycles, and the thematic resolution of parties or candidates. Resolution information answers questions about *how precise* spatial information is, for example, when taking decisions based on the information.

Resolution characterizes information about all concepts introduced so far: location is recorded at certain granularities, fields are recorded at certain sample spacings or cell sizes, and the choice of the types of objects (say, buildings vs. cities) and nodes (say, transistors vs. people) determines the spatial resolution of object and network information. The choice of the spatial, temporal, and thematic resolution at which spatial information gets recorded is primarily determined by the processes studied, because these involve phenomena of certain sizes, frequencies, and levels of detail. For example, migration, social networking, and the diffusion of technological innovations all involve people over months; embolism involves blood clots and vessels over hours; cancer involves cells and organs over years; climate change involves large air and water masses over decades; changing house prices involve land parcels and people over days or weeks.

Many processes need to be studied at multiple resolutions (for example, erosion) or they connect to processes at other resolutions. For example, one can think of all processes as involving some sort of motion at some resolution. All five core spatial concepts on our list can be represented at multiple resolutions: location descriptions are often hierarchical (for example, addresses); fields are often represented by nested rasters (called pyramids in the case of images); object hierarchies are expressed as part-whole relations between objects (for example, administrative subdivisions of countries); hierarchical network representations allow for more efficient reasoning (for example, in navigation), process models (for example, in medicine) are connected across levels of detail.

4.7 Accuracy

Accuracy, like precision, is a key property of information, capturing how information relates to the world. Information about accuracy answers questions about the *correctness* of spatial information. The location of a building, given in the form of an address, coordinates, or driving instructions, can in each case be more or less accurate. The spatial, temporal, and thematic components of spatial information are all subject to (in)accuracy.

Assessing the accuracy of information requires two assumptions: that there is in principle a well-defined correct value and that repeated measurement or calculation distributes in regularly around it. The first assumption requires an unambiguous specification of the reported phenomenon and of the procedure to assign values. For example, if temperatures are reported for different places, one may need to specify the level above ground to which they refer. The second assumption requires an understanding of measurement as a random process. Choosing a particular form of distribution (called a probability density function) allows for estimating the probability that a measured or computed value falls within a given interval around the correct value. Mean errors and any other accuracy data are based on these two assumptions.

Accuracy connects to resolution through the established practice of reporting all data at a resolution corresponding to the level of expected accuracy. If information is collected at multiple levels of resolution, one level can sometimes be considered as accurate when assessing the others. For example, positions determined from high-precision measurements serve as "fix points" for lower precision measurements, and objects get extracted from remotely sensed images by determining "ground truth" for parts of an image.

4.8 Semantics

Understanding the semantics of spatial information is crucial to its adequate use. When it comes to analyzing spatial information, determining whether the same things are called the same (and different things differently) is essential to producing meaningful results and making sense of them. The challenge is to capture what the producer means with some data or services and to guide the user on how to interpret them. For example, when navigation systems use road data, they make assumptions on what the data producer meant by "road width" (paved or drivable?, number of lanes or meters or feet?). When using a spatial information service, operational terms such as distance also have to be interpreted adequately.

Semantic information answers the question *how to interpret the terms* used in spatial information. It concerns the spatial, temporal, and thematic components. While the semantics of spatial and temporal data have long been standardized through spatial and temporal reference systems, the semantics of thematic data and operations remain hard to capture and communicate. What is meant with data about land use or body tissue, for example, depends on a complex interaction between defining the intended use of some terms (say, forest or muscle) and delineating the spatio-temporal extents of their application to land or tissue.

Data and services do not have a meaning by themselves, but are used to mean something by somebody in some context. Therefore, it is impossible to fix the meaning of terms in information. However, one can make at least some of the conditions for using and interpreting a term explicit. This is what *ontologies* do: they state constraints on the use of terms. But language use is flexible and does not always follow rules, even for technical terms. An empirical account of how some terms are actually used can therefore provide additional insights on intended meaning or actual interpretation. This is what *folksonomies* deliver: they list and group terms with which information resources have been tagged.

Semantic information consists of necessarily incomplete collections of constraints from ontologies and folksonomies on the use and interpretation of terms. The constraints can use binary logic (for example, stating that a term refers to a subset of the things that another term refers to) or fuzzy logic (where such a statement is neither true nor false, but possibly true). The latter is an attempt to account for the inherent vagueness of many terms.

Yet, all constraints depend on *context*. Terms are used by somebody to mean something in a given context. Ontologies and folksonomies capture some aspects of context, but spatial information is often used in other contexts than the ones it was produced in. For example, road width data produced by traffic engineers may be quite different from those needed for navigation. In order to map between different contexts, ontologies need to be *grounded*. This means that their constraints need to refer to something outside their context, to which the constraints of other contexts can then refer as well. Spatial information has successfully relied on grounding for centuries, through spatial reference systems. These systems refer coordinates to something outside their conceptual framework, such as physical monuments or stars. Generalizing this idea from location information to any terms used in spatial information leads to the idea of semantic reference systems [12]. These systems, once established in practice, are expected to support translations of terms used in spatial data and services from one context to another. Grounding is the basis for analytical translations of terms from one context to another. Since all constraints on meaning are non-deterministic, stochastic approaches to translation are a valid alternative to explicit grounding. For example, translations of terms across natural languages are now routinely and successfully achieved through stochastic methods.

4.9 Value

The final core concept proposed is that of value. Information about the value of spatial information answers questions about the many *roles* spatial information plays in society. The main aspect of value is *economic*, but the valuation of spatial information as a good in society goes far beyond monetary considerations. It includes its relation to other important social goods, such as privacy, infrastructure maintenance, or cultural heritage.

Setting policies on public *access* to spatial information, for example, is a pressing societal need requiring a better understanding of the many valuations involved. It is further complicated by the fact that information about indoor and geographic spaces can now be and is being collected and shared by almost everybody. This phenomenon of crowd-sourced or Volunteered Geographic Information (VGI, [8]) is profoundly altering the value of spatial information, from economic as well as institutional, ethical, and legal perspectives. For example, a key new challenge created by VGI is to understand and model *trust* in spatial information.

Given these wide ranging aspects of spatial information value, no coherent theoretical framework for it can be expected any time soon. Partial theories of value, for instance about the economic value of spatial information, are still sketchy and difficult to apply, because they involve parameters that are hard to control or measure. The cost of spatial information is no good guide to its value either, because it often reflects the high expenses for collecting the information, rather than the value of the result.

Value of information tends to accrue holistically and unpredictably, by new questions that can be asked and answered, new services that are provided. Partly for this reason, spatial information holdings have become significant assets, not only for scientists and governments, but also for enterprises in all sectors. Such assets need to be evaluated, for example in enterprise valuation, reinforcing the need for theories of spatial information value.

Even at the level of personal information management, the value of accessing and analyzing information through its spatial and temporal properties has barely been understood and tapped into yet [11]. For example, searching information by where or when it was collected or stored is highly effective, but still only weakly supported by the web, personal computers, and smart phones.

4.10 Also-ran

It may be useful to consider some arguments against core status for some other concepts. Obviously, these may have to be reconsidered, so that this list of also-rans is part of the material for discussion.

My earlier lists of concept candidates contained nearness, spatial relations, feature, map, layer, motion, path, uncertainty, and scale. Typical reasons to exclude them from the list were that they were too broad or too narrow. In particular:

- *nearness* got generalized to *spatial relations*, but these serve to specify location and are covered there;
- *features* are now treated together with objects;
- *maps* are visualizations of mostly geographic information that exists in other forms;
- *layers* structure the representations of several concepts (fields, objects, networks) and are dealt with there;
- *motion* is only one process in space, although the most important one;
- *paths* are covered as parts of networks;
- *uncertainty* covers several concepts, of which resolution, accuracy, and semantics are covered;
- *scale* is also a catch-all for several concepts, of which resolution is on the list, extent (of a study area) is rather trivial, and support is more specialized (belonging to measurement ontology).

5 Conclusions

Achieving a stronger role for spatial information in science and society requires explaining its uses and benefits at a higher level than that of technologies and acronyms. The small set of concepts of spatial information proposed in this paper indicates a possible basis for such explanations. While it may miss or misrepresent some concepts, it provides a starting point to reach a conceptual view of our field that is accessible and intelligible to outsiders. The main goal at the moment is, therefore, to receive critical feedback and suggestions of what to add, drop, or change.

The concepts chosen and revised based on the expected feedback will then be described in more detail over the coming year[3]. These descriptions will ask and answer four questions about each concept:

[3] To participate in the discussion, please visit http://ifgi.uni-muenster.de/services/ojs/index.php/ccsi/index

1. *what* is the concept, i.e., what phenomena does it capture?
2. *where* does information about the concept come from, i.e., what are typical sources of information about it?
3. *how* is the concept *represented*, i.e., what data structures and algorithms implement it?
4. *how* is information about the concept *used*, i.e., what reasoning and analyses does the concept support?

The expected result is a catalogue of core concepts that are meaningful and useful across disciplines - a vocabulary to talk about spatial information to non-specialists. Such vocabularies, when formalized, are referred to as ontologies. While formalization is not a primary goal here, treating the concepts as nodes in an ontology and relating them to an upper level ontology or embedding them in ontology patterns will certainly help to clarify them further. Starting this work as an ontology design exercise, however, would most likely not lead to a useful set of concepts, because their relation to actual data and computations would be too weak. A subsequent ontological analysis will produce an ontology of spatial information that allows for interfacing with other domains, while relating explicitly to information technology. As such, it will complement and benefit from existing ontologies of spatial information [6, 5].

Acknowledgments

Countless discussions over the years with many colleagues and friends have encouraged and influenced these thoughts. The members of http://musil.uni-muenster.de and the students of my *Introduction to Geographic Information Science* have been very helpful critics and supporters of this work. Some anonymous reviewers of GeoInfo2011 provided very useful comments.

References

[1] Ana Paula Dutra Aguiar, Gilberto Câmara, and Maria Isabel Sobral Escada. Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. *Ecological Modelling*, 209(2-4):169–188, December 2007.

[2] Peter A Burrough and Andrew U Frank. Concepts and paradigms in spatial information: Are current geographic information systems truly generic? *International Journal of Geographical Information Science*, 9(2):101–116, 1994.

[3] Peter A Burrough and Andrew U Frank, editors. *Geographic Objects with Indeterminate Boundaries*. Taylor&Francis, December 1996.

[4] Gilberto Câmara, Lúbia Vinhas, Clodoveu Davis, Fred Fonseca, and Tiago Carneiro. Geographical Information Engineering in the 21st Century. In Gerhard Navratil, editor, *Research trends in geographic information science*, pages 203–218. Springer Verlag, 2009.

[5] Helen Couclelis. Ontologies of geographic information. *International Journal of Geographical Information Science*, 24(12):1785–1809, November 2010.

[6] Andrew U Frank. Ontology for spatio-temporal databases. Spatiotemporal Databases. In Timos Sellis, editor, *The Chorochronos Approach*, pages 9–77. 2003.

[7] Michael F Goodchild. Geographical information science. *International Journal of Geographical Information Science*, 6(1):31–45, 1992.

[8] Michael F Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, November 2007.

[9] Michael F Goodchild. Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science*, 1(1):3–20, July 2010.

[10] Michael F Goodchild, May Yuan, and Thomas Cova. Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3):239–260, 2007.

[11] Krzysztof Janowicz. The Role of Space and Time For Knowledge Organization on the Semantic Web. *Semantic Web - Interoperability, Usability, Applicability*, 1(1-2):25–32, 2010.

[12] Werner Kuhn. Semantic Reference Systems. *International Journal of Geographic Information Science (Guest Editorial)*, 17(5):405–409, June 2003.

[13] Werner Kuhn. The sciences before the systems: towards a transdisciplinary role for spatial information. *International Journal of Geographical Information Science*, (under review):1–15, 2012.

[14] Y. Liu, Michael F Goodchild, Q. Guo, Y. Tian, and L. Wu. Towards a General Field model and its order in GIS. *International Journal of Geographical Information Science*, 22(6):623–643, 2008.

[15] David Sinton. The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. *Harvard papers on geographic information systems*, 6:1–17, 1978.

[16] Waldo Tobler. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46:234–240, 1970.

[17] Barney Warf and Santa Arias, editors. *The spatial turn: Interdisciplinary perspectives*, volume 26. Taylor & Francis, November 2009.