



Christian Rezende Freitas
GIS Specialist and Environmental Analyst with experience in coordination of projects of Geographic Information Systems and SDI (Spatial Data Infrastructure) using Geographic Database and Geographic ETL tools (Extract, Transform and Load). Knowledge in modeling, extraction and validation of spatial data for use in GIS and Business Intelligence tools



Ana Clara Mourão Moura
Graduated in Architecture and Urban Planning, Specialization in Territorial and Urban Planning from PUC-MG and University of Bologna, Master in Geography from UFMG and PhD in Geography from UFRJ. Professor at UFMG, Department of Urban Planning, and coordinates Geoprocessing Laboratory in the School of Architecture.

ETL Tools to Analyze Diagrams' Performance: Favoring Negotiations in Geodesign Workshops

Ferramentas de ETL Para Análise de Desempenho de Diagramas: Favorecendo Negociações em Workshops de Geodesign

Geodesign is a set of methods to shared construction of alternative futures to a territory. Among the existing methods, Steinitz's framework proposes analytical and propositional steps. In analytical step the goal is to compose maps identifying vulnerabilities and attractiveness of the area. In propositional step the users design diagrams of projects and policies to the place to be negotiated between the participants. To analyze the diagrams could help support the negotiations. The study presents the construction of methods, using ETL Tools, to describe the characteristics of the diagrams, mainly the locational and topological similarities. The similarities can be used to identify projects or policies that are equivalent, complementary or conflicting, helping the negotiation process, reducing the time spent in unnecessary discussions.

O Geodesign é um conjunto de métodos para a construção compartilhada de futuros alternativos para o território. Entre os métodos existentes, o framework proposto por Steinitz possui etapas analíticas e proposicionais. Na etapa analítica, o objetivo é compor mapas identificando vulnerabilidades e potencialidades. Na etapa proposicional, os usuários criam diagramas de projetos e políticas a serem negociados entre os participantes. Analisar os diagramas pode ajudar nas negociações. O estudo apresenta métodos, utilizando ferramentas de ETL, para descrever as características dos diagramas, principalmente as semelhanças locacionais e topológicas. Estas podem ser usadas para identificar projetos ou políticas que sejam similares, complementares ou conflitantes, auxiliando a negociação, reduzindo o tempo gasto em discussões desnecessárias.

Keywords:

ETL tools, Geodesign, Co-Design, Negotiation Methods, topological similarity, locational similarity

Palavras-chave:

Ferramentas de ETL, Geodesign, Co-Design, Métodos de negociação, similaridade topológica, similaridade posicional

1. INTRODUCTION

In the last 50 years, the studies in territorial planning are considering as a value and as a requirement the inclusion of citizens' expectations and decisions, in an approach of participatory planning. The sense of participatory or community planning is understood as a collaborative process, shared governance, bottom-up process, based on comprehensive inclusion of people of the place.

The interests in this paradigm started in the 60's, when some authors developed theories about space perception and cognition, in order to understand how people create their relation with the territory, the feelings about space and sense of place, the importance of topophilia in cultural and social identity (Lynch, 1960; Tuan, 1974). At the end of the decade Arnstein (1969) proposed a ladder of citizen participation as a theoretical reference, to encourage the inclusion of different sectors of society in planning, facing the issues of going from nonparticipation to citizens' power.

Participatory planning became the center of interests since then, and nowadays the interests are about methods and possibilities in E-Planning, including digital tools and web-based methods (Smith, 1973; International City Management Association, 1979). E-planning combines the resources of Geographic Information Systems (GIS) with the goals of Planning Support Systems (PSS), using web-based tools to participatory decisions (Hanzi, 2007; Brail and Klosterman, 2001; Saad-Sulonen, 2012). The idea of E-Planning started with the tasks of "Plans-on-the-map" and "Tell-it-on-the-map", providing feedbacks about preferences and expectations from people of the place, using the facilities of geo-questionnaires (Saad-Sulonen, 2012; Czepkiewicz, Jankowski and Młodkowski, 2017).

Geodesign also came in this era of participatory and community planning, with the goal to give support to opinion and decision making, but it's improving performances with the use of digital tools. It's a method to design with and for the geography, that's why it's geo+design. According to Flaxman (2010), Geodesign is a design and planning method which tightly couples the creation of spatial proposals, presenting impact simulations informed by geographic

contexts. Lee, Dias and Scholten (2014) explains it integrates designs and geospatial sciences.

Steinitz (2012) proposed a framework for Geodesign, composed by steps that go from the evaluation of the place to the construction of alternative futures in collective planning. According to Steinitz, the creation of alternative scenarios to a place must take into account 6 models, 3 of them to evaluate the area based on assessment, and 3 of them to propose new designs and interventions. The models of assessment steps are: representation, process and evaluation. The models of intervention steps are change, impact and decision.

In assessment step, the "Representation Model" has the goal to answer "how should the study area be described?". The "Process Model" has the goal to answer "how does the study area operate?". And the third model is the "Evaluation Model", that has the goal to answer "is the current study area working well?". Finishing this part, the organizers and participants may have a clear characterization of vulnerabilities and attractiveness about the place, which are the main issues to be faced.

Having a clear idea about the place and its needs, the participants start the intervention step, that is composed by other 3 models. The "Change Model" has the goal to answer "how might the study area be altered?". The "Impact Model" has the goal to answer "what differences might the changes cause?". Finally, the participants arrive to the last model, the "Decision Model" that has the goal to answer "how should the study area be changed?".

The author proposed to follow the guideline of 6 models in 3 iterations. The first iteration is based on "why are we studying the area?", in order to understand the main characteristics of the study area and to make people construct the first proposals of designs to the place. After finishing the first experiment, it's important to analyze the results and review all of the process, all of the models, in a second iteration, based on "how can we develop the case study?".

After reviewing all the steps to calibrate all the models, it's time for the third and last iteration, that follows again all the six models, with the

goal to define "what, where and when?" and perform the final designs to the study area.

During the steps of the framework, participants receive and produce information about the place, and construct together designs to the territory. Designs are sets of projects and policies they believe are the best ideas to the area. They do dynamics of co-construction of proposals, in consensus-seeking approaches. They draw their own ideas, but also draw new ideas observing the proposals from each other, or even draw new ideas in common sense.

The negotiation steps in Geodesign can be done in analogical mode, using printed maps and overlays. But, in the last years, Steinitz is working with Ballal, who developed a web-based tool to Geodesign, called "Geodesign Hub". The organizers have to produce the assessment models (representation, process and evaluation) to feed the web-platform with systems about the main characteristics of the place, classified according to potentiality of receiving proposals (levels of "feasible", "suitable" or "capable"), of inadequacy of receiving proposal (classified as "inappropriate") or even showing areas where the problem was already solved and it's no use to construct proposals there (classified as "existing"). (Figure 1).

It's also important to define a table of impacts, according to conflicts or suitability, in combinatorial evaluation between the systems. A diagram of project or policy can have parts colored from deep purple to deep orange, from most positive combination to most negative combination, highlighting territorial conflicts of interests (Figure 2). It's based on this matrix that the proposals of projects or policies in each thematic will be evaluated, classifying the combinations from +2 (most positive impact) to -2 (most negative impact) (Figure 3).

The participants draw diagrams per system, presenting ideas of projects and policies to solve problems or to develop potentialities. The participant can use as base the evaluation map of the system, a combination of some evaluation maps, or even base maps with spatial references. From the beginning of the workshop until the very last moment of constructing the final proposal, new diagrams can be drawn, as part of the process negotiation and exchange of ideas. (Figure 4).

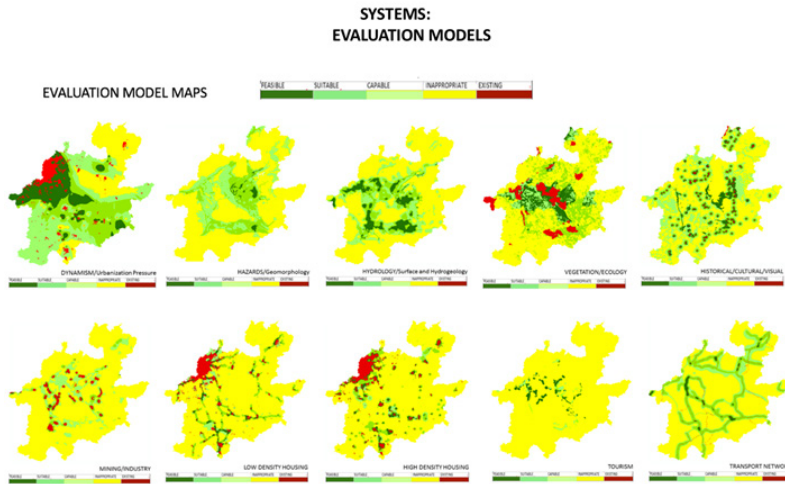


Figure 1 - Example of evaluation models organized in 10 systems that are used as references to draw diagrams about projects and policies. "Iron Quadrangle" case study, Brazil.

	A	B	C	D	E	F	G	H	I	J	K	L	M
				HCULV	ECO	HYDRO	GRISK	IND	UGRT	LDH	HDH	TRSM	TRANS
System #	SYSTEM CHANGE												
1	HCULV			N/A	1	0	1	-2	-1	-1	-1	2	-1
2	ECO			1	N/A	2	2	-2	-2	-2	-2	1	-1
3	HYDRO			0	2	N/A	2	-2	-1	0	-1	1	0
4	GRISK			0	2	2	N/A	-1	-2	-1	-2	0	-1
5	IND			-1	-2	-2	-1	N/A	0	0	0	-2	0
6	UGRT			-1	-1	-1	-1	1	N/A	1	1	0	2
7	LDH			-1	-1	-1	-2	1	1	N/A	1	0	1
8	HDH			-2	-2	-2	-2	0	2	-1	N/A	-2	2
9	TRSM			2	1	1	0	-1	1	0	0	N/A	1
10	TRANS			0	-2	-2	-1	2	2	1	2	0	N/A

RULE—if more than one polygon is in one location, use the most positive value. E.g high density housing

IMPACT MODEL

- 2 most positive, best
- 1 positive, good
- 0 neutral
- 1 negative, bad
- 2 most negative, worst

Figure 3 – Example of Impact Analysis Matrix. In the example, if the participant proposes a project of "ECO" (interests of ecology and to create protected areas) in an area that is important for "HYDRO" (to protect hydrological values), the impact is +2, because it's most positive, the best. But that project or policy can create problems in areas that are important to industry, high density or low density housing or areas of growth and dynamism ("IND", "LDH", "HDH", "UGRT"), and it will be considered -2, the most negative or worst.

The diagrams are named according to the main idea they represent, they have their area and perimeter dimensions calculated, they can receive hashtags to facilitate searches and organization of ontologies, they are colored according to systems impacts and it's also possible to analyze the cost and the achievement of targets' values proposed to the case study. (Figure 5).

Once elaborated a set of diagrams, to different themes, the participants are organized in teams that represent different sectors of society. In these teams they select those diagrams that fit best their ideas to the place, and they compose "Designs". Designs are compositions of projects and policies to the place, considering all the systems, and representing the main values and expectations according to the team. (Figure 6).

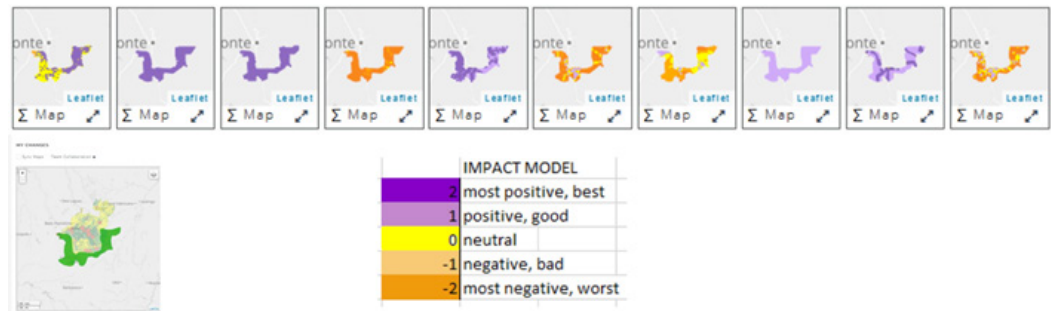


Figure 2 – Example of a diagram from the system "ECO" confronted with all the systems. Parts of it have very positive impact in some systems, parts of it have very negative impacts in other systems, but there are also neutral or not so positive or not so negative impacts. "Iron Quadrangle" case study, Brazil.

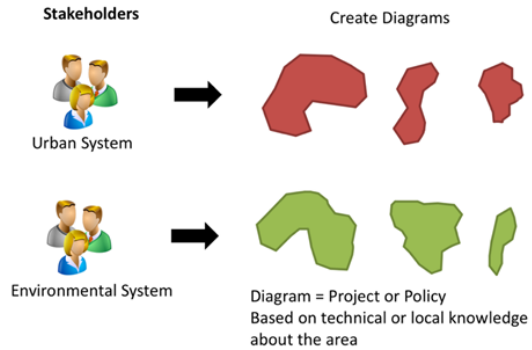


Figure 4 – The proposal of diagrams to each system.

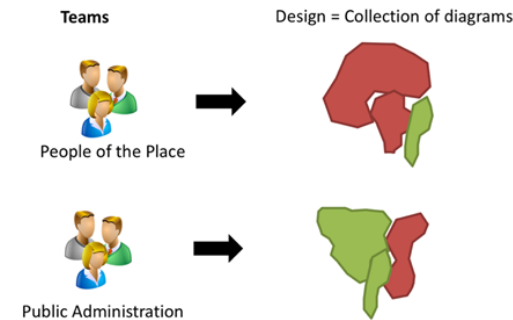


Figure 6 – The composition of designs based on sets of diagrams.

The comparison of designs proposed by the teams is a very important action during the workshop, because it tells how differently the groups think. At the end, the workshop will result in one final diagram, produced by negotiation, and it's important to recognize the similarities and conflicts to be faced. (Figure 7, 8).

The comparison of proposals helps the definition of a common solution, or to negotiate an acceptable solution, or even to refute ideas that are under conflicts of interests (Figure 9). During some workshops experiences we recognized some situations in which the teams arrived to a common proposal more easily, and some others in which a final design was almost impossible. But in some cases, they were against each other without noticing they were talking about the same thing, but believing they were on opposite sides.

<http://disegnarecon.univaq.it>

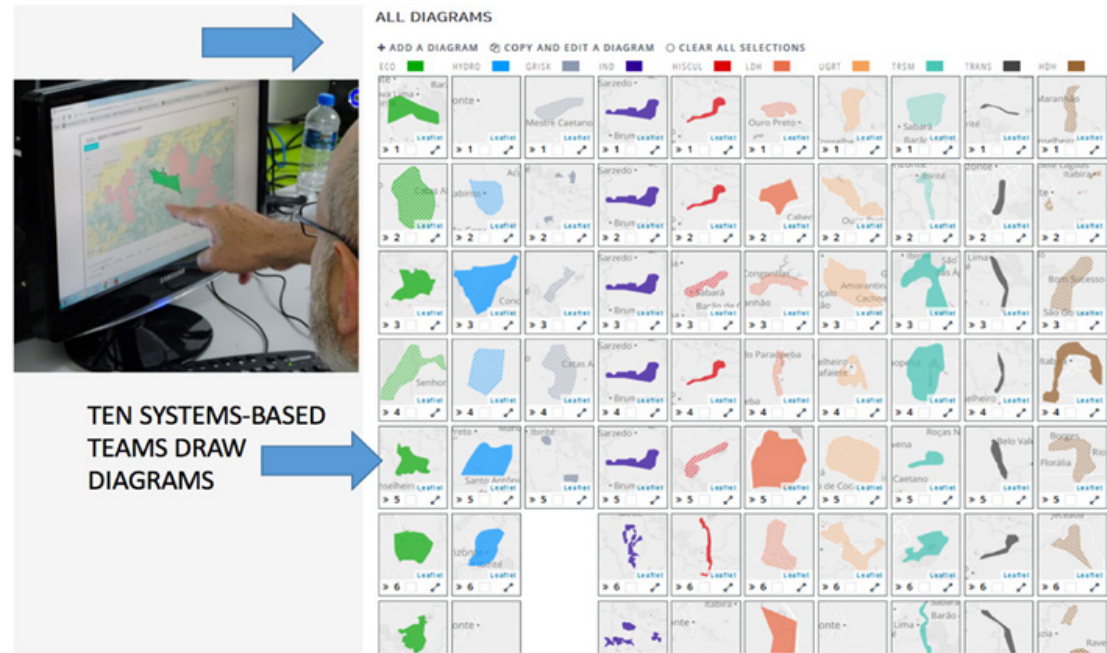


Figure 5 – The list of diagrams drew per system. In hatches colors the diagrams represent policies, and in colorful areas the diagrams represent projects. "Iron Quadrangle" case study, Brazil.



Figure 7 – The organization in teams to construct proposals of designs, that are sets of diagrams about policies and projects. "Iron Quadrangle" case study, Brazil.

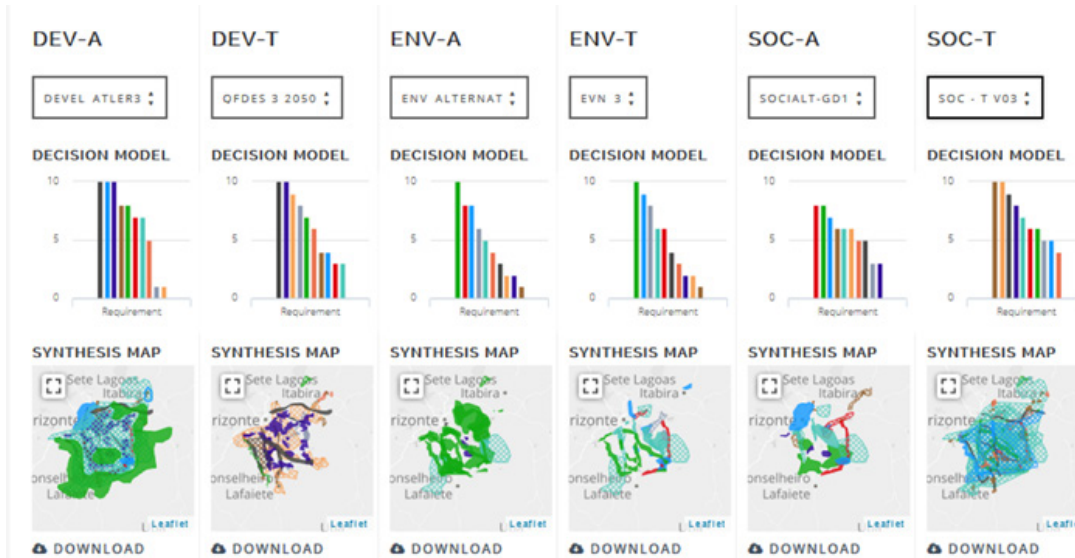


Figure 8 – Composition of designs. Example of 6 different designs, with sets of diagrams that better suited the expectation of each team. “Iron Quadrangle” case study, Brazil.

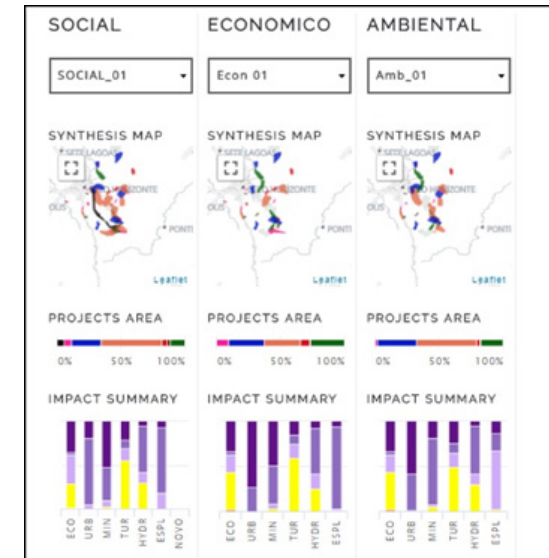


Figure 10 – Similar designs composed by different teams. Example of “Geology as the basis for planning alternative futures for the Iron Quadrangle”, developed by Casagrande (2018).

Two case studies made us think about the importance of understanding how similar or conflicted the designs can be: the case study of “Maria Tereza illegal settlement”, in Belo Horizonte, and the case study of “Geology as the basis for planning alternative futures for the Iron Quadrangle”, in Minas Gerais, both in Brazil. The first case study was developed under a collaboration of Geoprocessing Laboratory from the School of Architecture of Federal University of Minas Gerais with the municipality of Belo Horizonte (Araújo et al, 2018; Zyngier et al, 2017), and the second case study was a Master Dissertation of a student from Geoprocessing Laboratory (Casagrande, 2018).

The dissertation about geology studies as the basis for Geodesign in Iron Quadrangle area made us understand that, when you are working with teams that have the same social insertion and similar professional qualification (in the case study, all from Earth and Environmental Sciences), the proposals tend

to be quite similar, and in just few negotiation rounds it's possible to arrive to a final design. (Figure 10).

The case study of Maria Tereza illegal settlement was very important to understand the role of recognizing similarities in different designs. The municipality of Belo Horizonte invited the group from Geoprocessing Laboratory to test Geodesign method in a PRU (Plano de Regularização Urbana – Urban Regularization Plan) of an urban illegal settlement. As it was the first time the municipality was using the method, they decided to restrict the participation to public administration technicians, so that they could decide about the suitability of applying it in a real case study. They had to receive opinions from many technicians, and they were numerous, 70 participants. The plan was to use Geodesign Hub web-platform, but it was not possible to put everybody in the same room. The solution was to divide the group in two, and to develop to independent workshops, arriving to two final designs. (Figure 11).

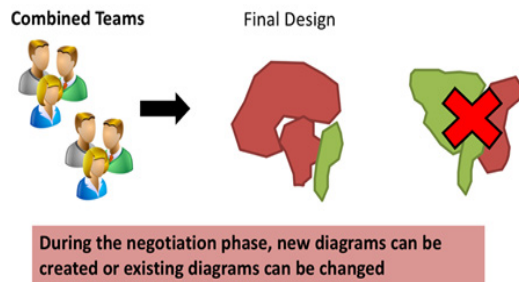


Figure 9 – The negotiation of the final design.



Figure 11 – Workshops from Maria Tereza case study, Belo Horizonte, Brazil, 2016.

The two groups did not have contact with each other, because they worked on two different days, but the final proposals were very similar. It can be said, therefore, that when the proposal is elaborated by social agents that are from similar context, the final proposals represent defensible and reproducible criteria. It's not difficult to arrive to a defensible collective proposal, based on main collective ideas and values.

At that time, we used visual inspection to recognize similarities among final designs, putting the diagrams side by side, by system, and comparing the proposals. In most systems the drawings were different, but the ideas were quite similar, and it was also possible to identify proposals with the same ideas but in different positions in the area. We understood that if we had a tool to analyze the diagrams, a feedback about similarities could be presented to the participants, during the workshop or in tables of discussions between the meetings. It was possible to see that sometimes they were against each other without recognizing they were talking about the same thing, as their ideas were quite similar. (Figure 12).

The interest in comparing diagrams was first planned to be applied in Geodesign, to find similarities in projects and policies, but it can also be useful in many other applications, whenever is important to identify how similar georeferenced polygons are. (Figure 13).



Figure 12 – Example of comparison of diagrams about projects and policies in two systems, in Maria Tereza case study, Belo Horizonte, Brazil.

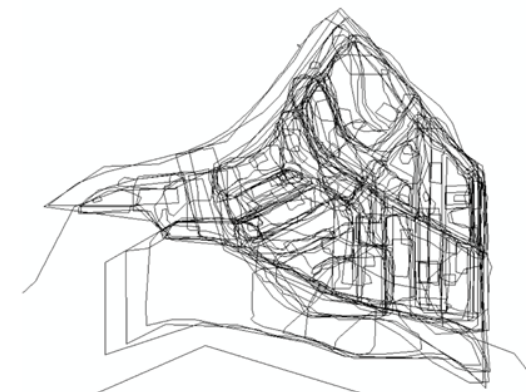


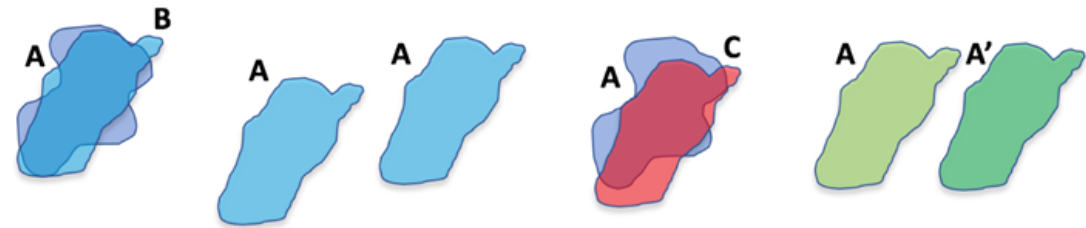
Figure 13 – How can we compare the similarity of diagrams? In the picture, example of Maria Tereza case study, presenting the collection of diagrams proposed.

2. THE PROPOSAL: ANALYZING THE SENSE OF SIMILARITY

Geodesign is a method to support decision making in the co-creation of alternative futures to a territory. The proposals are sets of diagrams representing projects and policies, so that the main decisions are about place (where?), dimension (how much?), main idea (what?) and possibilities (which?). In that sense, we were interested in developing studies to identify affinities or conflicts between diagrams: topologic similarity, positional similarity, conceptual similarity and taxonomy similarity. (Figure 14). The “conceptual” similarity is, somehow, already controlled in Geodesign Hub, as the coordinator of Geodesign is requested to input the impact matrix data in the system (Figure 3). When deciding about positive, negative or neutral impacts, the coordinator is defining the references that are going to be used to classify each diagram, colored from deep purple to deep orange (Figure 2). The results make the participants identify the conceptual conflicts or agreements, based on different ideologies or values expected to the same place. Our studies, to collaborate on this topic, are based on tools to share decisions about judgments and evaluations. In some case studies we conducted, the participants questioned about the defined values, and said they wanted to be consulted about their opinion. Since then, we are applying Delphi Method to decide about values in the matrix.

Delphi Method is based on the construction of consensus by approximations (LINSTONE and TUOFF, 1975). The participants answer to questionnaires in two rounds. In the first round they present their opinion about the impact from one system on the other, and the majority and media of opinions are calculated and presented to them. Comparing the media of opinions to their own opinion, they answer to the questionnaire in the second round, with the possibility of keeping or changing their votes, and a new media is calculated. This media is used in the final matrix. We tested digital web-based tools to do it: Survey Monkey and Google Forms.

The “taxonomy” similarity has the goal to understand if people use different words to the same idea. It's also interesting to connect similar ideas to each other,



Topologic similarity – different geometry and similar ideas in the same place

Geography – positional similarity, same ideas to different places in geographical proximity

Conceptual – different ideologies and values for the same place

Taxonomy – based on semiotics' studies, different words for the same ideas.

Figure 14 – The studies about similarities.

to compose ontologies, grouping families of ideas.

Geodesign Hub is already preparing possibilities to do these studies, because when creating the polygons, the authors can add hashtags to them. In some case studies we conducted (Faenza, in Italy, and Pampulha, in Brazil, both in 2017) we tested applications to collect previous ideas from the participants, before the workshop, to compose a list of possibilities, thinking about families of ideas and ontologies.

In Faenza case study we used crowdsourcing mapping to receive opinions about the place. In Pampulha and in Faenza case study we composed platforms of WebGis with the main variables of the place, and asked the participants to navigate on it before the workshop, to recognize vulnerabilities and attractiveness of the territory. After analyzing the WebGis, they were asked to answer to web-questionnaires, presenting keywords

of ideas to face the necessities of each system.

It was similar to the method of geoquestionnaires. During the workshop meeting, the list of keywords with ideas was shared with the participants. Future studies can be about the composition of ontologies of main words and ideas.

Our main studies were concentrated in topological and locational similarities, to which we constructed scripts to automatically calculate the relations among diagrams. The study of topological similarity was based on the intersection of diagrams: if the area of intersection is bigger than 80% of the total area of both diagrams, they are considered similar.

To support more complex analysis, it is also important to calculate if one diagram is in within the other, or if it contains the other. (Figure 15).

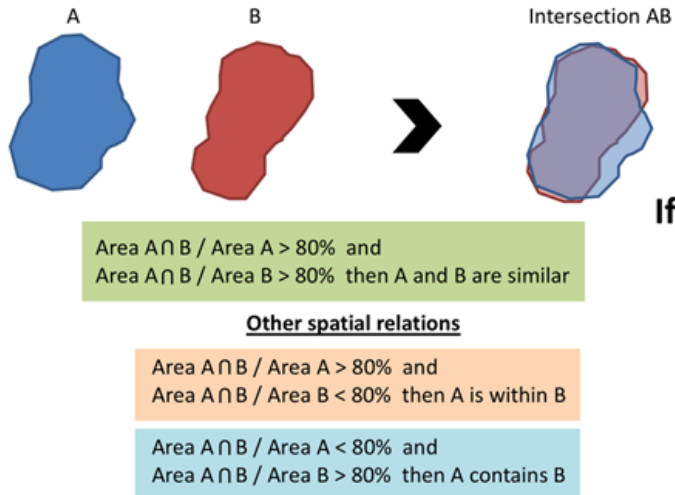


Figure 15 - Topological similarities and spatial relations about intersections of diagrams.

The studies about locational similarity, based on distances, was more difficult to define, because the sense of “close” or “far” depends on references. The sense of “close” is not the same of intersection or touching, because the forms of the diagrams can interfere in the positioning relation (Figure 16). We decided for the definition of a bounding box, based on the extend of the total area of the case study (Figure 17).

Once defined the bounding box, the smallest side of it is measured, and the centroids of the diagrams are identified. Two diagrams are considered close to each other if the distance between the centroids is less than 12,5% of the smallest side of the bounding box (Figure 18). This value is experimental, and requires more studies. Other studies that are under development are about the use regular distribution of points inside the diagrams and the use of tessellation to establish a network of neighborhoods, and in instead of measuring just the centroids, to measure the distance between all the points from different diagrams, resulting in the ranking of vicinities.

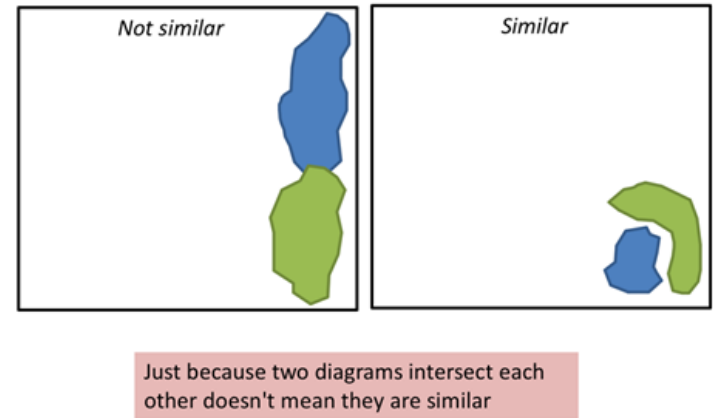


Figure 16 - Intersections or vicinities and not enough to say that diagrams have positioning similarity.

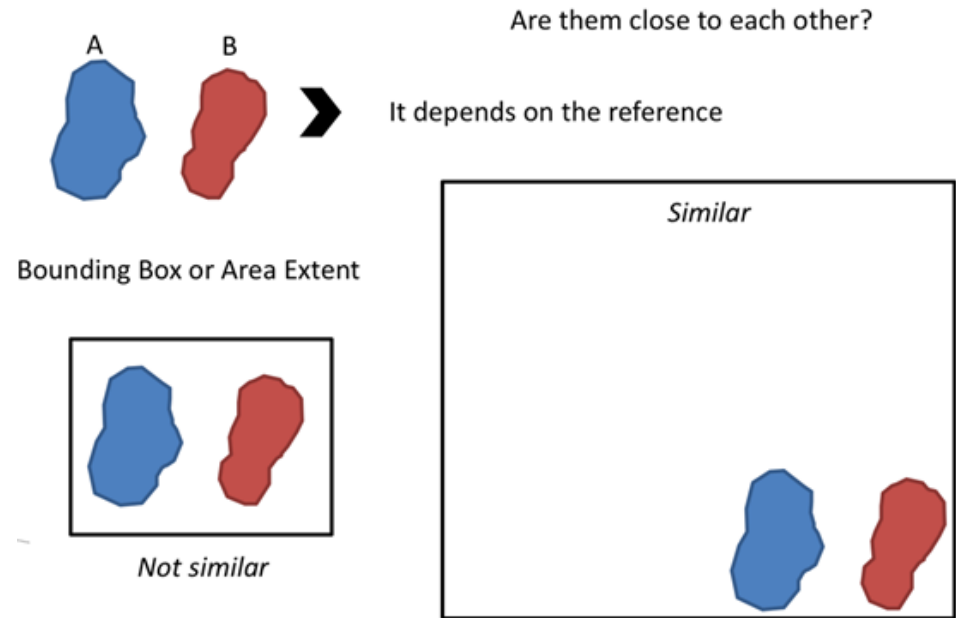
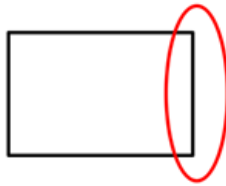


Figure 17 - The definition of a bounding box to present a reference to classify diagrams that are considered close to each other.

1 – Identify the smallest side of bounding box



2 – Use centroid of diagram to calculate the distance



Figure 18 - Measuring locational similarity.

Two diagrams are similar if the distance between its centroids is less than 12,5% of the smallest side of the bounding box

3. ETL TOOLS AND PROGRAMMING TO MODELING

The technological support selected to be used in the tasks were ETL - Extract, Transform and Load tools, in which were created scripts based on spatial models of topological and locational similarity analysis.

3.1. ETL TOOLS

ETL tools area generally associated to the use in database, in procedures of integration and validation of data, when is necessary to process big amount of information and to create routines to be applied without the supervision of the technicians.

The tasks of treatment, conversion and manipulate information can also be challenges to database tools, as the logical processing is based on data storage and rescue of information that are already structured. In this sense, these tools emerged as auxiliary devices that provides broad set of functionalities and algorithms to manipulate data.

Beyond the functionalities, ETL tools can work using

<http://disegnarecon.univaq.it>

client-server architecture, that can be used as part of the solution in data integration, according to existing technological components. Nowadays, these tools are used in the construction of Data Warehouse, non-structured database, directed towards final consumption of information, as BI – Business Intelligence. It requires the organization of information more adapted to experts use, rather than typical logical modeling of the information system for data storage.

In the last years, ETL tools started to be able to read and process spatial data, following the evolution that happened in database in the last decades. They are providing a set of spatial functionalities that, in some cases, can replace traditional GIS (Geographic Information Systems) tools. The main profit in the use of ETL is data interoperability, as they are able to ready and storage data from many different formats, in classical dataset or in geographic dataset.

In the presented case study it was applied a tool package of spatial ETL, the GeoKettle®, version 2.5, based on suite Pentaho Data Integration. (Figure 19).

The selection of the application took into account the characteristics:

- To be compatible with geospatial data and do transformations that are able to use topological concepts in spatial analysis;
- To be open source;
- To be scalable, working in distributed tasks, so that different requests can be started at the same time, using strategies of cloud computation or dedicated computer servers;
- To present better integration to Spatial Data Infrastructure (SDI) by service oriented;
- To present a friendly graphic interface that allows the use and changes in models without the need of specialized knowledge;
- The processes developed are stored in files and can be executed at any time, many times, and be easily distributed to the users.

3.2. SPATIAL MODELS

Since the tools were selected, the task was the composition of scripts, to compose models about topological and locational similarity analysis.

3.2.1 – TOPOLOGICAL SIMILARITY

The description of the model is presented in steps, to make easier the understanding of the whole process (Figure 20). The diagrams were downloaded from Geodesign Hub (because the script was developed to compare existing diagrams constructed during a Geodesign workshop) and converted to shapefile, so that GIS tools could be used.

The first step of the model is the reading of data, identifying and correcting existing possible topological problems in diagrams database, and generating a combining matrix of existing diagrams. If topological errors that come with the diagrams are not solved, some procedures are not executed, as they are based on the intersection of polygons. Geokettle tools are sensible to topological problems, what can cause the interruption of the analysis. (Figure 21).



Figure 19 - Conceptual design of geographic ETL Geokettle (©).

To solve topological problems, some applications are used to separate the diagrams that present errors, in which small buffers are constructed correcting the drawings without changing them so much (steps of Filter rows 2, error and Spatial Analysis).

Once corrected the problems, the dataset is duplicated in two groups, using the step “cross join”, to allow the construction of a cross combination matrix. It’s important to avoid combining a diagram with himself, and also to repeat combinations. To avoid the repetition of combinations (A with B is the same as B with A), subtractions of identifications numbers are done. As results, a set of data will be positive (id2 - id1) and another will be negative (id1 - id2), allowing the separation of them using the tool “filter combinations”.

The second step of the analysis is the spatial intersection of polygons, the calculation of proportions of resulting areas, and also the calculation of proportions of these areas of intersection. (Figure 22). It’s important to highlight that problematic shapes of polygons can result from diagrams’ intersections. To solve this, the tool can produce geometries called “collection” that are the conversion of very narrow or very small polygons to points or lines. After that it’s possible to eliminate the points or lines from the analysis.

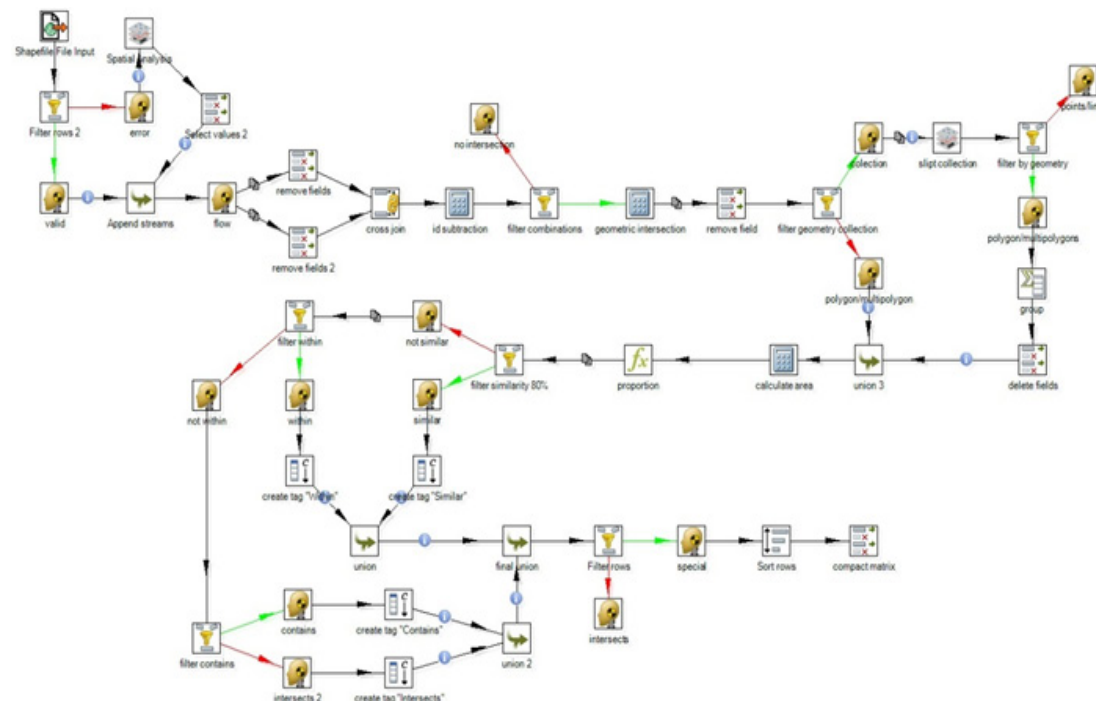


Figure 20 - Topological Similarity Model – global plan.

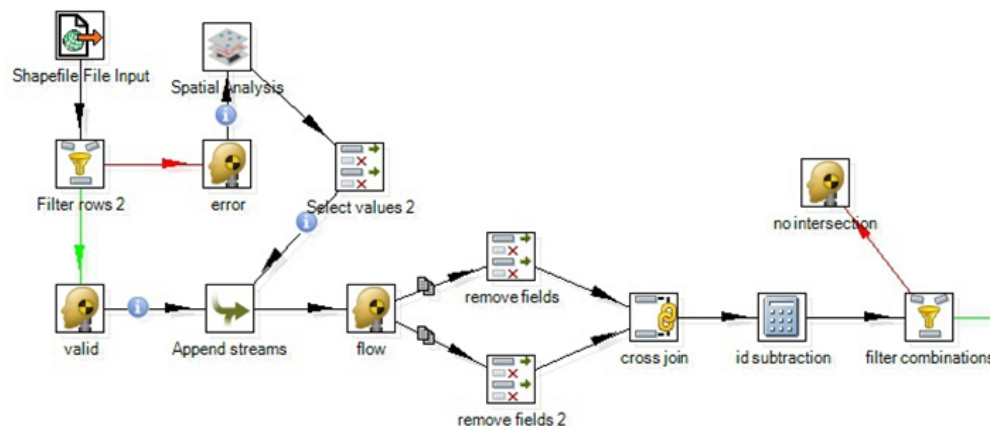


Figure 21 - First step of topological similarity analysis.

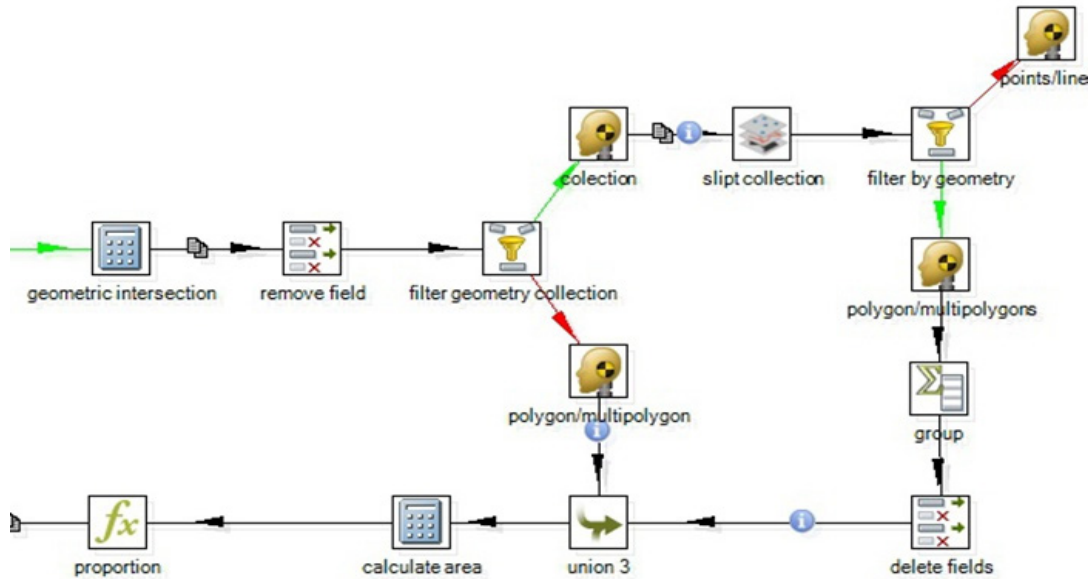


Figure 22 - Second step of topological similarity analysis.

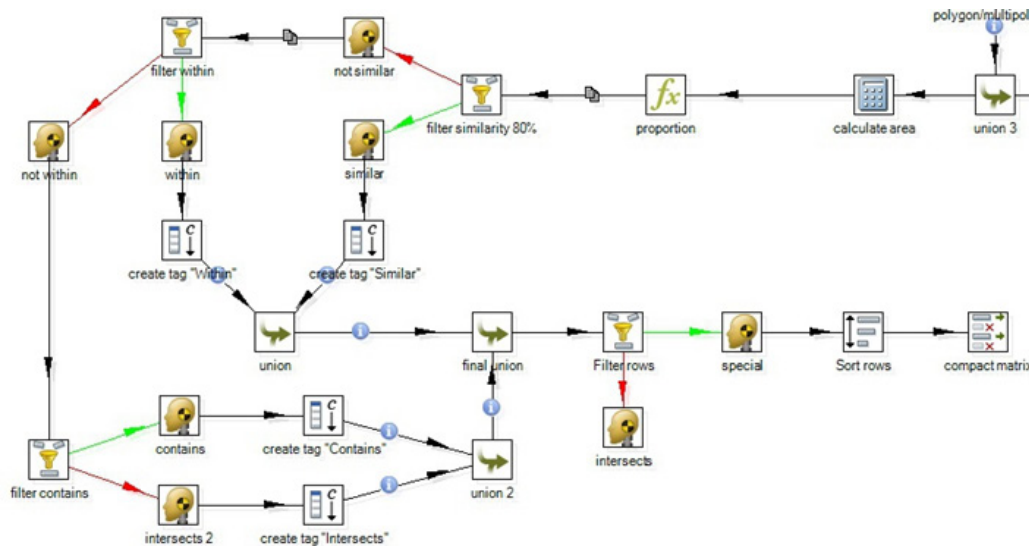


Figure 23 - Third step of topological similarity analysis.

The third and last step begins with the identification of the main relations observed from diagrams combinations (Figure 23). Diagrams are considered similar only if the area of intersection in both of them (area in A from the intersection with B, and also area in B from the intersection with A) is bigger than 80%.

The script calculates not only the level of similarity, but also spatial relations of contains and is contained (is within). The relation is recognized as “contain” when the area of intersection in the first diagram is less than 80% and in the second diagram is more than 80%, while it’s considered “within” when the area of intersection in the first diagram is more than 80% and in the second is less than 80%. As a result, it’s constructed a final matrix with all the existing combinations. (Figure 24).

Examine preview data

Standard view

Rows of step: compact matrix (832 rows)

#	id1	sys1	prop1	relacao	id2	sys2	prop2
1	4038	OLDH	64,3	Contains	4052	RISKS	92,7
2	4038	OLDH	96,6	Within	4084	HDH	5,2
3	4038	OLDH	96,9	Within	4111	HDH	5,1
4	4040	REC	100	Within	4041	SANI	1,2
5	4040	REC	100	Within	4043	ENV	1,7
6	4040	REC	100	Within	4045	TRANS	0,9
7	4040	REC	100	Within	4046	SANI	1,6
8	4040	REC	84,4	Within	4057	OLDH	1,8
9	4040	REC	100	Within	4058	ENV	4,6
10	4040	REC	100	Within	4064	SANI	1,6
11	4040	REC	100	Within	4072	SANI	1,2
12	4040	REC	100	Within	4084	HDH	0,9
13	4040	REC	91,6	Within	4105	OLDH	2,8
14	4040	REC	100	Within	4106	ENV	1,3

Figure 24 - Matrix of similarity relations between diagrams.

3.2.2 – LOCATIONAL SIMILARITY

The description of the model is also presented in steps, to make easier the understanding of the whole process (Figure 25).

The first step of locational similarity analysis is the construction of cross combination matrix of diagrams. It was applied the same processes of avoiding the diagram to be combined with himself and to avoid the repetition of combinations (considering A with B and eliminating B with A).

The second steps starts from the selection of a dimensional reference, defined by the bounding box that enfolds all the existing diagrams. It can be constructed in parallel to the first step (Figure 26). Due to specific requirements from Geokettle tools, the bounding box is converted to lines, and also from lines to points. Using the points, that represent vertices of the box, the tool calculates the distances in the faces of the boundary, to identify the smallest one.

The third step constructs a centroid to each diagram and calculates the distance between all the centroids (Figure 27).

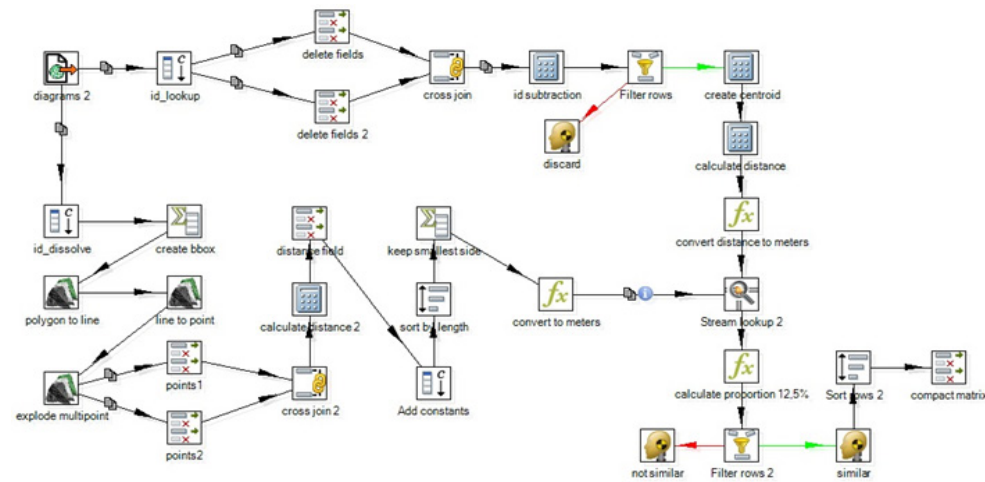


Figure 25 - Locational Similarity Model – global plan.

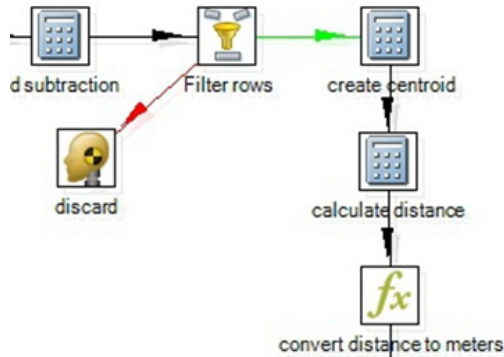


Figure 27 - Third step of locational similarity analysis.

In the fourth and last step, all the distances from the centroids are compared to the value of the smallest face of the boundary (Figure 28). It was defined that distances between centroids cannot be 12,5% bigger than the smallest side of the boundary. The tools applied are “calculate proportion” and “filter rows”.

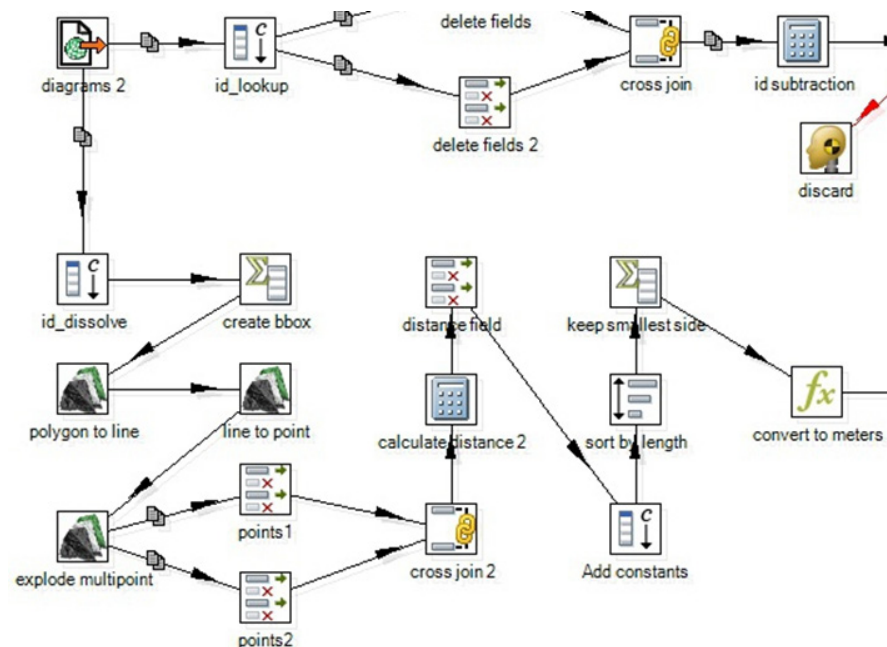


Figure 26 - Second step of locational similarity analysis.

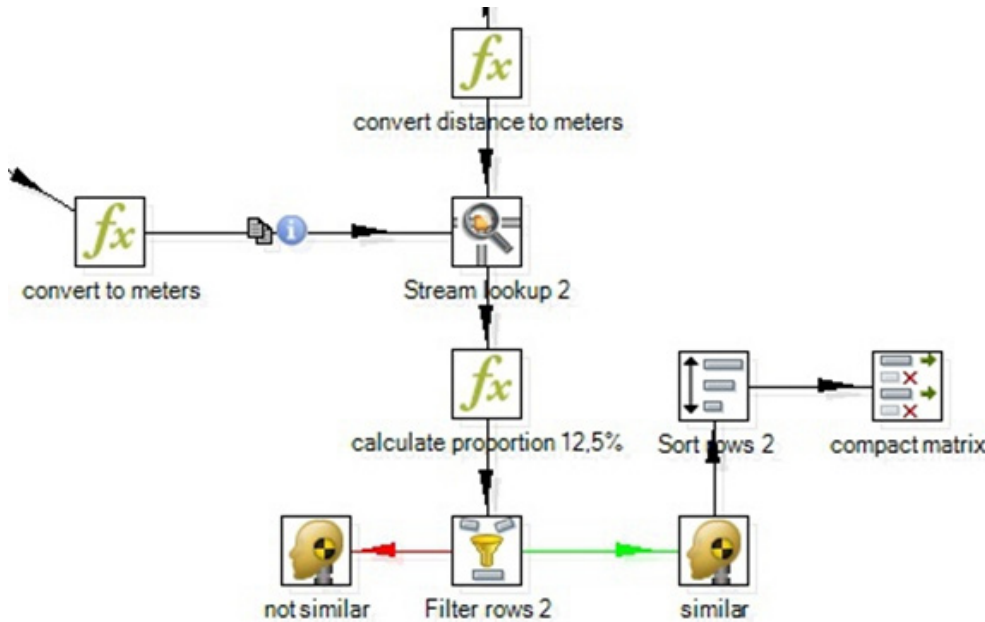


Figure 28 - Fourth step of locational similarity analysis.

As a result, a matrix of locational similarity is produced. (Figure 29).

#	id1	sys1	id2	sys2	smallest_side	distcentroid
1	4038	OLDH	4047	HDH	1047,6	111,2
2	4038	OLDH	4052	RISKS	1047,6	19,4
3	4038	OLDH	4055	PUBS	1047,6	102,3
4	4038	OLDH	4061	HDH	1047,6	128,2
5	4038	OLDH	4065	ENV	1047,6	130
6	4038	OLDH	4067	HDH	1047,6	78,2
7	4038	OLDH	4075	PUBS	1047,6	102,2
8	4038	OLDH	4090	REC	1047,6	90,8
9	4038	OLDH	4104	RISKS	1047,6	126,6
10	4038	OLDH	4115	RISKS	1047,6	110
11	4038	OLDH	4118	COM	1047,6	46,6
12	4038	OLDH	4126	COM	1047,6	37
13	4038	OLDH	4128	HDH	1047,6	98,5

Figure 29 - Matrix of locational similarity.

5. RESULTS AND DISCUSSIONS

Many tests applying these ETL scripts were done, using diagrams from different workshops, and the first results were considered very positive. To topological similarity script, based on the value of 80% of intersection, was considered acceptable, and it was also possible to clearly identify the reports of “contains” and “is within” (Figure 30).

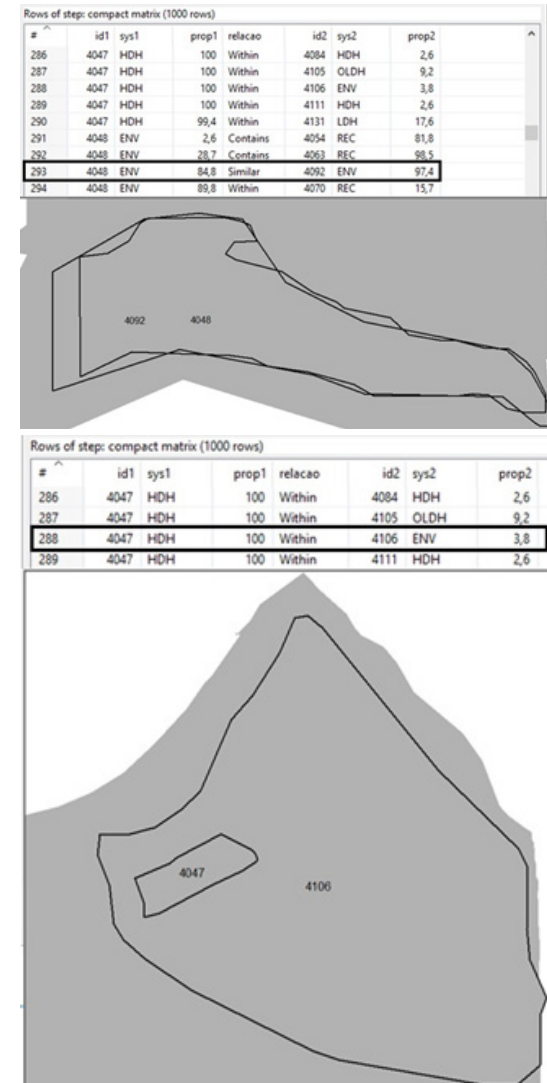


Figure 30 - Example of topological similarity over 80% and the “in within” relation.

The locational similarity script was also able to easily identify the diagrams that were close to each other. The locational similarity can be calculated between diagrams from the same system, but also using diagrams from systems that have conflicts between them. In Geodesign Hub, the application recognizes and classifies conflicts only when the diagrams are overlapped, but the new ETL script constructed can broaden the analysis that identifies neighborhoods' conflicts. (Figure 31).

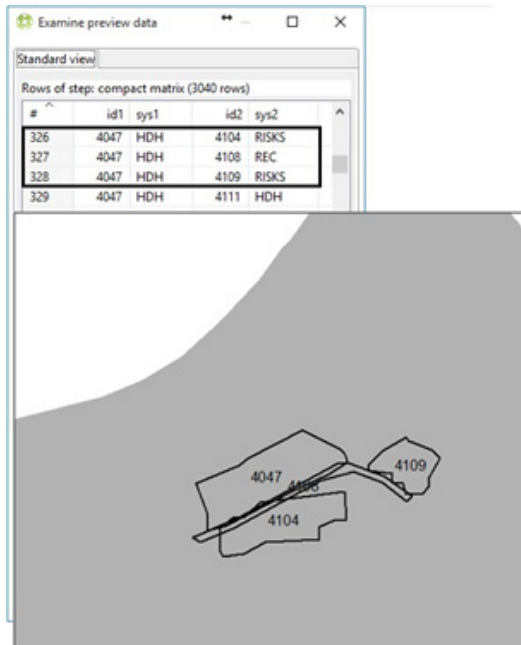


Figure 31 - Example of locational similarity and possible conflicts.

More complex relations between diagrams and systems can be identified and analyzed. Topological similarities between diagrams from different systems can be identified, and also the relations of "contains" and "is within", providing new analysis. In the example (Figure 32), in row 288, a diagram of high density housing (HDH) was identified as similar to environmental diagram (ENV) and situated inside in (within), what can be a conflict or a problem to be solved by planning. In the same table, in row 291, a

diagram from environmental system (ENV) contains a diagram from recreational system (REC), that can be considered complementary, giving support to planning decisions about the area. Also in the same table, row 293, there is a diagram from environmental system (ENV) that is similar to another from the same system (ENV), resulting in redundant proposals. (Figure 32).

#	id1	sys1	prop1	relacao	id2	sys2
286	4047	HDH	100	Within	4084	HDH
287	4047	HDH	100	Within	4105	OLDH
288	4047	HDH	100	Within	4106	ENV
289	4047	HDH	100	Within	4111	HDH
290	4047	HDH	99,4	Within	4131	LDH
291	4048	ENV	2,6	Contains	4054	REC
292	4048	ENV	28,7	Contains	4063	REC
293	4048	ENV	84,8	Similar	4092	ENV
294	4048	ENV	89,8	Within	4070	REC
295	4049	OLDH	10,1	Contains	4083	REC
296	4049	OLDH	60,1	Contains	4137	RISKS
297	4049	OLDH	100	Within	4041	SANI
298	4049	OLDH	100	Within	4045	TRANS
299	4049	OLDH	100	Within	4046	SANI
300	4049	OLDH	100	Within	4064	SANI
301	4049	OLDH	100	Within	4072	SANI
302	4049	OLDH	100	Within	4084	HDH
303	4049	OLDH	100	Within	4106	ENV

Figure 32 - Analysis of topological similarity and systems.

The results indicate the great potential in the exploration of similarity studies, that can be even more interesting if considering other characteristics of the diagrams. It's important to highlight that spatial analysis and spatial relations must be understood as a multidimensional study, combining different parameters.

6. CONCLUSIONS

The ETL models developed aimed to use simple analysis, that could be implemented even using other tools, according to users' preferences. The speed of execution was more important than the complexity of possibilities that could be employed. It's important to highlight that, using the example of 102 complex

diagrams from 10 systems, the script took less than 3 seconds to calculate the topological similarity and less than 1 second to calculate the locational similarity.

The cross combination matrix was able to easily identify more than 10.000 different combinations. The simplicity of the methods employed therefore is a great advantage in the diffusion of the studies.

The challenges observed during the process were related to the topological problems that may exist in the primary data source and that need to be addressed before performing the analyzes.

Two possible problems were identified in the studies about locational similarity. The first one is related to very big diagrams, that cover much of the studying area, because they will be considered similar to the majority of the diagrams. These very big diagrams should be treated differently, because it's quite strange that someone proposes something to cover all the territory. But it happens. They are generally related to policies that are applied to the entire area, or that could be reviewed and applied specifically to more suitable places. It also happens when very extremist proposals are constructed by participants that are not willing to negotiate with other ideas or values, and they demarcate the whole territory.

The second problem has to do with diagrams that are created as multipolygons or multipart. In this case, if the many parts are distributed in the territory, the set of diagrams have the same problem of a big diagram, because it will be considered close to everything, understood as one single proposal. During the workshop, it's generally explained to participants to avoid creating diagrams composed by more than one polygons, and to prefer to save different diagrams according to some classification or logic. The reason is due to negotiation process, because others can accept part of your proposal, but not all of the polygons designed. But it happens that someone prefers to draw a set of polygons in the same proposal, what requires previous filters, as the problems of very big polygons.

It would also be very interesting to adapt the tool so that the user could decide about values of parameters. For example, in topological similarity, instead of fixing the value in 80%, the user could be

able to define it, according to his knowledge about the case study. The same could happen in case of locational similarity and the defined value of 12,5%.

The validation of the methods developed in the studies opens new possibilities to investigate the relations in co-design methods. The recognition of similarities can be used in negotiation steps, in the construction of consensus or in the definition of priorities to be discussed.

As mentioned in the discussions and conclusion of the article, the objective of testing the described method of analysis was, at first, to identify the possibility of using spatial analyzes to improve the results obtained in a geodesign dynamics, trying to go beyond the discussions between the actors, from their interventions in territorial planning. Secondly, the objective was to identify the possibility of creating new technical functionalities that are able to extract such information quickly and that can be part of the technological set of tools available to the development of geodesign as a participatory planning method.

Due to this goal, the option of using simple but effective spatial analyzes to obtain the necessary data to identify the spatial properties of the participants' interventions becomes the main focus of the method, allowing it to be constructed using the most diverse technical resources available without representing high computational costs.

Particularly in Brazil the geodesign experiments are in the initial phase by the planners and therefore the limited number of case studies on which the method of extraction of similarities was used. We hope that with the greatest diffusion in the use of this framework and

other applications to co-creation in urban planning it is possible to expand and even change the methods of extracting the similarities according to the needs. However, the subjectivities found in the use of the method of similarities are related to the limitations regarding the quality of data origin used in the workshop. Topological problems and even more objective representations on the part of the participants should be corrected by the platforms on which the process will be carried out, as well as by the professional in charge of conducting the process, thus guaranteeing the clarity and the understanding of the presented opinions.

Even with all subjectivity, and in relation to the geographic scales where the territorial planning is applied, the method tries to leave open the possibility of changes in the indices of similarity, as described in the conclusion of the article.

The methods presented in the article open the researchers a new way to expand and extract the understanding of the interventions obtained in the participatory process of geodesign, without, at any moment, considering the topic closed. The article seeks to show the real possibility of expansion through the use of spatial intelligence on the data produced in a platform that objectively seeks to bring discussions about planning through the technological set typical of geographic information systems environments.

ACKNOWLEDGMENTS:

CNPq, Process 401066/2016-9, Edital Universal 01/2016. Fapemig, Process PPM-00368-18, PPM XII.

REFERENCES

- Araújo, R. P. Z., Moura, A. C. M., Nogueira, T. (2018). Creating collaborative environments for the development of slum upgrading and illegal settlement regularization plans in Brazil: the Maria Tereza neighborhood case in Belo Horizonte. *International Journal of E-Planning Research (IJEPR)*, 7(4).
- Arnstein, S. R. (1969): A Ladder Of Citizen Participation. *Journal of the American Institute of Planners*, 35(4), 216-224.
- Brail, R. K., & Klosterman, R. E. (2001). *Planning Support Systems: Integrating Geographic Information Systems, Models, and Visualization Tools*. ESRI Press, California, 443 p.
- Casagrande, P. B. (2018). *O Framework Geodesign aplicado ao Quadrilátero Ferrífero, Minas Gerais - Brasil: A Geologia como base de Planejamento de futuros alternativos para o Quadrilátero Ferrífero*. Master Dissertation. Instituto de Geociências, UFMG.
- Czepkiewicz, M., Jankowski, P., & Młodkowski, M. (2017). Geo-questionnaires in urban planning: recruitment methods, participant engagement, and data quality. *Cartography and Geographic Information Science*, 44(6), 551-567.
- Flaxman, M. (2010). *Geodesign: Fundamental Principles and Routes Forward*. Talk at GeoDesign Summit.
- Hanzl, M. (2007). Information technology as a tool for public participation in urban planning: a review of experiments and potentials. *Design Studies*, 28(3), 289-307.
- So, F. S. (1979). *Practice of local government planning*. Published in cooperation with the American Planning Association by the International City Management Association.
- Lee, D. J., Dias, E., & Scholten, H. J. (Eds.). (2014). *Geodesign by integrating design and geospatial sciences* (Vol. 111). Springer.
- Linstone, H. A., & Turoff, M. (Eds.). (1975). *The delphi method* (pp. 3-12). Reading, MA: Addison-Wesley.
- Lynch, K. (1960). *The image of the city* (Vol. 11). MIT press.
- Saad-Sulonen, J. (2012). The role of the creation and sharing of digital media content in participatory e-planning. *International Journal of E-Planning Research (IJEPR)*, 1(2), 1-22.
- Smith, R. W. (1973). A theoretical basis for participatory planning. *Policy sciences*, 4(3), 275-295.
- Tuan, Y. (1974). *Topophilia*. Columbia University Press.
- Tuan, Y. F. (1990). *Topophilia: A study of environmental perceptions, attitudes, and values*. Columbia University Press.
- Zyngier, C. M.; Casagrande, P. B.; Moura, A. C. M.; Ribeiro, S. R. (2017). O Geodesign como plataforma para co-design: Estudo de Caso Maria Tereza. In: *XXI Congresso Internacional de la Sociedad Iberoamericana de Gráfica Digital*, Concepción. Blucher Design Proceedings. São Paulo, Editora Blucher, 3, 403-409.

Ferramentas de ETL Para Análise de Desempenho de Diagramas: Favorecendo Negociações em Workshops de Geodesign

1. INTRODUÇÃO: GEODESIGN COMO MÉTODO E O INTERESSE NA ANÁLISE DE SEUS DIAGRAMAS

Nos últimos 50 anos, os estudos em planejamento territorial consideram como valor e como requisito a inclusão das expectativas e decisões dos cidadãos, em uma abordagem de planejamento participativo. O sentido de planejamento participativo ou comunitário é entendido como um processo colaborativo, governança compartilhada, processo de “baixo para cima”, baseado na inclusão abrangente de pessoas do lugar.

Os interesses desse paradigma começaram nos anos 60, quando alguns autores desenvolveram teorias sobre percepção e cognição do espaço, para compreender como as pessoas criam sua relação com o território, os sentimentos sobre espaço e sentido de lugar, a importância da toponímia na cultura e na identidade social (Lynch, 1960; Tuan, 1974). No final da década, Arnstein (1969) propôs uma escada

de participação cidadã como referência teórica, para encorajar a inclusão de diferentes setores da sociedade no planejamento, enfrentando as questões da não-participação até o poder dos cidadãos.

O planejamento participativo tornou-se o centro dos interesses desde então, e hoje em dia o foco está sobre os métodos e possibilidades no Planejamento eletrônico, incluindo ferramentas digitais baseados na Web (Smith, 1973; International City Management Association, 1979). O planejamento eletrônico combina os recursos dos Sistemas de Informações Geográficas (GIS) com os objetivos dos Sistemas de Suporte ao Planejamento (PSS), utilizando ferramentas baseadas na web para decisões participativas (Hanzi, 2007; Brail e Klosterman, 2001; Saad-Sulonen, 2012).

A ideia de E-Planning começou com as tarefas de “Planos no mapa” e “Mostre no mapa”, fornecendo insumos sobre as preferências e expectativas das pessoas do lugar, usando “geo-questionários” (Saad-Sulonen,

2012; Czepkiewicz, Jankowski e Młodkowski, 2017).

O Geodesign também surgiu nesta era de planejamento participativo e comunitário, com o objetivo de dar suporte à opinião e tomada de decisões, melhorando o desempenho com o uso de ferramentas digitais. É um método para projetar com e para a geografia, por isso é geo + design. Segundo Flaxman (2010), o Geodesign é um método de projeto e planejamento que acopla fortemente a criação de propostas espaciais, apresentando simulações de impacto informadas por contextos geográficos. Lee, Dias e Scholten (2014) explicam que o Geodesign integra projetos e ciências geoespaciais.

Steinitz (2012) propôs um framework para o geodesign, composto por etapas que vão desde a avaliação do local até a construção de futuros alternativos no planejamento coletivo. Segundo Steinitz, a criação de cenários alternativos para um local deve levar em conta 6 modelos, 3 deles para descrever a área com

base na avaliação, e 3 deles para propor novos projetos e intervenções. Os modelos da etapa de avaliação são: representação, processo e avaliação. Os modelos da etapa de intervenção são mudança, impacto e decisão.

Na etapa de avaliação, o “Modelo de Representação” tem como objetivo responder “como deve ser descrita a área de estudo?”. O “Modelo de Processo” tem como objetivo responder “como funciona a área de estudo?”. E o terceiro modelo é o “Modelo de Avaliação”, que tem como objetivo responder “a atual área de estudo está funcionando bem?”. Finalizando esta parte, os organizadores e participantes podem ter uma caracterização clara das vulnerabilidades e potencialidades sobre o local, que são as principais questões a serem enfrentadas.

A partir de uma ideia clara sobre o local e suas necessidades, os participantes iniciam o passo de intervenção, que é composto por outros 3 modelos. O “Modelo de Mudança” tem como objetivo responder “como a área de estudo pode ser alterada?”. O “Modelo de Impacto” tem como objetivo responder “que diferenças as mudanças podem causar?”. Finalmente, os participantes chegam ao último modelo, o “Modelo de Decisão”, que tem como objetivo responder “como a área de estudo deve ser alterada?”.

O autor propôs seguir a diretriz de 6 modelos em 3 iterações. A primeira iteração é baseada em “por que estamos estudando a área?”. A fim de compreender as principais características da área de estudo e fazer com que as pessoas construam as primeiras propostas de projetos para o local. Depois de terminar o primeiro experimento, é importante analisar os resultados e revisar todo o processo, todos os modelos, em uma segunda iteração, com base em “como podemos desenvolver o estudo de caso?”. Após revisar todos os passos para calibrar todos os modelos, tem início a terceira e última iteração, seguindo novamente todos os seis modelos, com o objetivo de definir “o que, onde e quando?” criando os projetos finais para a área de estudo.

Durante as etapas do framework, os participantes recebem e produzem informações sobre o local e constroem juntos projetos para o território. Designs são conjuntos de projetos e políticas que eles acreditam

serem as melhores ideias para a área. Eles fazem a dinâmica de co-construção de propostas, sempre buscando o consenso. Eles desenham suas próprias ideias, mas também desenham novas ideias observando as propostas uns dos outros, ou até mesmo desenham novas ideias com base no senso comum do grupo.

Os passos de negociação no geodesign podem ser feitos no modo analógico, usando mapas impressos e sobreposições. Mas, nos últimos anos, Steinitz em parceria com Ballal, desenvolveu uma ferramenta baseada na Web para a geodesign, chamada “GeodesignHub²”. Os organizadores produzem os modelos de avaliação (representação, processo e avaliação) para alimentar a plataforma web com sistemas sobre as principais características do local, classificados de acordo com a potencialidade para receber propostas (níveis de “factual”, “adequado” ou “capaz”), além da inadequação (classificada como “inapropriada”), assim como mostrando áreas onde o problema já se encontra resolvido, portanto, se a necessidade de construção de propostas (classificadas como “existentes”). (Figura 1).

Figura 01 – Exemplo de modelos de avaliação organizados em 10 sistemas usados como referência para criação dos diagramas de projetos e políticas. Estudo de caso “Quadrilátero Ferrífero”, Brasil.

Também é importante definir uma tabela de impactos, de acordo com conflitos ou adequações, em avaliação combinatória entre os sistemas. Um diagrama de projeto ou política pode ter partes coloridas do roxo ao laranja, da combinação mais positiva à combinação mais negativa, destacando conflitos territoriais de interesses (Figura 2). É com base nessa matriz que serão avaliadas as propostas de projetos ou políticas em cada temática, classificando as combinações de +2 (impacto mais positivo) a -2 (impacto mais negativo) (Figura 3).

Figura 02 – Exemplo de diagrama do sistema “ECO – Ecológico” confrontado com os diagramas dos demais sistemas. Parte do diagrama possui impacto “mais positivo” para alguns sistemas e parte “muito negativa” para outros, além de regiões de impacto neutro. Estudo de caso “Quadrilátero Ferrífero”, Brasil

Figura 03 – Exemplo da matriz de impacto. No exemplo, se o participante propuser um projeto do sistema “ECO” (interesse ligados a ecologia e criação de áreas de proteção) em áreas importantes para o sistema “HYDRO” (proteção dos recursos hídricos), o impacto é +2, por ser

considerado muito positivo. No entanto, o projeto ou política poderá criar problemas para as áreas de interesse para indústria, alta ou baixa densidade habitacional, assim como para áreas de expansão e dinamismo (“IND”, “LDH”, “HDH”, “UGRT”), sendo, portanto, considerada de -2, muito negativo.

Os participantes desenham diagramas por sistema, apresentando ideias de projetos e políticas para resolver problemas ou desenvolver potencialidades. O participante pode usar como base o mapa de avaliação do sistema, uma combinação de alguns mapas de avaliação ou mesmo mapas de base com referências espaciais. Desde o início do workshop até o último momento de construção da proposta final, novos diagramas podem ser desenhados, como parte do processo de negociação e troca de ideias. (Figura 4).

Figura 04 – Dinâmica de proposta de diagramas para cada sistema.

Os diagramas são nomeados de acordo com a ideia principal que representam, têm suas dimensões de área e perímetro calculadas, podem receber hashtags para facilitar a busca organização de ontologias, são coloridas de acordo com os impactos dos sistemas e também é possível analisar o custo e a obtenção dos valores das metas propostas para o estudo de caso. (Figura 5).

Figura 05 – Os diagramas com hachuras representam políticas, enquanto os diagramas sólidos representam projetos. Estudo de caso “Quadrilátero Ferrífero”, Brasil.

Uma vez elaborado o conjunto de diagramas, para diferentes temas, os participantes são organizados em equipes que representam setores da sociedade. Nessas equipes, eles selecionam os diagramas que melhor se adaptam às suas ideias e compõem “Designs”. Os projetos são composições de projetos e políticas para o local, considerando todos os sistemas, e representando os principais valores e expectativas de acordo com a equipe. (Figura 6).

Figura 06 – Composição de designs com base na coleção de diagramas.

A comparação de designs propostos pelas equipes é uma ação muito importante durante o workshop, pois demonstra a diferença de visão dos grupos. Ao final do workshop produzirá um diagrama final, produzido a partir da negociação, e é importante reconhecer as semelhanças e os

conflitos a serem enfrentados. (Figura 7, Figura 8).

Figura 07 – Organização das equipes para construção de propostas de design, coleção de projetos e políticas. Estudo de caso “Quadrilátero Ferrífero”, Brasil.

Figura 08 – Exemplo de 6 diferentes designs, com a relação de diagramas que melhor representam a expectativa de cada equipe. Estudo de caso “Quadrilátero Ferrífero”, Brasil.

A comparação de propostas ajuda na definição de uma solução comum, ou a negociar uma solução aceitável, assim como a refutar ideias conflituosas (Figura 9). Durante algumas experiências de workshops, reconhecemos algumas situações em que as equipes chegaram a uma proposta comum com mais facilidade e outras em que um projeto final era quase impossível. Mas, em alguns casos, a discordância ocorria sem que os participantes percebessem que estavam falando da mesma coisa, acreditando estar em lados opostos.

Figura 09 – Negociação do design final

Dois estudos de caso nos fizeram pensar sobre a importância de entender como os designs podem ser semelhantes ou conflitantes: o estudo de caso do assentamento ilegal Maria Tereza, em Belo Horizonte, e o estudo de caso da geologia como base para o planejamento de futuros alternativos do Quadrilátero Ferrífero”, em Minas Gerais, ambos no Brasil. O primeiro estudo de caso foi desenvolvido sob a colaboração do Laboratório de Geoprocessamento da Faculdade de Arquitetura da Universidade Federal de Minas Gerais com o município de Belo Horizonte (Araújo et al, 2018; Zyngier et al, 2017), e o segundo estudo de caso foi Dissertação de Mestrado de um estudante do Laboratório de Geoprocessamento (Casagrande, 2018).

A dissertação sobre os estudos geológicos como base para geodesign na área do Quadrilátero Ferrífero nos fez entender que, quando se trabalha com equipes que possuem a mesma inserção social e qualificação profissional similar (no estudo de caso, todas da Terra e Ciências Ambientais), as propostas tendem a ser bem parecidas e, em poucas rodadas de negociação, é possível chegar a um projeto final. (Figura 10).

Figura 10 – Exemplo de “Geologia como base para planejamento de futuros alternativos para o Quadrilátero ferrífero, por Casagrande (2018).

O estudo de caso de assentamento ilegal de Maria Tereza foi muito importante para entender o papel de reconhecer semelhanças em diferentes projetos. O município de Belo Horizonte convidou o grupo do Laboratório de Geoprocessamento para aplicação do método geodesign em um PRU (Plano de Regularização Urbana) de um assentamento urbano ilegal. Por ser a primeira vez que o município usava o método, foi escolhida a participação apenas dos técnicos da administração pública, para que pudessem decidir sobre sua adequação a um estudo de caso real. Foi necessário receber a opinião de muitos técnicos, e eles eram numerosos, 70 participantes no total. O plano era utilizar a plataforma web GeodesignHub², no entanto, era possível colocar todos em uma mesma sala. A solução foi dividir o grupo em dois, e desenvolver workshops independentes, chegando a dois projetos finais. (Figura 11).

Figura 11 – Workshops do estudo de caso Maria Tereza, Belo Horizonte, Brasil, 2016.

Os dois grupos não tiveram contato uns com os outros, já que os trabalhos foram realizados em dias diferentes, mas as propostas finais eram muito parecidas. Pode-se dizer, portanto, que quando a proposta é elaborada por agentes sociais que são de contexto semelhante, as propostas resultantes representam critérios defensáveis e reproduzíveis. Não é difícil chegar a uma proposta coletiva defensável, baseada nas principais ideias e valores coletivos.

Na ocasião, foi utilizada a inspeção visual para reconhecer semelhanças entre os projetos finais, colocando os diagramas lado a lado, por sistema e comparando as propostas. Na maioria dos sistemas, os desenhos eram diferentes, mas as ideias eram bastante semelhantes assim como era possível identificar propostas com as mesmas ideias, mas em posições diferentes na região. Entendemos que se tivéssemos uma ferramenta para analisar os diagramas, a informação sobre as similaridades poderia ser apresentada aos participantes, durante a oficina ou em tabelas de discussões entre as reuniões. A partir destas informações os participantes perceberiam que em vários momentos as divergências ocorriam sobre propostas semelhantes, já que

suas ideias eram bem parecidas. (Figura 12).

Figura 12 – Exemplo de comparação de diagramas sobre projetos e políticas em dois sistemas no estudo de caso de Maria Tereza, Belo Horizonte, Brasil.

O interesse em comparar diagramas foi planejado inicialmente para ser aplicado em geodesign, para encontrar similaridades em projetos e políticas, mas também pode ser útil em muitas outras aplicações, sempre que for importante identificar como polígonos georreferenciados são semelhantes. (Figura 13).

Figura 13 – Como comparar a similaridade dos diagramas? Na figura é apresentada a coleção de diagramas propostos no estudo de caso de Maria Tereza.

2. PROPOSTA: O SENTIDO DA SIMILARIDADE

Geodesign é um método para apoiar a tomada de decisão na co-criação de futuros alternativos para um território. As propostas são formadas por conjuntos de diagramas representando projetos e políticas, de modo que as principais decisões giram em torno dos seguintes aspectos, o lugar (onde?), dimensão (quanto?), ideia principal (o quê?) e possibilidades (quais?). Nesse sentido, busca-se o interesse em desenvolver estudos para identificar afinidades ou conflitos entre diagramas: semelhança topológica, semelhança posicional, semelhança conceitual e semelhança taxonômica. (Figura 14).

Figura 14 – Estudos sobre similaridades.

A similaridade “conceitual” é, de alguma forma, já controlada no GeodesignHub², já que o coordenador do geodesign é solicitado a inserir os dados da matriz de impacto no sistema (Figura 03). Ao decidir sobre impactos positivos, negativos ou neutros, o coordenador está definindo as referências que serão usadas para classificar cada diagrama, colorindo de roxo a laranja (Figura 02). Os resultados fazem com que os participantes identifiquem os conflitos ou acordos conceituais, baseados em diferentes ideologias ou valores esperados para o mesmo lugar. Nossos estudos, para colaborar neste tópico, são baseados em ferramentas para compartilhar decisões sobre julgamentos e avaliações. Em alguns

estudos de caso que conduzimos, os participantes questionaram sobre os valores definidos e disseram que queriam ser consultados sobre sua opinião. Desde então, estamos aplicando o método Delphi para decidir sobre valores na matriz.

O método Delphi é baseado na construção de consenso por aproximações (LINSTONE e TUROFF, 1975). Os participantes respondem aos questionários em duas rodadas. No primeiro turno, eles apresentam sua opinião sobre o impacto de um sistema no outro, a maioria e a média das opiniões são calculadas e apresentadas. Comparando a média de opinião com a sua própria opinião, eles respondem ao questionário novamente, com a possibilidade de manter ou mudar seus votos, após uma nova média é calculada. Esta é usada na matriz final. Testamos ferramentas digitais baseadas na Web para realizar o procedimento: o Survey Monkey e o Google Forms.

A similaridade da “taxonomia” tem o objetivo de entender se as pessoas usam palavras diferentes para a mesma ideia. Também é interessante conectar ideias semelhantes entre si, para compor ontologias, agrupando famílias de conceitos. O GeodesignHub² já está preparando funcionalidades para realizar esses estudos, pois ao criar os polígonos, os autores podem adicionar hashtags. Em alguns estudos de caso realizados (Faenza, na Itália, e Pampulha, no Brasil, ambos em 2017), testamos os aplicativos para coletar ideias prévias dos participantes, antes do workshop, para compor uma lista de possibilidades, pensando em famílias de ideias e ontologias.

No estudo de caso de Faenza, usamos o mapeamento colaborativo para receber opiniões sobre o local. No estudo de caso da Pampulha e em Faenza, foram criadas plataformas WebGis com as principais variáveis do local e instruímos os participantes que navegassem sobre os mapas antes do workshop, para reconhecer vulnerabilidades e potencialidades do território. Depois de analisar o WebGis, eles foram solicitados a responder aos questionários da web, apresentando palavras-chave de ideias para responder as necessidades de cada sistema. Método semelhante ao de geoquestionários. Durante o workshop, a lista de palavras-chave com ideias foi compartilhada com os participantes. Estudos futuros podem abordar a

composição de ontologias de palavras e ideias principais.

Nosso estudo foi concentrado na identificação das similaridades topológicas e locais, para as quais construímos scripts para calcular automaticamente as relações entre os diagramas. O estudo da similaridade topológica foi baseado na interseção dos diagramas: caso a área de interseção é maior que 80% da área total de ambos os diagramas originais, eles são considerados semelhantes. Para suportar uma análise mais complexa, também é importante calcular se um diagrama está contido, ou se contém o outro. (Figura 15).

Figura 15 – Similaridade topológica e relações espaciais na interação de diagramas

Os estudos sobre similaridade locacional, baseados em distâncias, foram mais difíceis de definir, porque o sentido de “perto” ou “longe” depende de outras referências espaciais. O sentido de “perto” não é o mesmo de interseção ou toque, porque as formas dos diagramas podem interferir na relação do posicionamento (Figura 16). Decidimos pela definição de um retângulo envolvente, com base na extensão da área total do estudo de caso (Figura 17).

Figura 16 – Interseções ou vizinhanças não são suficientes para definir a similaridade locacional.

Fig. 17 – A definição do retângulo envolvente como referência para classificação de proximidade entre diagramas

Uma vez definido o retângulo envolvente, o menor lado é medido e os centroides dos diagramas são identificados. Dois diagramas são considerados próximos um do outro se a distância entre os centroides for menor que 12,5% em relação ao menor lado do retângulo (Figura 18). Esse valor é experimental e requer mais estudos. Outros estudos que estão em desenvolvimento são sobre o uso da distribuição regular de pontos dentro dos diagramas e o uso do mosaico para estabelecer uma rede de vizinhanças, e, em vez de medir apenas os centroides para definir a distância entre todos os pontos de diferentes diagramas, resultando no ranking das vizinhanças.

Figura 18 – Medindo a similaridade locacional.

3. EMPREGO DE FERRAMENTAS DE ETL PARA MODELAGEM

Os suportes tecnológicos selecionados para serem utilizados nas tarefas foram as ferramentas ETL - Extração, Transformação e carga, nas quais foram criados scripts baseados nos modelos espaciais de análise de similaridade topológica e locacional.

3.1. FERRAMENTAS DE ETL

Ferramentas de ETL são geralmente associadas ao uso em banco de dados, para procedimentos de integração e validação de dados, quando é necessário processar grande quantidade de informações e criar rotinas a serem aplicadas sem a supervisão dos técnicos. As tarefas de tratamento, conversão e manipulação das informações também podem ser desafios às ferramentas de banco de dados, pois o processamento lógico é baseado no armazenamento de dados e no resgate de informações já estruturadas. Nesse sentido, essas ferramentas de ETL surgiram como dispositivos auxiliares que fornecem amplo conjunto de funcionalidades e algoritmos para manipulação dos dados.

Além das funcionalidades, as ferramentas de ETL podem trabalhar utilizando a arquitetura cliente-servidor, servindo, portanto, como parte de uma solução em integração de dados, de acordo com os componentes tecnológicos existentes. Atualmente, essas ferramentas são utilizadas na construção de Data Warehouse (Armazém de dados), banco de dados não estruturado, direcionado ao consumo final de informações, como o BI - Business Intelligence. Isso requer a organização das informações mais adaptadas aos especialistas, em vez da modelagem lógica típica do sistema de informações.

Nos últimos anos, as ferramentas de ETL passaram a ser capazes de ler e processar dados espaciais, acompanhando a evolução dos bancos de dados. Os ETL's fornecem um conjunto de funcionalidades espaciais que, em alguns casos, podem substituir as ferramentas tradicionais de SIG (Sistemas de Informações Geográficas). O principal ganho no uso de ETL é a interoperabilidade de dados, pois são capazes de preparar e armazenar dados de muitos formatos diferentes, em conjuntos de dados

clássicos ou em conjuntos de dados geográficos

No estudo de caso apresentado foi aplicado o pacote de ferramentas de ETL espacial, GeoKettle®, versão 2.5, baseado no pacote Pentaho Data Integration. (Figura 19)

Figura 19 – Modelo conceitual de ETL geográfica Geokettle®.

A escolha da aplicação levou em consideração as seguintes características:

- Ser compatível com dados espaciais e realizar transformações capazes de usar conceitos topológicos em análise espacial;
- Ser de livre domínio;
- Ser expansível, trabalhando de formar distribuída, permitindo a realização de várias tarefas ao mesmo tempo, além de permitir uma arquitetura baseado em computação em nuvem ou servidores dedicados;
- Apresentar melhor integração com infraestruturas de dados espaciais orientadas a serviços;
- Apresentar interface gráfica amigável que permite a construção e alteração dos modelos de análise sem a necessidade de conhecimentos específicos;
- Permitir que os processos desenvolvidos possam ser salvos em arquivos e executados a qualquer momento, quantas vezes necessárias, portanto, de fácil difusão entre usuários.

3.2. MODELOS ESPACIAIS

A partir da seleção das ferramentas, a tarefa foi a composição de roteiros, para compor modelos para as análises de similaridade topológica e locacional.

3.2.1 – SIMILARIDADE TOPOLÓGICA

A descrição do modelo é apresentada em etapas, para facilitar a compreensão de todo o processo (Figura 20). Os diagramas foram baixados do GeodesignHub® (o script foi desenvolvido para comparar os diagramas existentes construídos durante um workshop de geodesign) e convertidos para shapefile, para que as ferramentas GIS pudessem ser usadas.

Figura 20 – Modelo de Similaridade Topológica – plano global

A primeira etapa do modelo é a leitura de dados, identificando e corrigindo possíveis problemas topológicos existentes no banco de dados de diagramas e gerando uma matriz de combinação de diagramas existentes. Se os erros topológicos que acompanham os diagramas não forem resolvidos, alguns procedimentos não são executados, pois são baseados na intersecção de polígonos. A ferramenta Geokettle® é sensível a problemas topológicos, o que pode causar a interrupção da análise. (Figura 21).

Figura 21 – Primeira etapa de análise da similaridade topológica.

Para resolver problemas topológicos, algumas funcionalidades são usadas para separar os diagramas que apresentam erros, nos quais são construídos pequenos buffers corrigindo os desenhos sem alterá-los em demasia (passos Filter rows 2, error e Spatial Analysis). Uma vez corrigido os problemas, o conjunto de dados é separado em dois grupos idênticos. O passo “cross join”, é empregado para permitir a construção da matriz de combinação entre os diagramas. É importante evitar a combinação de um diagrama consigo mesmo, além eliminar a repetição combinações já realizadas. Para evitar a repetição de combinações (A com B é o mesmo que B com A), são feitas subtrações usando o número de identificação dos diagramas. Como resultados, um conjunto de dados será positivo (id2 - id1) e outro será negativo (id1 - id2), permitindo assim a separação dos mesmos através do passo “filter combinations”.

A segunda etapa da análise é a intersecção espacial dos polígonos, o cálculo das áreas resultantes e o cálculo das proporções dessas áreas de intersecção. (Figura 22). É importante destacar que formas problemáticas de polígonos (erros topológicos) podem resultar das intersecções dos diagramas. Para resolver essas ocorrências, a ferramenta pode produzir geometrias chamadas “coleção”, conversão de polígonos muito estreitos ou muito pequenos em pontos ou linhas. Depois disso, é possível eliminar as geometrias indesejáveis da análise.

Figura 22 – Segunda etapa da análise de similaridade topológica

O terceiro e último passo começa com a identificação das principais relações observadas a partir das combinações de diagramas (Figura 23). Os diagramas são considerados semelhantes somente se a área de intersecção em ambos (área em A da intersecção com B e também área em B da intersecção com A) for maior que 80%.

Figura 23 – Terceira etapa de análise da similaridade topológica

O script calcula não apenas o nível de similaridade, mas também as relações espaciais de contém e está contido (está dentro). A relação é reconhecida como “conter” quando a área de intersecção no primeiro diagrama é menor que 80% e no segundo diagrama é maior que 80%, enquanto é considerada “dentro” quando a área de intersecção no primeiro diagrama é maior que 80% e no segundo é inferior a 80%. Como resultado, é construída uma matriz final com todas as combinações existentes. (Figura 24).

Figura 24 – Matriz de similaridade e relações entre os diagramas

3.2.2 – SIMILARIDADE LOCACIONAL

A descrição do modelo também é apresentada em etapas para facilitar o entendimento do processo como um todo (Figura 25).

Figura 25 – Modelo de Similaridade Locacional – plano global

A primeira etapa da análise de similaridade locacional é a construção da matriz de combinação cruzada de diagramas. Aplicaram-se os mesmos processos para evitar que o diagrama fosse combinado consigo mesmo, assim como, evitar a repetição de combinações (considerando A com B e eliminando B com A).

A segunda etapa começa com a seleção da referência dimensional, definida pelo retângulo envolvente para o conjunto de diagramas existentes. O processo pode ser construído em paralelo ao primeiro passo (Figura 26). Devido a especificidades da ferramenta Geokettle®, retângulo é convertido em linhas e as linhas convertidas em pontos. Utilizando os pontos, que representam os

vértices do retângulo, a ferramenta calcula as distâncias dos lados, com o objetivo de identificar o menor.

Figura 26 – Segunda etapa de análise da similaridade locacional

A terceira etapa constrói os centroides para cada diagrama e calcula as distâncias para todos os demais diagramas (Figura 27).

Figura 27 – Terceira etapa da análise de similaridade locacional

Na quarta e última etapa, todas as distâncias dos centroides são comparadas com o valor da menor face do retângulo envolvente (Figura 28). Foi definido que as distâncias entre os centroides não podem ser maiores que 12,5% que o menor lado. As ferramentas aplicadas são “calculate proportion” e “filter rows”.

Figura 28 – Quarta etapa da análise de similaridade locacional

Como resultado a matriz de similaridade locacional é produzida (Figura 29).

Figura 29 – Matriz de similaridade locacional

5. RESULTADOS E DISCUSSÕES

Foram realizados diversos testes empregando os scripts de ETL, usando diagramas de diferentes workshops, e os primeiros resultados foram considerados muito positivos. Para o modelo de similaridade topológica, baseado no valor de 80% da interseção, foi considerado aceitável, e também foi possível identificar claramente as relações de “contém” e “está dentro” (Figura 30).

Figura 30 – Exemplo de similaridade topológica acima de 80% e da relação “está contido”.

O script de similaridade locacional também foi capaz de identificar facilmente os diagramas que estavam próximos uns dos outros. A semelhança locacional pode ser calculada entre diagramas do mesmo sistema, mas também usando diagramas de sistemas que possuem conflitos entre si. No GeodesignHub², o aplicativo reconhece e classifica conflitos apenas quando os diagramas são sobrepostos, mas o novo

script ETL construído pode ampliar a análise que identifica os conflitos das vizinhanças. (Figura 31).

Fig. 31 – Exemplo de similaridade locacional e possíveis conflitos

Relações mais complexas entre diagramas e sistemas podem ser identificadas e analisadas. Semelhanças topológicas entre diagramas de diferentes sistemas podem ser identificadas, e também as relações de “contém” e “está dentro”, fornecendo novas análises. No exemplo (Figura 32), na linha 288, um diagrama de habitação de alta densidade (HDH) foi identificado como semelhante ao diagrama ambiental (ENV) e situado dentro (dentro), o que pode representar um conflito ou um problema a ser resolvido pelo planejamento. Na mesma tabela, na linha 291, um diagrama do sistema ambiental (ENV) contém um diagrama do sistema recreativo (REC), que pode ser considerado complementar, dando suporte às decisões de planejamento sobre a área. Ainda na mesma tabela, linha 293, um diagrama do sistema ambiental (ENV) que é semelhante a outro do mesmo sistema (ENV), resultando em propostas redundantes. (Figura 32).

Figura 32 - Analysis of topological similarity and systems

Os resultados indicam o grande potencial na exploração dos estudos de similaridade, que pode ser ainda mais interessante se considerar outras características dos diagramas. É importante ressaltar que a análise espacial e as relações espaciais devem ser entendidas como um estudo multidimensional, combinando diferentes parâmetros.

6. CONCLUSÕES

Os modelos de ETL desenvolvidos visavam a utilização de análises simples, que poderiam ser implementadas mesmo utilizando outras ferramentas, de acordo com as preferências dos usuários. A velocidade de execução era mais importante que a complexidade de possibilidades a serem empregadas. É importante destacar que, usando o exemplo de 102 diagramas complexos de 10 sistemas, o script levou menos de 3 segundos para calcular a similaridade topológica e menos de 1 segundo para calcular a semelhança locacional.

A matriz de combinação foi capaz de identificar facilmente mais de 10.000 combinações diferentes. A simplicidade dos métodos empregados é, portanto, uma grande vantagem na difusão dos estudos.

Os desafios observados durante o processo foram relacionados aos problemas topológicos que podem existir na fonte de dados primária e que precisam ser abordados antes de realizar as análises.

Dois possíveis problemas foram identificados nos estudos sobre similaridade locacional. O primeiro está relacionado a diagramas muito extensos, que cobrem grande parte da área de estudo, pois serão considerados semelhantes à maioria dos diagramas. Estes diagramas devem ser tratados de forma diferente, é raro que alguém possa propor algo que cubra todo o território. Mas isso acontece. Eles geralmente estão relacionados a políticas aplicadas a toda a área ou que podem ser revisadas e aplicadas especificamente a lugares mais adequados. Além disso, propostas muito extremas são construídas por participantes que não estão dispostos a negociar com outras ideias ou valores e terminam por demarcar todo o território.

O segundo problema está relacionado a diagramas que são criados como multipolígonos ou multipartes. Neste caso, se as muitas partes são distribuídas no território, o conjunto de diagramas tem o mesmo comportamento que um diagrama de grande extensão, porque será considerado próximo de tudo, entendido como uma única proposta. Durante o workshop, geralmente é explicado aos participantes para evitar a criação de diagramas compostos por mais de um polígono. Os participantes são orientados a optar por salvar diferentes diagramas de acordo com alguma classificação ou lógica. O motivo principal é devido ao processo de negociação, porque outros podem aceitar apenas parte de sua proposta, mas nem todos os polígonos são projetados. Apesar disso, é comum a existência de participantes que preferem desenhar um conjunto de polígonos na mesma proposta, o que requer filtros anteriores a análise de similaridade, como os problemas de polígonos muito grandes.

É recomendável também adaptar a ferramenta para que o usuário possa decidir sobre os valores dos parâmetros usados para identificação das similaridades.

Por exemplo, na semelhança topológica, em vez de fixar o valor em 80%, o usuário pode defini-lo, de acordo com seu conhecimento sobre o estudo de caso. O mesmo poderia acontecer no caso da similaridade locacional em que o valor definido é de 12,5%.

A validação dos métodos desenvolvidos nos estudos abre novas possibilidades para investigar as relações em métodos de co-design. O reconhecimento de semelhanças pode ser utilizado nas etapas de negociação, na construção de consenso ou na definição de prioridades a serem discutidas.

Conforme foi mencionado nas discussões e conclusão do artigo, o objetivo em testar o método de análise descrito foi, em um primeiro momento identificar a possibilidade de emprego de análises espaciais para enriquecer os resultados obtidos em uma dinâmica de geodesign, procurando ir além das discussões entre os atores, a partir das intervenções destes no ordenamento espacial.

Em segundo lugar o objetivo foi identificar a possibilidade de criação de novas funcionalidades técnicas que sejam capazes de extrair tais informações de maneira rápida e que possam fazer parte do conjunto tecnológico das ferramentas disponíveis ao desenvolvimento do

geodesign enquanto método de gestão participativa.

Nesse aspecto a opção pelo emprego de análises espaciais simples, mas eficazes na obtenção dos dados necessários para identificação das propriedades espaciais das intervenções dos participantes, torna-se o foco principal do método permitindo que este possa ser construído pelos mais diversos recursos técnicos disponíveis, sem representar custos computacionais elevados.

Particularmente no Brasil as experiências em geodesign estão em fase inicial pelos planejadores e por isso o número limitado de estudos de caso sobre o qual o método de extração de similaridades foi utilizado. Esperamos que com a maior difusão no emprego deste framework seja possível expandir e mesmo alterar os métodos de extração das similaridades segundo as necessidades encontradas. No entanto, como informado pelos autores, as subjetividades encontradas no emprego do método de similaridades estão relacionadas as limitações quanto a qualidade de origem do dado usado no workshop. Problemas topológicos e mesmo de representações mais objetivas por parte dos participantes devem ser corrigidas pelas plataformas sobre a o processo será realizado, assim como pelo profissional encarregado

de conduzir o processo, garantindo assim a clareza e o entendimento das opiniões apresentadas.

Mesmo com toda a subjetividade, e em relação as escalas geográficas em que o planejamento é aplicado, o método procura deixar em aberto a possibilidade de alterações nos índices de similaridade, conforme descrito na conclusão do artigo.

Consideramos a intervenção do revisor pertinente, no entanto, também consideramos que os métodos apresentados no artigo abrem aos pesquisadores um novo caminho para ampliar e extrair o entendimento das intervenções obtidas no processo participativo do geodesign, sem, em momento algum, considerar o tema encerrado.

O artigo procura mostrar a possibilidade real de ampliação através do emprego da inteligência espacial sobre os dados produzidos em uma plataforma que busca objetivamente trazer as discussões sobre o planejamento através do conjunto tecnológico típico dos ambientes de sistemas de informação geográfico

AGRADECIMENTOS

CNPq, Process 401066/2016-9, Edital Universal 01/2016. Fapemig, Process PPM-00368-18, PPM XII.

NOTAS

[1]. <http://www.openstreetmap.org>

[2]. <https://developers.google.com/places>

[3]. <https://foursquare.com>

[4]. <https://www.yelp.com>

[5]. <http://www.waze.com>

[6]. <http://ushahidi.com>

[7]. <http://www.dpi.ufv.br/projetos/clickonmap/>

[8]. <http://aqui.io/themerise>

[9]. Lei 12.527, de 18 de novembro de 2011. http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/12527.htm (in Portuguese)

[10]. <https://opengovdata.org>

[11]. <http://gtfs.org>. A especificação GTFS foi inicialmente proposta pelo Google e pela autoridade de trânsito para Portland, Oregon, como um padrão de dados abertos. Seu nome original era Google Transit Feed Specification. Ver <http://beyondtransparency.org/chapters/part-2/pioneering-open-data-standards-the-gtfs-story> for a complete background on GTFS.

[12]. <https://prefeitura.pbh.gov.br/bhtrans/informacoes/dados/dados-abertos>

[13]. <http://www.inde.gov.br>

[14]. <http://datageo.ambiente.sp.gov.br>

[15]. <http://geoportal.ide.ba.gov.br>

[16]. <http://dados.prefeitura.sp.gov.br>

[17]. <http://dados.pbh.gov.br>

[18]. <http://inspire-geoportal.ec.europa.eu>