Descriptive Sentence Extraction for Text to 3D scene Generation

Valentina Bova
Department of Informatics, Modeling, Electronics and System Engineering
University of Calabria
Rende, Italy
e-mail: valentina.bova@dimes.unical.it

Elena Cardillo
Institute of Informatics and Telematics
National Research Council
Rende, Italy
e-mail: elena.cardillo@iit.cnr.it

***Abstract*—Three-dimensional objects (3D) allow extensive and heterogeneous information to be stored in single models which can be exploited by users to satisfy various research and study needs. Moreover, the 3D visualization would be even more interesting if it were the result of the "materialization" of descriptive sentences extrapolated from texts related to the subject matter. In other words, a direct connection between 3D models and the associated texts or drawings could provide a useful and stimulating explication of the case of study. The extrapolation of specific information from texts, however, is time-consuming and it requires the user to have in-depth knowledge of the referring domain. An innovative solution to the problem, then, would be to develop a system that can analyse and "comprehend" the documents in order to automatically provide, as output, portions of text containing geometric and spatial information useful for the 3D scenes generation. In this paper, the analysis of the framework of the above-mentioned system is presented and its implementation on a specific corpus concerning the "World City" project, is evaluated.**

Keywords-Sentence Extraction; 3D Models; World City project; Text-to-Scene Conversion.

I. INTRODUCTION

Three-dimensional (3D) reconstruction is an emerging methodology used in several areas, such as art, education, robotics and, nowadays, it is also developing in the Cultural Heritage field. According to the principles for the Conservation and Restauration of Built Heritage "In the protection and public preservation of archaeological sites, the use of modern technologies, databanks, information system and virtual presentation techniques should be promoted" [1]. But that is not all, because the relationship between Cultural Heritage and Information and Communication Technologies (ICTs) could be exploited not only to represent sites and artworks of the past, still existing or no longer existing, but also to spread the knowledge of projects never realised. Thanks to virtual models, in fact, researchers have the possibility to study and formulate conclusions about sites and object placements, characteristics and configurations, while, more general users may discover and understand their beauty and importance.

The aim of this work is to spread knowledge of a worldwide known project developed by the French bibliographer Paul Otlet in collaboration with international architects such as Andersen, Hébrard, Le Corbusier and Heymans [2]: the "World City" [3]. The complex project has been described in several drawings and textual documents which contain, in a dispersed and fragmented way, information of heterogeneous nature. Conversely, a 3D representation could include, in a single composite model, all the data that could be immediately transferred to the users. In fact, the 3D model has a strong communicative power in addition to the great advantage of being a "universal datum". This means that the transfer of information does not depend on the used written language, weakness of the texts, or on the users' domain knowledge and imagination, which often cause erroneous or different interpretations of the information. On these bases, this work aims to exploit a virtual 3D of the *World City* in order to describe the project in an innovative way: the creation of 3D scenes starting from spatial descriptions contained and extracted from a referent corpus of texts. The project, in fact, was the result of numerous years of study and research attested by primary and secondary sources.

The paper is divided into sections as follow: Section II provides an overview of similar systems. Section III summarizes the 5 different phases of the methodology and describes the first two in detail (Domain Analysis and Corpus Compilation, Information Extraction). Section IV shows the results of the sentence extraction system. In section V, concluding remarks are presented.

II. BACKGROUND

The work proposes to achieve its aim by a methodology that, through a set of rules, will lead to the reconstruction of virtual scenes regarding the World City project. The idea is to generate associations between terms and 3D elements writing some scripts able to analyse the texts, extract the relevant information from them and handle the location of the three-dimensional objects in the virtual environment.

A. Related works

Several systems that can interpret natural descriptions to create a visual representation have already been developed in the past. A text-to-scene conversion system, frequently cited as the first in this field, is the tool SHRDLU [4], which allows users to insert, as input, commands written in natural language with the purpose of using them to move virtual objects around in a small "blocks world". Likewise, Natural Language Image Generation (NALIG) [5] is able to generate scenes by tacking sequences of descriptions which concern the spatial relations between objects. The pioneering WordsEye [6] is one of the main works that focused on this theme, in fact, it creates three-dimensional scenes from input descriptions by converting a parse tree to a dependency representation that, in turn, is transformed into a semantic one. Other similar systems are [7] [8], which use manual links between language and objects in order to create three-dimensional scenes in a virtual space.

Figure 1. Perspective of the project, Archives of the "Fondation Le Corbusier"; Paris; (24525).

Chang et al. [9] present an advanced 3D generation approach able to learn from data; indeed, the system is capable of mapping, in an automatic way, textual terms to objects and of creating a 3D representation. Furthermore, other works like those described in [10] [11] addressed the development of systems able to infer the presence of implicit objects, or constraints on the objects, in a described scene and to generate a final 3D representation. Sproat [12], instead, focuses on the possibility of inferring the environment of the textual descriptions; for example, taken into account the sentence "Carl is having a shower" the system tries to infer that the scene is set in a bathroom and not in a bedroom or in a kitchen. Big efforts have also been spent, in the literature, to propose solutions for the automatic or semi-automatic extraction or annotation of spatial relations and spatial objects from texts in order to generate 3D scenes. The SpatialML annotation scheme, for example, aims to mark places, including buildings, mentioned in a text (indicated with PLACE tags) and map them to data from gazetteers and other databases. Semantic attributes, such as country abbreviations, subdivision and dependent area abbreviations, and geo-coordinates are used to help establish such a mapping. Rules are language-dependent for marking up SpatialML tags, while are language-independent for marking up semantic attributes of tags [13]. Also, Klien and Lutz [14] proposed a method based on spatial relations for automatizing the semantic annotation process. In particular, they showed how the use of spatial relations at the data level, thus expressed through spatial processing methods (e.g., the calculation of the topology, direction or distance between two spatial entities) can be exploited for the semi-automatic semantic annotation of geodata. Concept definitions and spatial relations here can be extracted from geographic domain ontologies. Other works are focused on the use of a markup language, based on the Text Encoding Initiative (TEI) Guidelines, for semantically annotating raw texts and, in particular, for the task of Named Entities Recognition and Spatial Role Labeling [15]. Regarding the use of ontologies, in [16] they are proposed to bridge between cognitive–linguistic spatial concepts in natural language and multiple qualitative spatial representation and reasoning models. To make this mapping, authors developed a novel global machine learning framework for ontology population.

Finally, a more recent initiative, using also ontologies and semantic web technologies is the Digital 3D Reconstruction in Virtual Research Environment project [17], which aims to define standards for the web-based delivery, e-documentation and presentation of 3D data sets of destroyed architectural landmarks and artworks. The results are concerned with indexing of sources, documentation, semantic modelling, and visualization of 3D data sets using WebGL-technology. Here the main contribution is the development of the Cultural Heritage Markup Language (CHML), a human and machine-readable XML Schema for semantic annotation and for the digital 3D reconstruction of the lost and/or never existing Cultural Heritage 3D objects. The advantages of this new data model is that it is mapped to CIDOC CRM, which is the referent ontology in the Cultural Heritage domain, and is ready to use for annotating and indexing cultural objects within a Virtual Research Environment [18].

B. *The World City*

The concept of the World City was born in the period that goes from 1910 to 1941.This city should have been a spatial non-space, i.e., a "home for men" after war and a world centre for the accumulation, organization, and dissemination of knowledge. The idea was to create a place where people could live in harmony and in universal cooperation and a site whose urban organization responds to precise philosophical, scientific and rational criteria. The City should have included numerous buildings which are the World Museum (1) [19], the Hall of Modern Times (2), the International Association (3), the Library (4), the University (5), the Stadium (6), the Pavilions (7) and the cité hôtelière (8). In Figure 1, a plan realised by the architect Le Corbusier is shown to facilitate the understanding of the composition.

III. Materials and Methods

In the few above-mentioned works (see Section II-A), as well as in others, the 3D scenes are generated starting from sentences written by the users or manually selected from a text. On the contrary, nothing has been done with regard to the automatic extraction of descriptive sentences from a large corpus containing various information. Therefore, the challenge of the present work is to develop a system capable of selecting, without human help, specific sentences from texts concerning the "World City" with the final purpose of spreading the knowledge of the project thanks to the generation of 3D virtual scenes.

The approach can be subdivided in 5 different steps, which are listed below:

1) *Domain Analysis and Corpus compilation*: collection of the relevant sources related to the "World City";
2) *Information Extraction*: extraction of sentences which contain spatial description from the corpus;
3) *Semantic Annotation and Spatial Roles labelling*: sentences annotation and automatic detection of the spatial roles;
4) *3D Objects modeling*: creation of the models which compose the city through a graphical tool;
5) *3D Scenes generation*.

In this paper, we focused on points 1 and 2, namely, on the study of the domain and on the possibility of automatically extracting all the spatial information from the heterogeneous texts that compose the corpus. Further works will deal with the Semantic Annotation and the automatic 3D scenes generation.

C. Domain Analysis and Corpus Compilation

The domain analysis consists in the study of the architectural and historical sources present in the literature in order to discover the specific terms or the multi-words used in the architectural domain, with particular attention to the terminology used in the descriptions of the "World City". For this purpose, a corpus has been created focusing on two important aspects [20]:

- Definition of the characteristics of the reference population from which a significant sample was extracted;
- Definition of qualitative and quantitative criteria for determining the representativeness of the corpus.

The corpus, in fact, must fulfil the role of a representative sample, in the statistical sense of the term, because all the observations obtained from its analysis must be valid and extendable to all the individuals of the population. Specifically, the extracted terminology must be as representative as possible of the one used in the reference domain, but, on a smaller scale. One of the criteria that contributes to ensure the representativeness of the corpus is its size. However, there are not yet precise directives concerning the right dimension. Moreover, given the large amount of information and documents in all areas of knowledge, determining the number of all the available sources, for a precise domain, is a difficult and significant issue.

During the first phase, we collected sources of heterogeneous nature, which include images or technical drawings, videos and textual documents but, for our corpus composition, we only used the texts, because we are interested in the generation of 3D scenes starting from texts written in natural language. On these bases, aiming to respect the criterion of quality, we selected the texts which respect the following requirements [21]:

1. The texts should be representative of the period taken into consideration that goes from the beginning of the nineteenth century to the present day.
2. The texts should be written by different authors: firstly by Paul Otlet, but also domain experts both of the time and contemporary;
3. The texts should be original and not translations;
4. The texts should be complete and not fragments of the whole documents.

All these criteria should guarantee the maximum reliability of all the texts which form the corpus that is composed by primary sources (in the specific case, Paul Otlet's publications or the correspondence between him and the architects), and secondary sources (articles and books of other authors related to the referent domain).

At the end of the analysis, we collected 29 texts in total, for an amount of 411,401 tokens. These documents, originally of different formats (.doc, .pdf, .jpeg) have all been transformed into .txt files through the software ABBYY [22], able to provide optical character recognition and document capture. The whole corpus could be subdivided into four sub-corpora, one for each language handled. In particular, 17 documents are in English (201,747 tokens), 6 documents are in French (85,738 tokens), 4 documents are in Italian (105,930 tokens) and only two documents are in Spanish (17,986 tokens). Once the corpus was created, we carried out an analysis on the natural language constructions useful to talk about spatial configurations. The study was carried out on the four above-mentioned languages because we wanted to work on original texts and not on translations to ensure that the terminology is original and appropriate with reference to the specific domain.

D. Spatial Information Extraction

One of the most important functions of natural language is to describe spatial relationships between objects through linguistic constructs containing spatial information. The latter are easily understandable by the human mind, but machines, on the contrary, have not the same cognitive capabilities and they cannot distinguish spatial and non-spatial data from . This means that it is difficult for computers to identify and extract from texts only the information useful for 3D scenes construction, which is our final task. Therefore, we developed a sentence extractor capable of parsing large data to automatically extract specific sentences from the aforementioned corpus [23]. The aim is to provide a great help to people that, instead of reading the whole texts to select precise information, may automatically obtain the required data.

A way to guide the computer to a correct extraction is to identify how the spatial information is expressed in natural language. We ascertained that is is generally provided by the use of prepositions that establish a relationship between two or more objects However, it is not sufficient, because the same preposition could also be used to talk about events or situations which do not involve spatial descriptions. For example, the preposition *on* could be used to describe spatial configurations like "the picture is *on* the wall" or "the bottle is *on* the table" but, the same preposition is also used in different contexts such as in the sentences "They need to concentrate *on* their studies" or "The discussion will be *on* a topic you have studied recently".

The first two sentences express spatial concepts, concepts of verticality or objects overlapping, while the last two sentences do not have any kind of reference to entities locations. This means that the usage of prepositions depends on several aspects like the entities involved in the scene or the general situation of speech. In other words, prepositions are often used in front of nouns or pronouns to show the relationship between them and other words in the sentence, but they can also be used to describe the time when something happens (*They arrived* **on** *Sunday*), the way in which something is done (*We went* ***by*** *train*) and even more. On this basis, the identification and the extraction of particular sentences from a large number of texts, characterised by heterogeneous information, is a big and interesting challenge. Therefore, in this work, a Python script able to read and analyse a text has been created with the purpose of extracting all the sentences containing spatial descriptions. The latter are composed by three central concepts belonging to the Holistic Spatial Semantic Theory [24] in which the three main spatial roles are defined. They are trajector(s), landmark(s) and spatial indicator(s) that, linked together, generate a spatial triplet. In particular:

- the *trajector (TR)* is a spatial role label assigned to a word that denotes a central object of a spatial scene;
- the *landmark (LM)* is a spatial role label given to a word that indicates a secondary object of a spatial scene (to which a possible spatial relation between two objects can be established);
- the *spatial indicator (SI)* is a spatial role label assigned to a word that indicates a spatial relation between objects (TR and LM) of a spatial scene.

To better understand, in the sentence "*The* [*car*]$_{TR}$ *is* [*under*] $_{SI}$ *the* [*tree*] $_{LM}$" the "car" is a trajector, the "tree" is a landmark and "under" is a spatial indicator. A key element of the system is a set of text extraction rules that identify relevant information to be extracted from the input text. In particular, the adopted extraction process is based on some word lists belonging to a "spatial domain" and on specific rules which guide the system to identify only the sought information.

The system is subdivided in four sections which, by successive and interconnected steps, gradually lead to a more accurate result. The general idea is to start from an entire document, written in one of the four languages mentioned above, and to parse it in order to identify the possible spatial indicators and labels (TR and LM) existing and interrelated. Although in this step, a difference between roles is not carried out, they are considered regardless of whether they are a trajector or a landmark. The assignment of the correct roles, in fact, will be done subsequently (see step 3 of the methodology described in section III), by using a probabilistic programming language called Saul [25]. For this purpose, two main operations are executed by the system: i) the division of the whole text into single sentences and, ii) their analysis to discard, step by step, those that are not useful for the final goal. The four scripts of the system are the following:

1- *Sentence_Split.py* = the system opens a document, in a .txt format, and separates strings using two specific delimiters, a dot ('.') or a long sequence of blank spaces. Then, it creates a .txt file (*Sentences.txt*) that collects all the sentences of the document.

2- *Search.py* = the system reads, as input, a .txt file containing all the possible prepositions or expressions that may be spatial indicators (e.g., *on, in the centre of, to the right*, etc.). Then, it opens the *Sentences.txt* file and it analyses if each sentence contains or does not contain one or more spatial indicator(s). Finally, it creates, as output, a .txt file (*outfile.txt*) including only the phrases with spatial indicator(s) inside.

3- *CountOccurrences.py* = the system reads two .txt files, the previous *outfile.txt* and *domainNouns.txt*, which is a list containing terms or multi-words that may be TR(s) or LM(s) of the sentences. The latter are all written with lowercase letters so, the next step of the system is to replace all the capital letters, present in *outfile.txt* file, with lowercase letters in order to be case-insensitive. In the end, the system counts the number of labels occurrences (number of time that TR(s) and/or LM(s) appear in each sentence) and, if it is greater than two, it extrapolates the sentences from the *outfile.txt* file and writes them into another .txt file (*labelSentences.txt*). In this way, it is hoped that, at least, one label is the TR of the sentence and another one is its LM and that they are connected through a spatial indicator detected in the previous step.

4- *Last.py* = in this phase the system improves the results establishing which are the positions that can be occupied by the labels in order to be possible TR(s) and LM(s). The locations are set up with reference to the positions of the spatial indicators taking into account some sentence structure rules. In Table 1, the main rules adopted to identify a possible landmark in a sentence are illustrated. More specifically, by making reference to the location of the spatial indicator (signed as s_i) the LM may be located in the next position (s_i + 1), in (s_i + 2) or in (s_i + 3) position. Sample sentences are:

a) A [*zoo*] $_{TR}$ designed for children is located [*in the south of*] $_{SI}$ [*London*] $_{LM}$;

b) The [*fork*] $_{TR}$ is [*to the left of*] $_{SI}$ [*the*] $_{article}$ [*plate*] $_{LM}$;

c) Put the [*box*]$_{TR}$ [*on*]$_{SI}$ [*the*]$_{article}$ [*wooden*]$_{adjective}$ [*table*]$_{LM}$!

TABLE I. LANDMARK POSSIBLE POSITION

a	***SI***	LM		
b	***SI***	article	LM	
c	***SI***	article	adjective	LM
	SI Position: s_i	***LM position: s_i + 1***	***LM position: s_i + 2***	***LM position: s_i + 3***

A further case regards the lack of the landmark in the sentence. This happens when the description is characterized by the presence of an implicit LM like in the sentence: The [balloon] TR flown [up]SI where the LM is NIL (void). Once these rules have been taken into account, the system makes a screening selection of the input file (*labelSentences.txt*) and generates a final file (Last.txt),as output, which shall collect all the sentences containing spatial descriptions and spatial relationships. Table 2 summarizes the four abovementioned steps:

TABLE II SUMMARIZATION OF THE FOUR SCRIPT STEPS

Script	Action	Input	Output
1	**Subdivision** of the whole document into **single sentences.**	*Original document (File.txt)*	*Sentences.txt*
2	**Selection** of the sentences which **contain** one or more **spatial indicators.**	*Spatial_ indicator.txt* *Sentences.txt*	*Outfile.txt*
3	**Count** of the number of times that **words** representing TR or LM appear in each sentence. If num>2 the sentence is extracted.	*Domain Nouns.txt* *Outfile.txt*	*Label Sentences.txt*
4	**Application of rules**: assignment of the **positions** that can be occupied by the labels in order to be possible TR(s) and LM(s).	*Label Sentences.txt*	*Last.txt*

Since at step 3 of the methodology, the automatic assignment of roles (trajector, landmark) and spatial indicator in the sentences will be carried out using Saul, which works only for English, it has been necessary to translate of all the sentences extrapolated from the Italian, French and Spanish documents into English. More precisely, 146 sentences have been manually translated from French, 241 from Italian, and 48 from Spanish.

IV. RESULTS

In total, the system extrapolated 705 different sentences, originally in English (270 sentences) or translated into English (435 sentences), that may be classified into two different categories: "useful extractions" and "unuseful extractions". In particular, the sentences belonging to the first category are those that have been extracted by the system and that really contain spatial information, while, the ones belonging to the second category, are the sentences extracted even if they do not contain spatial data. The number of "useful extraction" sentences is 578, while, the amount of "unuseful extraction" sentences is equal to127 units. In addition, there is a third category that includes all the sentences not detected by the system but which contains spatial information; in other words, the missing data. They are, in total, only 75, as depicted in Figure 2.

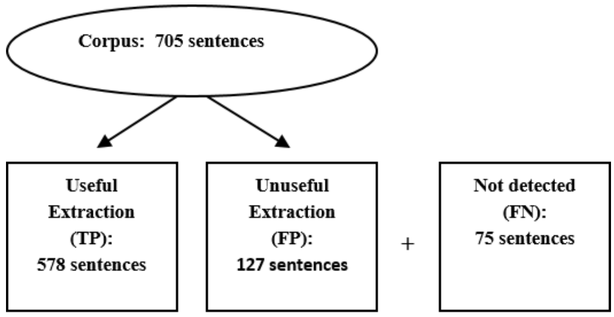

Figure 2. Schematization of the results.

The results have been estimated by the evaluation metrics of precision and recall defined as:

$$precision = \frac{TP}{TP+FP} \qquad recall = \frac{TP}{TP+FN}$$

where:
TP = is the number of system-extracted sentences that contain spatial information;
FP = is the number of system-extracted sentences that do not contain spatial information;
FN = is the number of sentences containing spatial information that the system does not extract from the texts.

The count of the exact number of sentences containing spatial information has been manually performed. They have been detected from the four *Sentences.txt* files (one for each language handled) and compared to the files automatically generated by the system: *Label Sentences.txt* and *Last.txt.* On this basis, the parameters of precision and recall have been estimated. Table 3 shows the results of the system run:

TABLE III. EVALUATION METRICS OF PRECISION AND RECALL

PRECISION	**0.820**
RECALL	**0.885**

These results certify that the used system has an adequate level of reliability because it is capable of responding to the task with a limited margin of error.

V. DISCUSSION AND CONCLUSION

Three-dimensional virtual reconstructions are used in various fields, such as playful or didactic, but play a predominant role in the Cultural Heritage sector. The 3D elements, in fact, allow complete and exhaustive knowledge of the object under consideration because they can be viewed from all points of view, rotated, inspected within them and exploded in their various components. Despite the multitude of 3D models designed to spread the knowledge of our Cultural Heritage, the innovative nature of this work is the connection between three-dimensional data and one-dimensional data. In other words, the will to materialize 3D scenes from descriptive sentences contained in texts of historical, artistic and cultural value that regard the project of the World City. Such interrelationship between reference textual sources and three-dimensional models would enrich the original texts with information that are often implied or scattered in the texts and cannot be viewed together. Furthermore, thanks to the system created and described in this work, it is possible to overcome the problem associated with the manual extraction of spatial descriptions that is a time consuming operation, which becomes more and more complicated in the growth of the size of the examined corpus. On the contrary, the automatic extrapolation of sentences to materialize in virtual space is the key to the problem and the innovation of the proposed work. These sentences will then become the starting point for virtual reconstructions of more or less complex scenes containing spatially related entities. Given the simple task evaluated, a future test will be to train a machine learning document classifier to see if the evaluation is improved with respect to the use of our rule-based approach.

The proposed methodology will provide a three-dimensional view of the World City, just as Paul Otlet and Le Corbusier designed it, and it will allow a simpler and quicker understanding of the complex project that, although it has great historical, artistic and cultural value, was never realized. Finally, we hope that the abovementioned methodology will be exploited in other contexts where there is a need to extrapolate sentences, containing entities and relationships between them, from documents. Especially when the documents, due to their large-scale, cannot be manually read and analysed by the user. Further work will test the potentiality of semantic web technologies and ontologies to improve the results as experimented in [18].

REFERENCES

[1] K. Charter, “Principles for conservation and restoration of built heritage”, Marsilio, Venice, 2000.

[2] G. Gresleri and D. Matteoni, “La città mondiale”. Andersen, Hébrand, Otlet, Le Corbusier, first edition, Polis/Marsilio Editori, 1982.

[3] P.Otlet, Cité Mondiale. Geneva : World Civic Center : Mundaneum, Publication n. 133 de l’Union des Associations Internationales, Palais Mondial, Bruxelles, 1929.

[4] T. Winograd, “Procedures as a representation for data in a computer program for understanding natural language (No. MAC-TR-84),” Massachusetts Inst of Tech Cambridge Project Mac, 1971.

[5] G. Adorni, M. Di Manzo and F. Giunchiglia, “Natural language driven image generation,” in Proceedings of the 10th International Conference on Computational Linguistics and 22nd annual meeting on Association for Computational Linguistics, pp. 495-500, 1984.

[6] B. Coyne and R. Sproat, “WordsEye: An Automatic Text-to-Scene Conversion System,” in Proceedings of the 28th annual conference on Computer graphics and interactive techniques (SIGGRAPH ‘01), ACM, New York, pp. 487–496,2001.

[7] L. M. Seversky and L. Yin, “Real-time automatic 3D scene generation from natural language voice and text descriptions,” in Proceedings of the 14th Annual ACM International Conference on Multimedia, 2006.

[8] M. Savva, A. X. Chang, G. Bernstein, C. D. Manning, and P. Hanrahan, “On being the right scale: Sizing large collections of 3D models,” in SIGGRAPH Asia, Workshop on Indoor Scene Understanding: Where Graphics meets Vision, 2014.

[9] A. Chang, W. Monroe, M. Savva, C. Potts and C.D. Manning, “Text to 3d scene generation with rich lexical grounding,” arXiv In section V concluding remarks are presented.preprint arXiv: 1505.06289, 2015.

[10] A. Chang, M. Savva and C. D. Manning, “Learning Spatial Knowledge for Text to 3D Scene Generation,” in Empirical Methods in Natural Language Processing (EMNLP), pp. 2028–2038, 2014.

[11] A. Cropper, “Identifying and inferring objects from textual descriptions of scenes from books,” in OASIcs- OpenAccess Series in Informatics (Vol. 43), Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2014.

[12] R. Sproat, “Inferring the environment in a text-to-scene conversion system,” in Proceedings of the 1st international conference on Knowledge capture, ACM, pp. 147-154, 2001.

[13] I. Mani, J. Hitzeman, J. Richer, D. Harris, R. Quimby and Ben Wellner, “SpatialML: Annotation Scheme, Corpora, and Tools”, in Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), Marrakech, Morocco, May 28-30, 2008.

[14] E. Klien and M. Lutz, “The Role of Spatial Relations in Automating the Semantic Annotation of Geodata, in Cohn, Anthony G. and Mark, David M. (eds.)”, in Proceedings of the International Conference on Spatial Information Theory 2005, Springer Berlin Heidelberg, pp. 133—148, 2005.

[15] L. Moncla and M. Gaio, “A multi-layer markup language for geospatial semantic annotations”, in Proceedings of the 9th Workshop on Geographic Information Retrieval, Paris, France, November 26-27, 2015, Article n. 5, ACM New York (USA), 2015.

[16] P. Kordjamshidi and M.F. Moens, “Global machine learning for spatial ontology population”, Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 30, January 2015, pp. 3-21, 2015.

[17] Digital 3D Reconstructions in Virtual Research Environments. Leibniz Association 2013-2016. [Online]. Available from: http://www.patrimonium.net. [retrieved: March, 2018]

[18] P. Kuroczynski, O. Hauck and D. Dworak, “3D Models on Triples Paths – new Pathways for documenting and Visualizing Virtual Reconstruction”, in 3D Research Challenges in Cultural Heritage II: How to Manage Data and Knowledge Related to Interpretative Digital 3D Reconstructions of Cultural Heritage, LNCS, Spriger Verlag, pp. 149-172, 2016.

[19] W. Van Acker, “Opening the Shrine of the Mundaneum: The Positivist Spirit” in the Architecture of Le Corbusier and his Belgian ‘Idolators’, in A. Brown & A. Leach (Eds.), Proceedings of the Society of Architectural Historians, Australia and New Zealand: 30, Open: Gold Coast: SAHANZ, Vol. 2, pp. 791-805, 2013.

[20] A. Caruso and A. Folino, “Corpus-based knowledge representation in specialized domains,” in Corpus-based studies on language verieties, Peter Lang, pp. 11-35, 2016.

[21] J. Pearson, “Terms in Context,” Amsterdam-Philadelphia: John Benjamins Publishing Company, pp. 58-62, 1998.

[22] ABBYY FineReader, URL: < https://www.abbyy.com/it-it/> [retrieved: March, 2018]

[23] F. Bonin et all, “A Contrastive Approach to Multi-word Extraction from Domain-specific Corpora,” in Proceedings of LREC'10, Valletta, Malta, 17-23, pp. 3222–3229, 2010.

[24] J. Zlatev, “Holistic spatial semantics of Thai, Cognitive Linguistics and Non-Indo-European Languages”, pp. 305–336, 2003.

[25] P. Kordjamshidi, R. Dan, and W. Hao "Saul: Towards Declarative Learning Based Programming." *IJCAI*, 2015.