Towards Knowledge-Based Spatial Planning

Chapter · January 2019

DOI: 10.4018/978-1-5225-7927-4.ch001

Robert Laurini
Institut National des Sciences Appliquées de Lyon

Towards Knowledge-Based Spatial Planning

Robert Laurini
Knowledge Systems Institute, USA, and INSA-Lyon, University of Lyon, France

ABSTRACT

From millennia, spatial planning has been based on human knowledge about the context and its environment together with some objectives of development. Now, with artificial intelligence and especially knowledge engineering, practices of spatial planning can be renovated. Presently due to the capacity of automatic reasoning, novel practices can be designed. In addition to human collective knowledge either verbal or visual, some new chunks of knowledge can be introduced, coming from physical laws, administrative regulations, standards, data mining and best practices. In addition, by means of different techniques, huge sets of data are daily collected in smart cities. By big data analytics, some regularities and patterns can be discovered, which again will lead to new actions towards cities: in other words, there is a virtuous circle linking smart territories and big data, which can be the basis for novel spatial planning. The role of this chapter will be to analyze those new chunks of knowledge and to explain how human knowledge, possibly coming from different stakeholders, can be harmonized with machine-processable knowledge as to be the basis for territorial intelligence.

Keywords: Geographic Knowledge, Urban Knowledge, Smart Cities, Urban Planning, Smart Planning, Governance, Geographic Rules, Urban Rules, Territorial Intelligence.

INTRODUCTION

The use of computing in urban planning began with the Baxter's book (BAX, 1976) which was including several statistical and mathematical modeling aspects. Then gradually cartography and databases were integrated to give the well-known Geographic Information Systems in the 80s. Little by little some spatial analysis tools were included. Now the context is different for several reasons.

- The existence of many sensors, especially for environment and traffic controls is now mainstream; the consequence is that the novel GIS must capture those data in real time.
- The new development of visualization, giving the so-called geo-visualization has greatly renovated cartography by integrating animation, research of salient features and so on.
- The development of big data has led local authorities to envisage new methods to use those data by enriching their knowledge about city's evolution.
- The trend of volunteered geographic information and the will of people to participate in decision making, under the banner of crowdsourcing have implied the necessity of dealing efficiently with those characteristics.
- Empowerment of people has increased the importance of several stakeholders who can have different logics; those logics can interfere with local authorities and can be in contradiction.
- The advances in artificial intelligence can also be integrated, transforming GIS into Geographic Knowledge systems.
- The multiplicity of experiences in various cities throughout the world has led to new organizations not only for technology watching, but overall by sociological watching, i.e. examining novel experiences which can be imported.

The objective of this chapter will be to study how urban knowledge can help smart urban planning. After presenting the background of this study, we will examine how to combine various types ok knowledge.

BACKGROUND

In this section, the concept of knowledge and especially geographic knowledge will be examined. Then a structure of geographic knowledge base will be studied.

Knowledge and Geographic Knowledge

In information technology, it is common to distinguish data, information, knowledge and wisdom. According to Gurteen (1998), the cake metaphor can help understand the differences between those concepts: "data" corresponds to molecular components of the cake; "information" to the ingredients; "knowledge" to the recipe (how to make the cake); and finally, "wisdom" corresponds to know why and for whom to make the cake. In other words, for urban issues, "data" come from sensors, census, aerial or drone photos, satellite images, pools, etc.; "information" corresponds to the meaning of the data as usual, and finally "knowledge" and "wisdom" for whom and with what objectives to make the city.

All those previous issues lead to consider new approaches in urban planning, especially based on knowledge. So, for our purpose "geographic knowledge corresponds to information potentially useful to explain, manage, monitor, understand the past, plan a territory and innovate". Let me develop.

a/ Geographic Knowledge to Explain. It corresponds more or less to Gollege's definition. Synonyms can be to understand, to explore, to assess the context and to detect problems. Existing books and monographs can help a lot from an historical point of view. Techniques such as geographic text mining (Salaberry 2013) and, when databases are existing, spatial data mining can be the sources of this kind of geographic knowledge. An extension of this category is to consider knowledge to reconstitute the past landscape or to simulate future evolution.

b/ Geographic Knowledge to Manage. One of the goals of local authorities is to manage the territory under their jurisdiction. The management could range from street and engineering network repairs to school and other public services such as waste collection. The knowledge they have to use is essentially coming from laws, by-laws and best practices. In other words, knowledge is identified in some natural language sentences and must be transformed to become machine-processable. Often, can knowledge be here seen as an extension to business intelligence applied to local authorities?

c/ Geographic Knowledge to Monitor. This kind of knowledge can be seen as an extension of the previous one, but its nature is totally different. Indeed, local authorities in order to reduce pollution or regulate traffic, install sensors as previously explained to get raw data which are possibly transformed into knowledge by real time data mining.

d/ Geographic Knowledge to Plan. In my understanding, this is the ultimate goal of geographic knowledge engineering, to plan smart cities or territories. It means to design scenarios of evolution, study alternatives and take citizen's opinions into account within the scope of sustainable development.

e/ Geographic Knowledge to Understand the Past. This is another usage of knowledge, for instance in archeology or in history. By examining excavation findings, already-known ancient knowledge can be used to understand, but also the findings can suggest novel theories, for instance for commercial exchange.

f/ Geographic Knowledge to Innovate. This kind of knowledge is important for the future of the smart cities; it will be essentially based on technological and sociological watches.

Rules constitute a very important component of knowledge. In artificial intelligence, the representation of rules is based on several mathematical theories, such as classical logics. Moreover, according to Graham (2006), Morgan (2008), and Ross (2011), business rules should be considered as first-class citizens in information technology, and, why not, in urban planning? But before examining geospatial rules, it is necessary to revisit certain aspects of geographic object modeling so that to introduce them into rules. For them, rules can be modeled as `IF-THEN-Fact` and `IF-THEN-Action`, the first one for generating new data, for instance the calculus of VAT in business, and the second for launching a new action, this action being made by a computer, a robot or a human.

But, for geospatial rules, some other models are necessary (Laurini, 2018), namely:

- `IF-THEN-Zone`: for the creation of a zone, based for example of its coordinates;
- Co-location rules: "if something here, then another thing nearby";
- Bi-location rule: "IF something holds in place P, then something else in place Q", to model the so-called "butterfly effect";
- Located rule: "IF in a place B, THEN apply RuleB";
- Metarule: "IF some conditions hold, THEN apply RuleA".

Structure of a Geographic Knowledge Base

In Laurini (2017) a new structure of geographic knowledge bases has been proposed as depicted in Figure 1. They will consist of a set of geographic objects, a set of geographic relations, an ontology, a gazetteer, a set of geographic structures, a set of physico-mathematical models and a set of rules; in addition, external knowledge can also be very useful. For more details, please refer to Laurini (2017a) and Laurini-Favetta (2017). Remember that a gazetteer is a repository for storing placenames, and ontologies for the description of all types and classes of geographic objects. By external knowledge, we mean two things: (*i*) neighboring knowledge to ensure reasoning continuity, and (*ii*) knowledge coming from outside for both technological and sociological watches which will be the background for innovation.

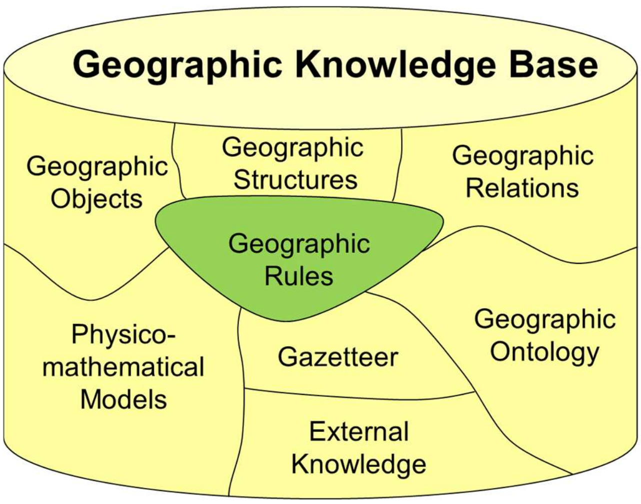

Figure 1. Structure of a Geographic Knowledge Base

Big data and Big Data Analytics

According to Snijders *et al*. (2013), “Big Data” is a loosely defined term used to describe data sets so large and complex that they become awkward to work with, using standard statistical software. Now, with telecommunications and sensors, cities can get lots of data, from which characteristics, regularities or patterns may be discovered through sets of algorithms generally named Big Data Analytics. Among those characteristics, novel knowledge chunks can be discovered.

According to Mitchell (2014), Big Data Analytics has several characteristics; among them, mention the necessity to use cloud computing, the importance of “big data lakes”, predictive analysis, new data models and deep learning. Even if the development of those technologies is outside the goal of this chapter, let me only explain the following; in the past, we first designed a data base model and then populated it; but now we do have a “data lake”, *i.e*. without any a priori model. The challenge of big data analytics is so more complex to be achieved. As a consequence, a sort of virtuous circle can be defined as depicted in Figure 2: throughout the smart city, lots of sensors (including user-generated) can measure data which can generate knowledge about the smart city. By accumulating knowledge, it is expected to design new micro-theories regarding cities (Batty 2013).

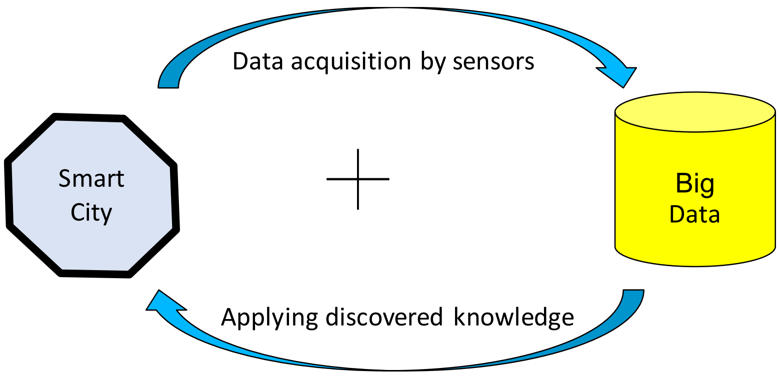

Figure 2. The Smart City/Big Data virtuous circle

But there is a great difference with big data relative to the business industry, it is the importance of GIS data. In the general discourses regarding big data, GIS data are rarely mentioned, maybe they are implicit, but rarely explicit as an essential component for smart cities. However, in addition to coordinates which can be easily assigned to sensors, either fixed or mobile, one must and needs to make connections with the shape and location (often 2D, but more and more 3D) of several geographic and urban features.

ORIGIN OF KNOWLEDGE IN SMART CITIES

In a smart city, knowledge chunks can come from different origins and have different semantic meanings. Let's examine some of them, namely physical knowledge, legal knowledge, practical knowledge, associative knowledge and correlative knowledge.

Physical laws

Those rules are mandatory and must not suffer any exception; they can be considered as components of cause-effect chains. In this category, let us mention any physical geography rules such as rules linked to rivers (floods), diffusion of noise, pollution, earthquake, etc. In other words, they are true implications without any exception.

However, we have to face two categories, permanent conditions and incidental occurrences.

- Consider a city built on a steep mountain or another one on a flat plain: those physical conditions will imply different decisions regarding the structure and the organization of the city, essentially facing mitigation or not of possible disasters.
- Whereas a river traversing a city, the consequences are totally different whenever it is quiet or generating a flood. Similarly, remember the quiet Roman city of Pompei when suddenly the Vesuvius volcano erupted in AD 79. As nobody can forecast the date of the future disasters, mitigation actions must be developed as soon as possible. Another class corresponds to physical events for which some environmental regulations exist, for instance for air pollution. Among incidental occurrences, let also mention rain and hail storms which can damage houses and plantings. Sensors can help a lot to give early alarms.

Those types of knowledge chunks must always hold. However, the second category needs to integrate additional sensor information (often as thresholds) in the chunks.

As some rules can be stated as simple implications, the use of differential equation can often help model those mechanisms.

Rules coming from legal issues

Those rules are mandatory whereas some of them can lead to several alternative options. For instance, in several countries, vehicles are running right. It means that it is mandatory to drive right, but it is possible that some vehicle do not follow this law. In this case, if there is a policeman, they will be subject to a ticket. Anyhow, when structuring a transportation plan, this law is used as a basis.

In urban planning, consider a law in which there are limitations for construction, for instance the number of floors. When a citizen is applying for a building permit, this rule, together with other rules, is mandatory. In addition, in a zone in which this law applies, it is possible than elder buildings are higher than actually planned, for instance a very old listed monument (for instance a cathedral). Do we have to demolish them to follow the new planning rule, or to consider them as an exception?

In this category, we can also consider standards and norms relative to socio-economic ratios, such as the number of elementary school pupils.

Various styles of laws can be considered:

- Mandatory laws for which usually sanctions are defined whenever they are not followed;
- Laws allowing certain possibilities; for instance, the possibility to build a new camping site; maybe during years, there will be no camping site.
- Laws opening several alternatives; in this case, one day, one of them will appear.
- Laws prohibiting some types of construction, for instance a stadium downtown.

Those types of knowledge chunks must always hold, except if there are well-known derogations.

Best Practices

By considering a rule such as "when designing a new metroline, move underground networks", it could be classified as an engineering rule or a best practice. It can look as a legal rule, but it is more a recommendation. In other words, the implication is not mandatory.

Associative Knowledge

The origin of associative knowledge comes from data mining in which one tries to exhibit so-called frequent associations such as (A, B) → C meaning that "if we have A and B, then we have often C", often being characterized by various measures. For instance, Shekhar- Huang (2001) and Shekhar- Zhang, P. (2006) have found they have discovered the following spatial associations:

- (West Nile disease, stagnant water sources, dead birds, mosquitoes)
- (Tow, police, ambulance)
- (Cold front, warm front, snow fall)

Those types of knowledge chunks must hold only if the global performances are better than without using them.

Knowledge Coming from Big Data Analytics, Correlative Knowledge

Knowledge chunks coming from data analytics have a different status. As the causality is either physical, legal or practical, for knowledge chunks originating from data analytics, they are linked by correlation. And correlation does not mean causality. Two well-known Latin sophisms can help us, "Post hoc, ergo propter hoc" and "Cum hoc, ergo propter hoc" which respectively mean "after this, therefore because of this" and "together with this, therefore because of this". In other words, as the antecedents are clear, and the consequents are also clear, the signification of the implication; maybe it is only statistical coincidence. Under the name of correlative knowledge, they can be considered as a sort of extension of associative knowledge.

But everybody has to be cautious of possibly detected spurious correlations. The website http://www.tylervigen.com/spurious-correlations gives a set of spurious correlations such as "the number of divorces in the State of Maine" and the "consumption of margarine" with a correlation coefficient of 0.992558!

Suppose one is interested to reduce the number of road accidents. He is facing two results:

> 1 – The following relation was discovered: "the consumption of sour cream, and the number of motorcycle riders killed in accidents" (http://www.dailymail.co.uk/sciencetech/article-2640550/Does-sour-cream-cause-bike-accidents-No-looks-like-does-Graphs-reveal-statistics-produce-false-connections.html).
>
> 2 – Kumar and Toshniwal (2015) have discovered the following cluster: "*It consists of 69 % of two wheeler accidents which are distributed on intersections near markets, hospitals, local colonies across highways and non-highway roads. Those accidents which occurred on intersections and curves on highways involved one injury only. Two-wheeler accidents at non-highway locations are mostly involved two injuries".*

Does it mean that, if we want to reduce those accidents, we must either impose limitations of selling or buying sour cream in all shops of the city, or redesign some roads in specific locations? For anybody the second solution looks more relevant. But when we are facing several results from big data analytics, what could be the procedure to detect the more relevant chunks of knowledge? Are they criteria or rules to govern such a selection?

As a consequence, it is interesting when knowledge chunks coming from data analytics hold in one context; but in contrast, no matter when they are not taken into consideration. In other words, the various origins imply various statuses and various semantics. More developments look necessary to clarify those semantics and to imagine their outcomes.

A General Model for Rules

Based on the previous descriptions, a general diagram for rules can be designed regrouping all aspects, their origin, components, temporal dimension, mathematic tools, their management, their usage and the various modes of implication.

- For physical rules, the implication is mandatory;
- For legal rules, the implication is also mandatory, but the sanctions may or may not exist;
- For best practices, the implication is more or less a kind of recommendation; in other words, nobody is obliged to follow this kind of rules; perhaps some additional conditions could be considered; at a first approximation, random variates can help.

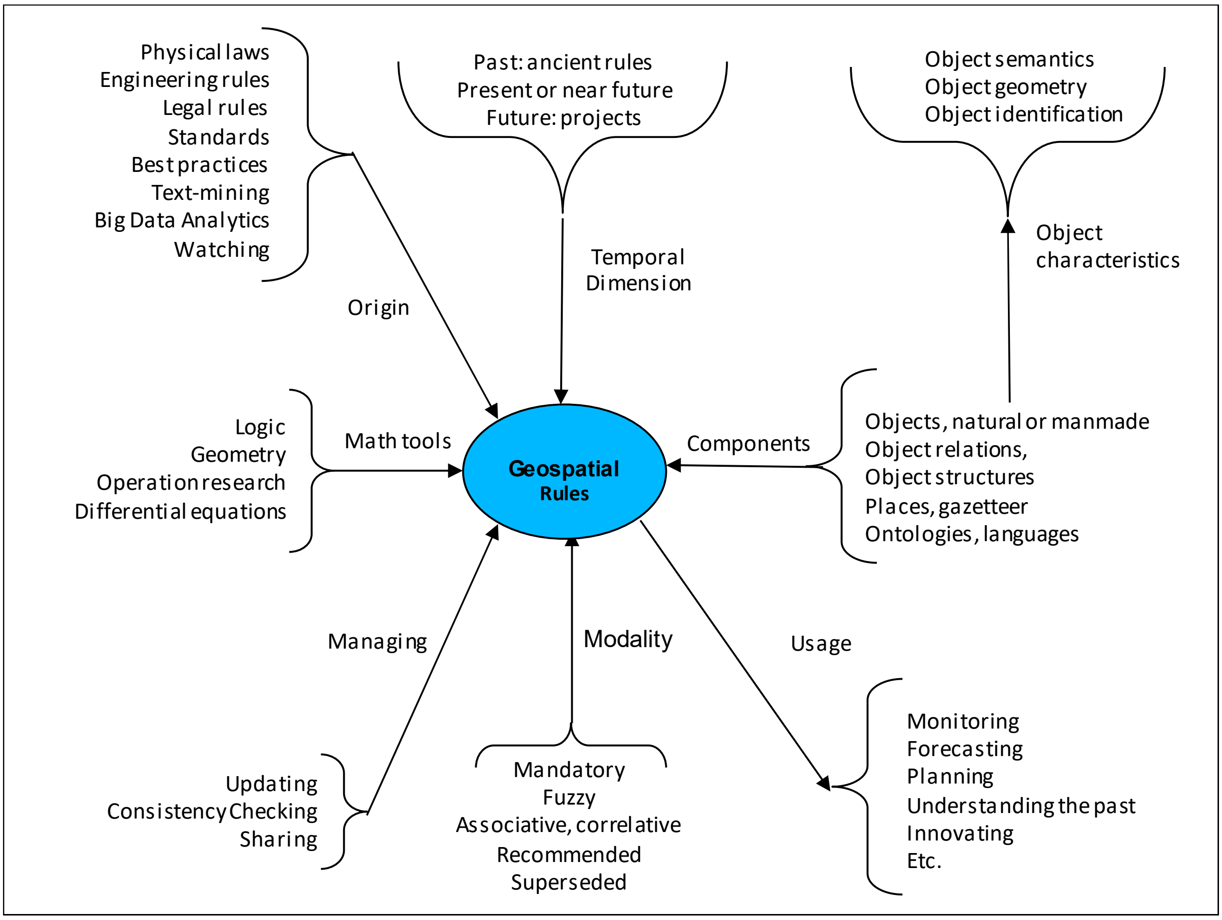

Figure 3. Main characteristics of geospatial rules

How to combine them?

Suppose we deal with many knowledge chunks coming from different origins, some of them being physical, legal, practical, associative or correlative, which can be useful to solve a problem. The big question is "how to combine them?" After eliminating the manual combination, we can face two solutions, the fully automatic, and a hybrid approach.

As far as we know, the solution to combine automatically knowledge chunks for different semantics is not in easy task. If all the chunks are logic-style, Boolean logic is a good background already implemented in system such as Prolog. If we include geometric considerations, the task is a little bit complex. Let's say that more investigations are necessary to envision such an approach

However, by hybrid solutions, we mean solutions based on a man-machine cooperation; and the solution seems possible by some trial and error method by looking for consequences for each knowledge chunks. Let's develop a few cases in the next section.

SPATIAL KNOWLEDGE ON THE MOVE

As previously told, spatial knowledge, and especially geospatial rules can have different modalities and can be used for several key applications. Let us examine few of them.

Knowledge for the Formal Evaluation of Plans

Once a plan is designed, all knowledge chunks are considered to state whether they are followed. A solution is to present all chunks, one by one, against the designed plan. As previously explained, depending upon the variety of chunks, the outcomes can be different.

Concerning permanent physical laws, all the laws must be applied, but concerning incidental physical laws, some priority order may be selected according to some local policies. Those which in the top of the list must be followed. Let's call them mandatory chunks.

Regarding legal chunks, once having listed the exception, they must be followed. In the contrary, the system must inform the planners. They are also considered as mandatory chunks

Regarding best practices, associative and correlative knowledge chunks, it is nice if they are followed, but no direct consequences when they are not followed. Let's call them informative chunks.

As a consequence, three solutions are possible.

- According to the contexts, some chunks are not applicable, for instance rules concerning high mountains in a flat city.
- If all mandatory chunks are followed, one can state that the designed plan is formally validated. If they are not, either the planners must modify the plan accordingly, or be aware of the possible sanctions; the latter could be said negatively followed.
- But regarding informative chunks, it can be interesting to notify the planners when they are followed.

As an output of this formal evaluation, a possibility could to list all chunks with some comments, especially those negatively followed.

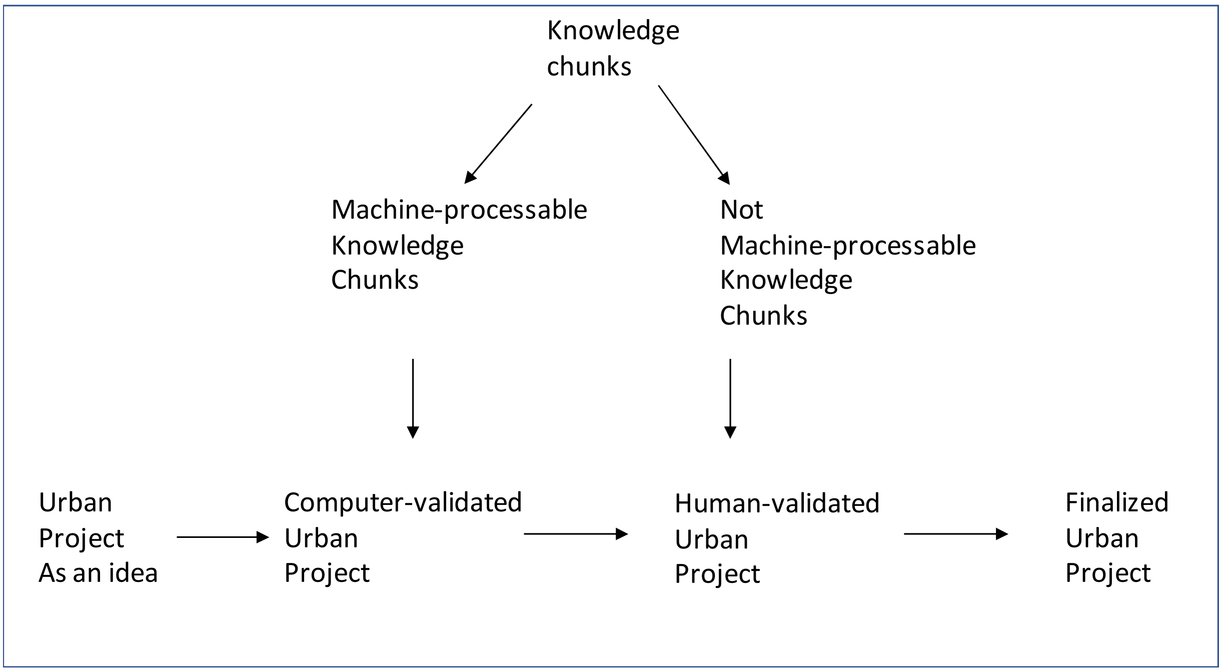

Figure 4. Integrating various knowledge chunks for an urban project

Knowledge for Monitoring

A lot of legal laws and best practices are linked to thresholds, especially for environmental monitoring based on sensors for measuring the quality of air, of water, or the approach of a big storm, and for traffic control. Here the rules can be easily applied, for instance such as the announcement of alarms. But the big problem is that nobody is sure that drivers will follow those announcements. For instance, in traffic management, three types of consequences can be figured out after the announcement of a jam:

- If nobody obeys the announcement, the jam problem will increase;
- If only a good proportion of drivers follows the announcement, the problem will be solved;
- If a very high proportion follows the announcement, the local problem will be solved, but other problems will emerge, such as a new jam in a parallel route.

In the last case, some additional best practices must be designed.

Knowledge for Specific Plan Design

Let consider rapidly some examples.

- A touristic city wants to optimize its resources by using big data. The problem is as follows: if we know tourists' itineraries in the city, we may propose ideal locations for shops, restaurants and hotels. How to solve the problem, or at least to recommend best practices? The first issue is to determine tourist's itineraries. Let's suppose we want to use mobile phone tracks based on GPS data; but how to distinguish tourist mobile phones from local phone numbers? An idea is only to use mobile phone numbers coming from other countries, so neglecting domestic tourists coming from other cities. After this step, we can extract trajectories, look for patterns, and possibly find some hints for international tourists.

- Another example is to find the optimal locations of advertising panels. Similarly, based on GPS data, vehicle trajectories can be discovered.
- Considering free access to Wi-Fi, a city wants to determine the open squares where it can be installed.

CONCLUSIONS

The goal of this paper is to show that big data, coupled with GIS data, can help urban planning, but only by combination with other sources of knowledge with different statuses. As previously mentioned, since often the antecedents and consequents are clear, the meaning of the implication can be various.

Knowledge discovery can be set as the goal of big data analytics; but as soon as a new associative or corelative chunk of knowledge is detected, how to determine its meaning? Is it casual, a spurious correlation, or the premices of the discovery of a new generic hidden law? We mean to find solutions and procedures to identify those new laws, so they can help urban planners in planning the city of the future within the context of the knowledge society.

REFERENCES

Batty, M. (2013.) Big data, smart cities and city planning, *Dialogues in Human Geography*, 3(3) pp. 274–279.

Baxter, R. (1976). *Computer and Statistical Techniques for Planners*; Methuen Press 1976, 336 pages.

Champeau, S. (2015). *Big Data et Smart Cities : perspectives et problèmes*. Can be downloaded from https://champeau.files.wordpress.com/2015/04/big-data-smart-cities.pdf

Golledge, R. (2002). The Nature of Geographic Knowledge. *Annals of the Association of American Geographers*, 92 (1), pp. 1-14.

Graham, I. (2006). *Business Rules Management and Service Oriented Architecture: A Pattern Language*. London, John Wiley.

Gurteen, D. (1998). *Knowledge, Creativity and Innovation*, *Journal of Knowledge Management*, Vol. 2, Issue 1, pp. 5-13.

Kumar, S., Toshniwal, D. (2015). *A data mining framework to analyze road accident data*. Journal of Big Data 2:26. DOI 10.1186/s40537-015-0035-y. Can be download from https://core.ac.uk/download/pdf/81196256.pdf

Laurini, R., Servigne, S., Favetta, F. (2016). An Introduction to Geographic Rule Semantics. 22nd *International Conference on Distributed Multimedia Systems,* DMS 2016, Salerno, Salerno, Italy, November 25-26, 2016, Published by Knowledge Systems Institute, ISBN: 1-891706-40-3, pp. 91-97.

Laurini, R. (2017). *Geographic Knowledge Infrastructure for Territorial Intelligence and Smart Cities*. Wiley-ISTE, Elsevier. April 2017.

Laurini, R. (2018). *Knowledge Society, Territorial Intelligence, Smart Cities, Knowledge Infrastructure*. Paper presented at Cities4People: Smart Cities and Data Analytics, April 18-20, 2018, JADS, s'Hertogenbosch, The Netherlands. In press, edited by K. Kourtit and N. Nijkamp.

Mitchells, R. (2014), *Eight big trends in big data analytics*. Can be accessed from: https://www.computerworld.com/article/2690856/big-data/8-big-trends-in-big-data-analytics.html

Morgan, T. (2008). *Business Rules and Information Systems: Aligning IT with Business Goals*. Addison-Wesley. 384 pages.

Shekhar, S., Huang, Y. (2001). *Discovering Spatial Co-location Patterns: A Summary of Results*, in Proc. of the7th Int'l Symposium on Spatial and Temporal Databases, 2001, pp. 236–256.

Shekhar, S., Zhang, P. (2006). *Spatial Data Mining: Accomplishments and Research Needs*. Can be downloaded from: http://www.spatial.cs.umn.edu/paper_ps/giscience.pdf

Snijders, C., Matzat, U., Reips, U.D. (2012). Big Data: Big Gaps of Knowledge in the Field of Internet Science, *International Journal of Internet Science*, 2012, 7 (1), pp. 1–5, ISSN 1662-5544.

Ross, R. G. (2011). *More on the If-Then Format for Expressing Business Rules: Questions and Answers*, Business Rules Journal, vol. 12, No. 4 (Apr. 2011), URL: http://www.BRCommun 2002ity.com/a2011/b588.html.

KEY TERMS AND DEFINITIONS

Big Data: ****************
External Knowledge: knowledge describing issues located outside the jurisdiction of a local authority.
Gazetteer: a database for place names or toponyms
Geographic Knowledge: to information potentially useful to explain, manage, monitor, understand the past, plan a territory and innovate.
Geographic Knowledge Base: a repository for geographic knowledge.
Geographic Object: computer description of a geographic feature as stored in a geographic data or knowledge base.
Geographic Ontology: a semantic network describing the relationships between types of geographic objects.

Geospatial Rule/Geographic Rule: in information technology, a rule in which geographic places are involved.
Internal Knowledge: knowledge describing issues located inside the jurisdiction of a local authority**.**
Knowledge: in information technology, an information able to be used to solve a problem.
Spurious correlations: a spurious correlation is a mathematical relationship in which two or more events or variables are not causally related to each other, yet it may be wrongly inferred that they are, due to either coincidence or the presence of a certain third, hidden factor.
Urban Knowledge: information potentially useful to explain, manage, monitor, understand the past, plan a city or a metropolis and innovate.