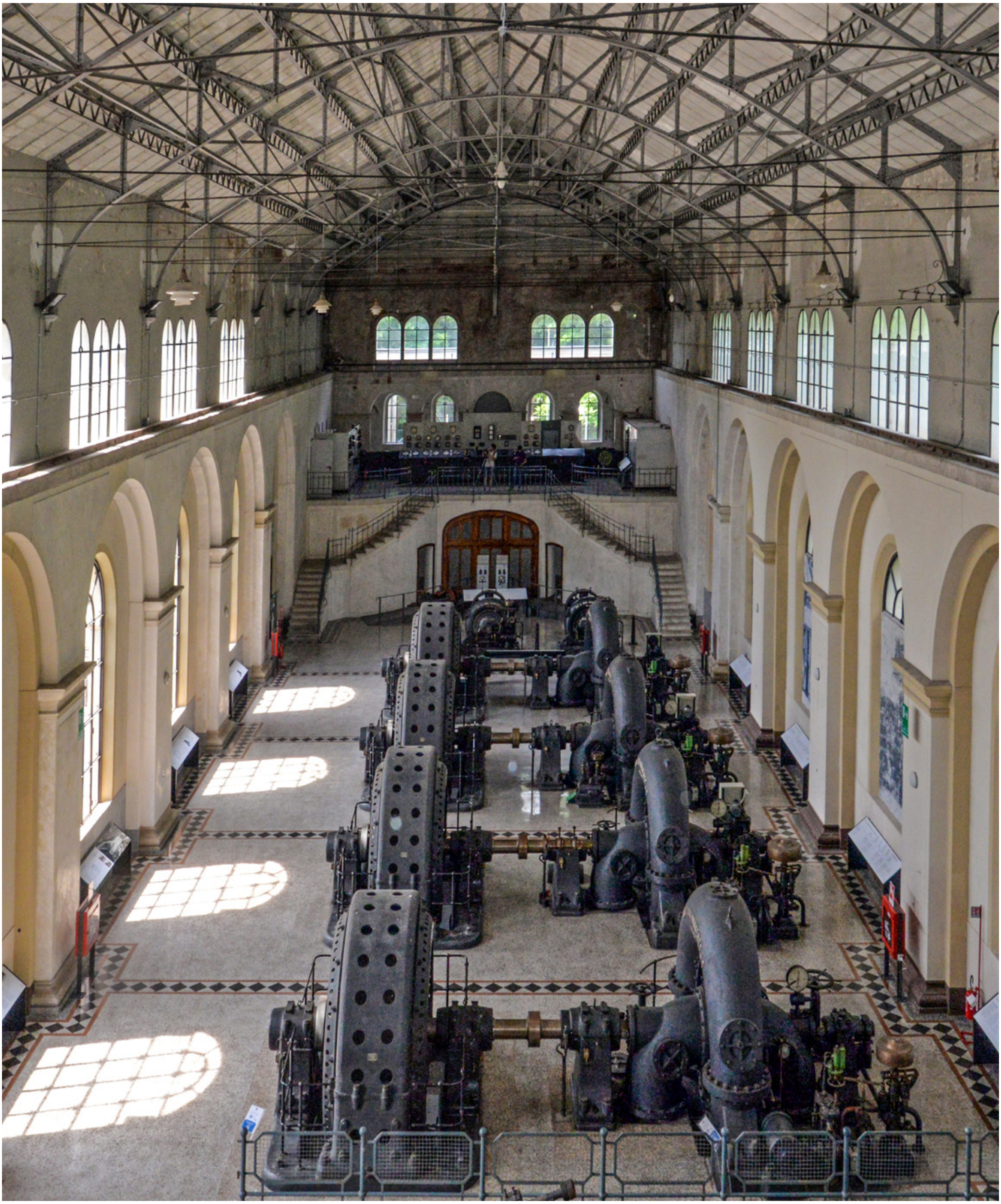
AUTOMATED SEMANTIC AND SYNTACTIC BIM DATA VALIDATION USING VISUAL PROGRAMMING LANGUAGE

Original
AUTOMATED SEMANTIC AND SYNTACTIC BIM DATA VALIDATION USING VISUAL PROGRAMMING LANGUAGE / Barbero, Andrea; Vergari, Riccardo; Ugliotti, Francesca Maria; Del Giudice, Matteo; Osello, Anna; Manzone, Fabio. - In: TEMA. - ISSN 2421-4574. - ELETTRONICO. - 7:N.2 (2021)(2021), pp. 122-133. [10.30682/tema0702m]

Publisher:
Edicom Edizioni

Published
DOI:10.30682/tema0702m



TEMA

Technologies
Engineering
Materials
Architecture

Journal Director: R. Gulli

e-ISSN 2421-4574

Vol. 7, No. 2 (2021)

Issue edited by Editor in Chief: R. Gulli

Cover illustration: Antonio Pitter power plant, interior view (Malnisio). © Francesco Chinellato, Livio Petriccione, 2019

Editorial Assistants: C. Mazzoli, D. Prati

TEMA
Technologies
Engineering
Materials
Architecture

e-ISSN 2421-4574
Vol. 7, No. 2 (2021)
Year 2021 (Issues per year: 2)

Editor in chief
Riccardo Gulli, Università di Bologna

Assistant Editors
Annarita Ferrante – Università di Bologna
Enrico Quagliarini – Università Politecnica delle Marche
Giuseppe Margani – Università degli Studi di Catania
Fabio Fatiguso – Università Politecnica di Bari
Rossano Albatici – Università di Trento

Special Editors
Luca Guardigli – Università di Bologna
Emanuele Zamperini – Università degli Studi di Firenze

Associated Editors
İhsan Engin Bal, Hanze University of Applied Sciences – Groningen
Antonio Becchi, Max Planck Institute – Berlin
Maurizio Brocato, Paris – Malaquais School of Architecture
Marco D'Orazio, Università Politecnica delle Marche
Enrico Dassori, Università di Genova
Vasco Peixoto de Freitas, Universidade do Porto – FEUP
Stefano Della Torre, Politecnico di Milano
Marina Fumo, Università di Napoli Federico II
José Luis Gonzalez, UPC – Barcellona
Francisco Javier Neila Gonzalez, UPM Madrid
Alberto Grimoldi, Politecnico di Milano
Antonella Guida, Università della Basilicata
Santiago Huerta, ETS – Madrid
Richard Hyde, University of Sydney
Tullia Iori, Università di Roma Tor Vergata
Raffaella Lione, Università di Messina
John Richard Littlewood, Cardiff School of Art & Design
Camilla Mileto, Universidad Politecnica de Valencia UPV – Valencia
Renato Morganti, Università dell'Aquila
Francesco Polverino, Università di Napoli Federico II
Antonello Sanna, Università di Cagliari
Matheos Santamouris, University of Athens
Enrico Sicignano, Università di Salerno
Claudio Varagnoli, Università di Pescara

Editorial Assistants
Cecilia Mazzoli, Università di Bologna
Davide Prati, Università di Bologna

Journal director
Riccardo Gulli, Università di Bologna

Scientific Society Partner:
Ar.Tec. Associazione Scientifica per la Promozione dei Rapporti tra Architettura e Tecniche per l'Edilizia
c/o DA - Dipartimento di Architettura, Università degli Studi di Bologna
Viale del Risorgimento, 2
40136 Bologna - Italy
Phone: +39 051 2093155
Email: info@artecweb.org - tema@artecweb.org

Media Partner:
Edicom Edizioni
Via I Maggio 117
34074 Monfalcone (GO) - Italy
Phone: +39 0481 484488

TEMA: Technologies Engineering Materials Architecture
Vol. 7, No. 2 (2021)
e-ISSN 2421-4574

CONSTRUCTION HISTORY AND PRESERVATION

CONSTRUCTION AND BUILDING PERFORMANCE

BUILDING AND DESIGN TECHNOLOGIES

AUTOMATED SEMANTIC AND SYNTACTIC BIM DATA VALIDATION USING VISUAL PROGRAMMING LANGUAGE

TEMA
Technologies
Engineering
Materials
Architecture
e-ISSN 2421-4574
Vol. 7, No. 2 - (2021)

Andrea Barbero, Riccardo Vergari, Francesca Maria Ugliotti, Matteo Del Giudice, Anna Osello, Fabio Manzone

DOI: 10.30682/tema0702m

Highlights

In this contribution, the Data Validation concept is introduced as an innovative methodological approach in terms of semantic and syntactic check to verify alphanumerical content, exploiting the flexibility of VPL tools (e.g., Dynamo). Data reliability and its usage during the building life cycle are crucial according to BIM procurement documents used by the owner to achieve BIM model uses. The conversion of standards in checklist and rulesets enrich the area covered by the current meaning of BIM Model Checking (BMC).

Abstract

Building Information Modeling (BIM) is part of a digitalization process that, in recent years, has been revolutionizing the way buildings and infrastructures are designed, built, and maintained. Compared to traditional processes, BIM enhances the production and the management of data related to buildings and infrastructures throughout their life cycle. It is founded on a three-dimensional graphical model based on the specificity of project goals following the "level of information need" defined in BIM procurement documents. In this framework, an automated process for checking information within a BIM model plays a role of fundamental importance. Although this increases the model's reliability, on the other hand, it decreases the time of working. Therefore, this research aims to develop a working methodology based on Visual Programming Language (VPL) for an automated BIM Data Validation process. This workflow aims to meet the growing need of owners to centralize data relating to their real estate assets to always have the appropriate one at the operational level. This methodology has been tested in different case studies to evaluate the strengths and weaknesses of using a standardization protocol in a large portfolio and complex buildings. This allows the huge amount of data from BIM models to be checked and summary reports to be produced, sharing with the various stakeholders involved in the knowledge process.

Keywords

Building Information Modeling, Data Validation, Visual Programming Language, Checklist, Rulesets.

Andrea Barbero*
DISEG – Dipartimento di Ingegneria Strutturale, Edile e Geotecnica, Politecnico di Torino, Torino (Italy)

Riccardo Vergari
DISEG – Dipartimento di Ingegneria Strutturale, Edile e Geotecnica, Politecnico di Torino, Torino (Italy)

Francesca Maria Ugliotti
DISEG – Dipartimento di Ingegneria Strutturale, Edile e Geotecnica, Politecnico di Torino, Torino (Italy)

Matteo Del Giudice
DISEG – Dipartimento di Ingegneria Strutturale, Edile e Geotecnica, Politecnico di Torino, Torino (Italy)

Anna Osello
DISEG – Dipartimento di Ingegneria Strutturale, Edile e Geotecnica, Politecnico di Torino, Torino (Italy)

Fabio Manzone
DISEG – Dipartimento di Ingegneria Strutturale, Edile e Geotecnica, Politecnico di Torino, Torino (Italy)

** Corresponding author:*
e-mail: andrea.barbero@polito.it

1. INTRODUCTION

BIM is becoming increasingly common in the Architectural Engineering Construction and Operational (AECO) sector. It improves information flow both for the delivery phase of a project (Project Information Model - PIM) and for the operational phase of an asset (Asset Information Model - AIM) [1, 2]. It is based on a shared and digitalized repository environment that collects geometric and alphanumeric information related to complex buildings and infrastructures [3]. In these terms, the paradigm behind this study is that BIM is all about the data [4], as the "I" of information is the central point to achieve management goals. For this reason, a BIM base system includes the fulfillment of numerous project requirements such as those dictated by building codes, regulations, standards, owner specifications [5]. In this framework, the key to success is the correct definition of what information is relevant according to the client's needs and their characteristics and structure.

Different meanings have been used in the international BIM context to indicate the "Level of Definition" by which information must be entered into a BIM model. The American Institute of Architects (AIA) first proposed LOD definition as Level of Detail, which later evolved into Level of Development [6]. In the UK, the PAS 1192-2 introduced the differentiation between the geometric Level of Development (LOD) and the alphanumeric one. It has split the LOD between the Level of Geometry (LOG) and the Level of Information (LOI). Italian norm defines instead the Level of Development (LOD) for each object to be inserted within the BIM model, mentioning the double definition of LOG and LOI [7].

Nowadays, the International Standard Organization (ISO) seeks to standardize this definition internationally with ISO 19650-1:2019 standard by introducing the concept of the "level of information need" concerning the specific BIM uses. Regarding the distinction between geometric and alphanumeric content, the ISO suggests that the granularity of alphanumeric information should be at least as important as that of geometric information [2].

Once defined, the level of information needed should be used to determine the requirements in procurement documents such as Employer Information Requirements (EIR), BIM Execution Plan (BEP), or BIM guidelines. Therefore, it has become essential to define procedures to ensure the correctness of information within BIM models in the contractual documents mentioned above. The whole framework is referred to in international literature as BIM Model Checking (BMC) [8].

This study aims to contribute to this research area by proposing a methodology to check the alphanumeric content defined and requested according to specific semantic and syntactic logics by setting a Dynamo tool to automatize the Data Validation process and easily extrapolate analyses and visualize data.

1.1. BACKGROUND AND LITERATURE REVIEW

BMC is expected to be one of the major contributors to BIM utilization in the AECO industry [9]. Nevertheless, a joint understanding of the different types of checking and the use of terms about BMC is still lacking [10]. BMC uses geometric and alphanumeric information as input to process different types of checking on BIM models. This activity is based on BIM models and algorithms for processing data against a set of predefined rules defined by BIM procurement documents.

This contribution starts from the reflections carried out by Hjelseth concerning BMC definition. The author proposed a framework for the classification of BMC concepts following an ontology-based analytical approach [11], [8], [10]. According to Hjelseth, the most commonly used concept type of BMC is "validation checking" [10]. This one aims to verify the quality of the information in the BIM model within defined logic rules. Rulesets can be based on different sources, such as norms, codes, standards, contracts, or other requirements. Figure 1 shows the process of validation checking based on processing predefined criteria where the outcome can be: "pass", "fail", or "not checked". The output of this process is generally a report of the detected errors to be shared with the stakeholders involved in the project, enriching collaboration, and communication.

Within this validation area, the most relevant checking method mentioned in the literature includes a geometry-based approach, best known as Clash Detection,

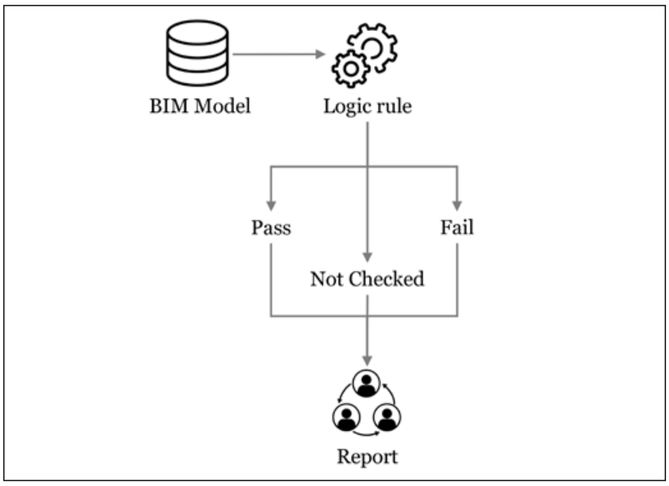


Fig. 1. *Logic of validation checking applied in this research.*

and an information-based one for verifying codes, regulations, and standards, best known as Code Checking. Clash detection aims to verify eventual clashes between geometric instances within BIM models [10], while Code Checking aims to control whether objects are following codes, regulations, and external standards [12].

Table 1 summarizes the most significant articles concerning our research area, listed in chronologic order and weighted based on a parameter of relevance for this contribution. An overview of the state of the art on BMC is provided by all the references with relevance 1, dealing with this matter of definition [13–15] [10] [11] [8] [16–19] [3] [20] [21]. The authors summarized different approaches and best practices on this topic, highlighting how the field of automated rule checking in the AECO sector using BIM models is an increasingly common practice. References classified with relevance 2 discuss VPL in the context of BMC, focusing the research on checking BIM models against codes, norms, standards, and regulations [22–25] [5] [26].

Based on this classification, checking the BIM database against a set of rules has been a significant topic in BIM research in the past few years [3]; this is because it is an uphill, costly, cumbersome, and error-prone task [11, 12]. Some model checking software already exists, but these solutions are often considered "black-boxes" in the literature since the output is not always clear to the user due to the lack of transparency on rules processing [22].

Although standards and codes differ from country to country, the alphanumeric content varies from project to project, even within the same country, because it depends on its purpose. Indeed, all stakeholders have different aims that imply the impossibility of using an immutable set of rules applicable to all contexts. At the same time, the development of a method able to check the correctness of the alphanumeric content within the BIM procurement specifications is challenging. For these reasons, the methodological aspect of the Data Validation activity concerned the identification of the most relevant information, according to the client's management, integration purposes, and the transposition of standards in specific rulesets. In this frame, VPL represents a way of programming using a code-based language by manipulating program elements graphically and alphanumerically rather than by specifying them textually [27].

1.2. CASE STUDIES

The methodology, shown in section 2, has been applied to different national research projects. The following have been chosen for their complexity in terms of the amount of data to be managed and analyzed regarding existing private and public infrastructure. Case A concerns a football stadium, case B refers to a high-rise building, while Case C considers a real estate portfolio owned by a municipality. For the modeling part, the Autodesk Revit platform has been selected due to its widespread use in the international market and its multidisciplinary nature. Specifically, the models examined refer to different disciplines (i.e., architectural, MEP, functional distribution of spaces) and may be articulated differently depending on the specific project.

Consequently, Autodesk Dynamo is used synergistically as an open-source computational design platform, supporting Data Validation activities through algorithms. It is grounded on the use of Python language and, for this reason, its potential is unlimited in the field of data analysis related to BIM models. Its visual, systemic, and geometric approach enhances collaboration between the parts of a project, from the process of shape concept to its management once it has been created.

The BIM models developed are designed to meet owners' real needs and set up advanced management systems. According to the specific model uses, operational

Author	Title	Year	Relevance	
			1	2
Eastman, Lee, Jeong, Lee	Automatic rule-based checking of building designs	2009	▨	
P. Pauwels, D. Van Deursen, R. Verstraeten, J. De Roo, R. De Meyer, R. Van de Walle, J. Van Campenhout	A semantic rule checking environment for building performance checking	2011	▨	
B.T. Zhong, L.Y. Ding, H.B. Luo, Y. Zhou, Y.Z. Hu, H.M. Hu	Ontology-based semantic modeling of regulation constraint for automated construction quality compliance checking	2012	▨	
J. Dimyadi, R. Amor	Automated Building Code Compliance Checking - Where is it at?	2013	▨	
W. Solihin, C. Eastman	Classification of rules for automated BIM rule checking development	2015	▨	
M. Preidel, A. Borrmann	Automated Code Compliance Checking Based on a Visual Language and Building Information Modelling		▨	▨
M. Preidel, A. Borrmann	Towards code compliance checking on the basis of a visual programming language	2016	▨	▨
A. Ciribini, Ventura, M. Bolpagni	Informative content validation is the key to success in a BIM-based project		▨	
E. Hjelseth	Classification of BIM-based model checking concepts		▨	
M. Preidel, Daum, A. Borrmann	Data retrieval from building information models based on visual programming	2017	▨	▨
V. Donato , M. Lo Turco, M. Bocconcino	BIM-QA/QC in the architectural design process		▨	
J. Reinhardt, M. Matthews	The Automation of BIM for Compliance Checking: A Visual Programming Approach		▨	▨
Ghannad, Lee, Dimyadi, Solihin	Automated BIM data validation integrating open-standard schema with visual programming language	2019	▨	▨
Sydora, Stroulia	Towards Rule-Based Model Checking of Building Information Models		▨	
Kim, Lee, Shin, Choi	Visual language approach to representing KBimCode-based Korea building code sentences for automated rule checking		▨	▨
Nawari O. Nawari	A Generalized Adaptive Framework (GAF) for Automating Code Compliance Checking		▨	
Y.C. Lee, C. M. Eastman, W. Solihin	The Mechanism and Challenges of Validating a Building Information Model regarding data exchange standards		▨	

Tab. 1. *Classification and analysis of BIM Model Checking references.*

BIM guidelines have been developed for each project to define the level of information needed as defined in the introduction, according to the specific model uses [2]. BIM guidelines have been structured straightforwardly, starting from the analysis and the application of the national and international standards mentioned before

In detail, Case A, represented by a football stadium, is characterized by many aspects that require specific maintenance activities. Architectural, structural, and space configurations are essential for the management of its operation and events. Particular attention is paid to the constituent aspects of modeling analysis in terms of defining the building registry and entities. Case B is part of a large urban transformation site consisting of 5 buildings. It is a new 209m – high public administration office building with 42 floors above ground and two underground. The building is characterized by a complex mechanical system consisting of plant engineering stations located in the technical rooms in the basement and above the building. The main lines serve the Air Handling Units (AHU) located on each floor and are connected to their terminals by a series of ducts from these stations. Case B has been selected because of its complexity due to numerous BIM objects being managed and verified. Starting from the vertical analyses of specific disciplinary domains carried out in the previous context, case C aims to define a basic dataset of information transversal over the entire building stock. It is crucial to adopt the same logic and rigorous coding and classification system to establish a consistent database to compare the data of different buildings across the board. As an example, the mapping of spaces is a key aspect for all management activities as well as the design or re-functionalization of an asset. 30 public buildings with different uses have been considered, including offices, schools, and archives [28]. In all these scenarios, the definition and validation of the alphanumeric content of each model become very important, defining a standardized validation procedure suitable for complex buildings, regardless of the nature of their owners.

Within the framework of the selected projects, a VPL-based procedure has been developed to investigate the Data Validation effectiveness and replicability.

2. METHODS

Given the increasing adoption of BIM in the construction sector, automated model checking is a pragmatic approach to expeditiously identifying errors that may otherwise cause issues later in the building life cycle. The main issues concerned identifying with the owner and the stakeholders involved in the digitalization process of meaningful information for the specific model uses. This phase implies the definition of appropriate requirements in the BIM procurement documentation to guarantee the effective usage, sharing, and integration with external platforms and domains of the BIM database. According to this scenario, the BMC procedure can be enriched with an additional declination and step, called Data Validation, which considers the internal standards of a project. In this way, it is possible to overcome the traditional checking approach based mainly on geometric features, Clash Detection, and external standards, Code Checking, used for different projects.

The innovative aspect of using the VPL for this validation activity lies in creating a flexible control tool, which could be defined as a white-box since it is possible to manage and implement every control, as the opposite of the "black-box" concept [22] mentioned before. The proposed BIM Data Validation process follows the structure illustrated in Figure 3. It can be configured as an iterative procedure to validate the compliance of the BIM model, where the main steps are comprised by:

- Identification of project's meaningful information.
- Definition of a checklist.
- Definition of rulesets.
- Processing of rules.
- Output generation.

2.1. PROJECT'S MEANINGFUL INFORMATION

As mentioned before, identifying the key information of a specific project contributes to the setting of internal standards, which are subject to validation. Within this check, both semantic and syntactic logics are included to verify the reliability of the information. Specifically, semantic

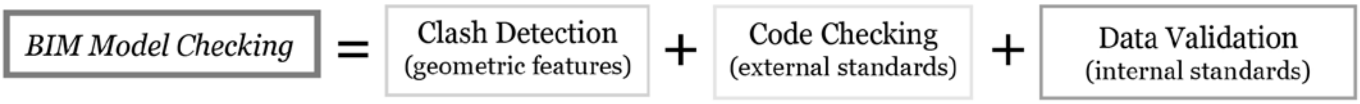


Fig. 2. *BIM Model Checking enrichment proposal.*

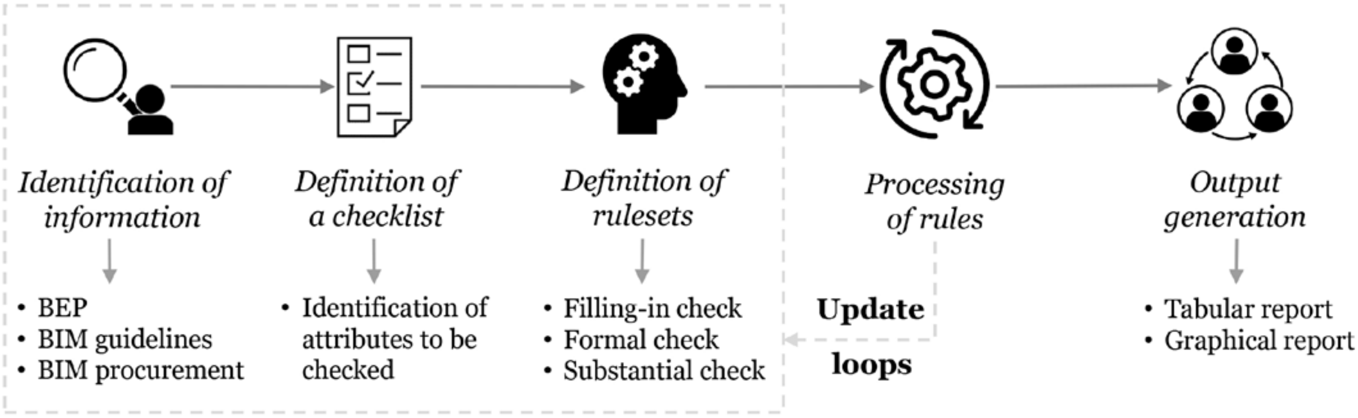


Fig. 3. *Methodology workflow for Data Validation process.*

rules for checking output consistency according to the BIM procurement documents can be introduced (checklist). It allows users to check whether the information contained in a specific parameter is relevant to an object. On the other side, the syntactic aspects are essential to properly manage data according to the specific type of checking (rulesets). For this reason, the declaration of the alphanumeric parameters associated with objects and how they should be compiled become an integral part of the procurement documents.

2.2. CHECKLIST

Checklists are used to identify what parameter and object relationships have to be checked within BIM models. For the research study, five checking clusters were identified: object naming, classification and coding, geometrical features, specific characteristics, and hierarchy. These groupings correspond, in fact, to the main specifications and parameters related to objects explained in the BIM documents. These have been defined to better manage a large amount of heterogeneous data and to simplify the semantic rules definition of the Data Validation process, highlighting the relevance of each information constituting the BIM database. From this setting, the attributes of the objects are then evaluated according to the disciplines, modeling software categories, and specific components.

The first group evaluates the terminology for each object to ensure coherence, for example, of the family and type name of the elements within each project standard. The classification and coding cluster includes all the relevant parameters for the unique classification and identification of the BIM database items. The check aims to verify the uniqueness of codes and their correct structure to avoid possible data loss during integration phases. The geometrical features overcome the clash detection activity by checking parameters used for management aspects. On the other side, the fourth cluster contains all the specific characteristics of each element, defined on the established information content. It is the broadest due to its specificity, and the semantics of the control are closely linked to the model use (e.g., maintenance management, simulations, quantity takeoff, etc.). Finally, the cluster hierarchy concerns the parameters required for system definition: the appropriate belonging of an object to a specific system typology or the relationship among components.

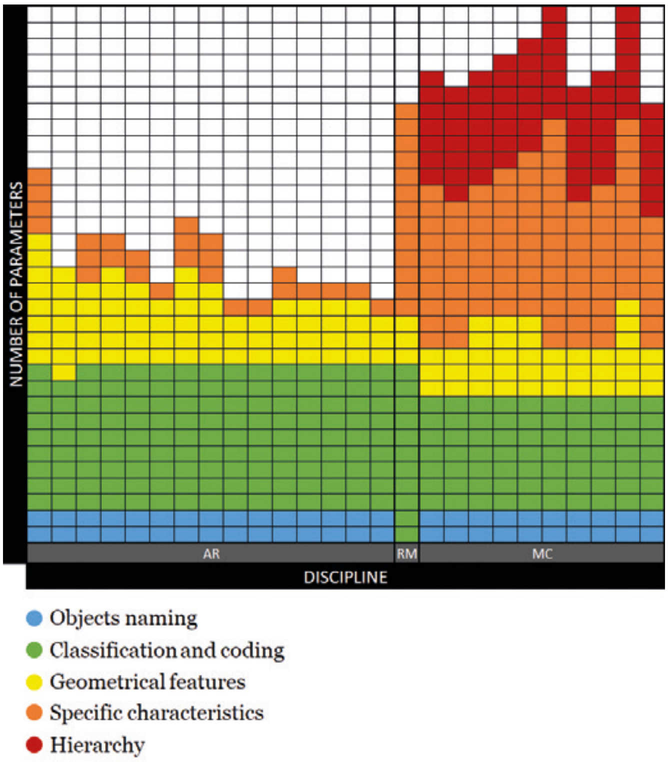


Fig. 4. *Graph of the developed checklist.*

Figure 4 shows grouping the attributes according to the discipline analyzed: the Architectural aspects (AR) for case study A, rooms (RM) for case study C, and the Mechanical (ME) for case study B. Object naming and classification, and coding are distributed for all the categories of objects related to the different disciplines concerning the cluster distribution. At the same time, geometrical features are more related to the categories of the architectural discipline, while system components are better characterized by specific characteristics and hierarchy, which include the functional relationships between objects.

2.3. RULESETS

Based on the checklist, which enables the structure of the validation process and identifies the parameters involved, rulesets for the BIM Data Validation process were defined for each parameter. These syntactic aspects were based on three main types of checking rules to manage data properly, moving from a general control to a detailed one:

- check *if the parameter has been compiled or not*;
- check whether the parameter has been filled out as provided for in the BEP or BIM guidelines. For this purpose, a Regular Expression (RegEx) must be defined by logical and/or mathematical constants and operators able to cover the different ways of populating alphanumeric content within the BIM model since there is no predefined naming for all the objects involved. This type of check could be defined as a *formal check*. An example could be the attributes "Type Name", for which is important to verify the compliance of the structure used in terms of letters, numbers, and symbols defined in BIM documents to guarantee proper sharing and integration.
  RegEx adopted: ^(E|I)(_)(S)[0-9]{1,2}$|^(E|I)(_)(S)[0-9]{2}(_)[0-9]{1,2}(x)[0-9]{1,2}$
- check if the parameter has been input accordingly to the specific content required by the BEP or BIM guidelines. The control rules do not admit possible contents for evaluating a parameter other than what is defined. This type of check could be defined as a *substantial check*, and it could be used, for example, for the objects' identification code for which it is necessary to respect exactly the value defined, enabling its usage.
  RegEx adopted for a specific discipline: TRP_TR_AR_LXX_AA_XXXXX.

The adoption of control rules was based on the importance of the single BIM model discipline parameter, and its definition should be done with the owner.

2.4. PROCESSING OF RULES

Using a native VPL environment in the BIM authoring platform enables a wide personalization of control rules based on the BIM model uses of the specific case study, directly querying the large amount of data. Moreover, developing a VPL tool enhances the replicability and the iterability of the process itself, allowing the possibility of

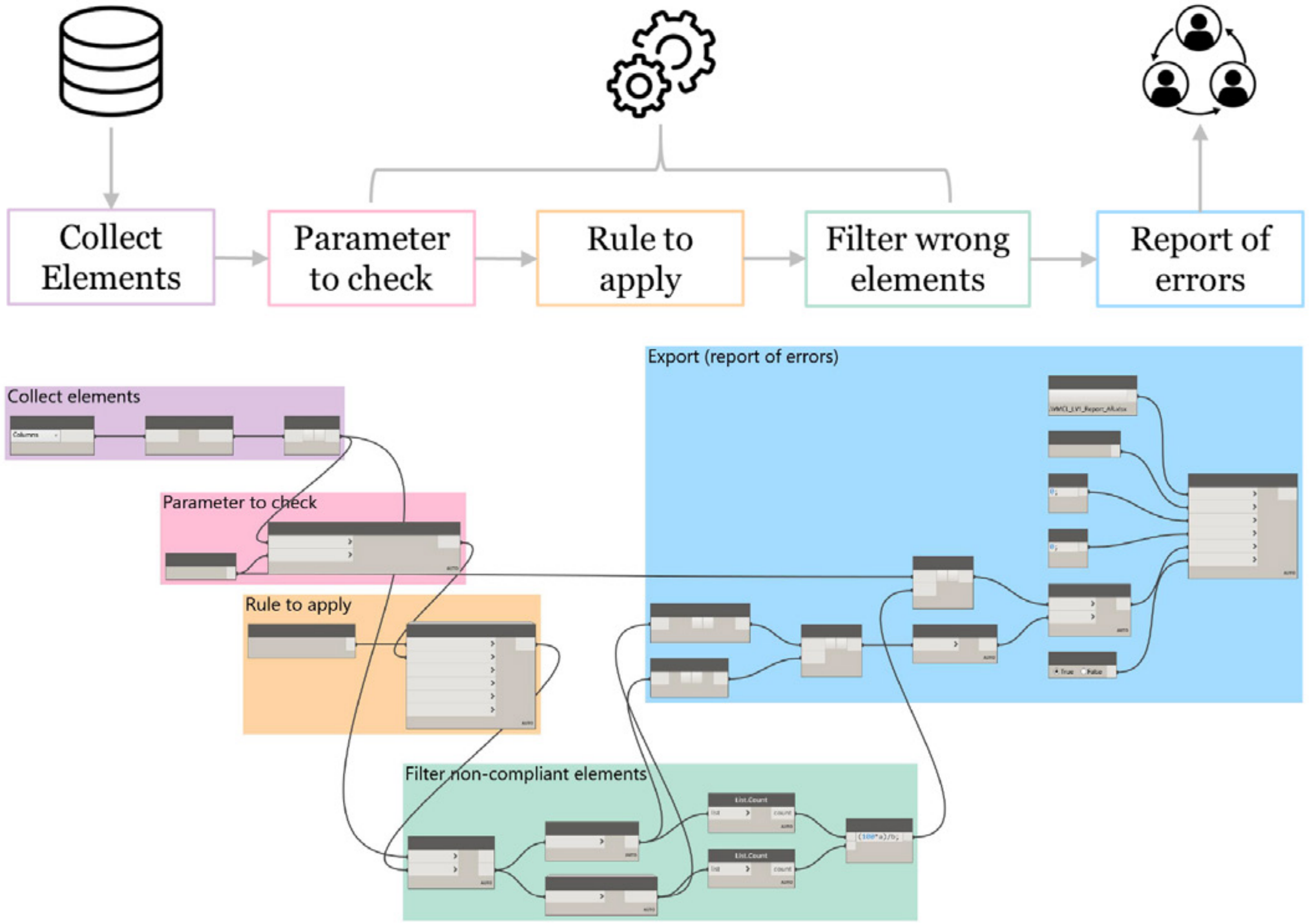


Fig. 5. *Logical pattern used for structuring the Data Validation algorithm within BIM models.*

changing rules depending on the BIM uses updating. The structuring and changing of rules are facilitated by the visual character of the VPL itself, which allows the stakeholder involved to view every single step of the control.

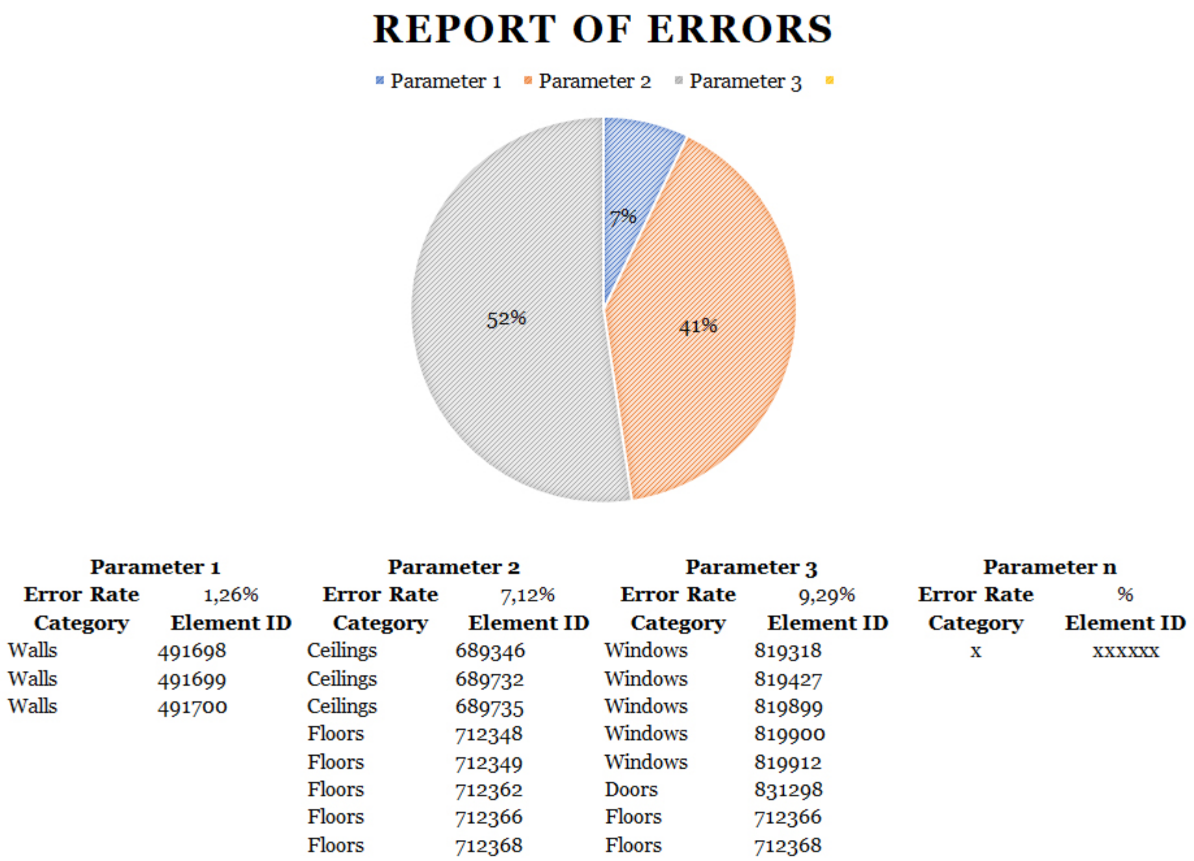
For the research study, the algorithm for Data Validation was created following the semantic and syntactic features of the checklist and rulesets. It can essentially be represented as a logical sequence comprised of five operations, as shown in Figure 5. The algorithm provides an output of non-compliant elements, Using a Boolean mask, listed and then exported in the errors report.

Figure 5 represents the methodological pattern used for each type of checking adopted within the proposed Data Validation process. The use of a pattern allows its customization and repeatability in the setting of more complex scripts. At the same time, this pattern has been accomplished through an iterative process, consisting of several subsequent test phases, which allowed defining it as a control tool for BIM models, also during the development phases. The same implemented logic could be used during the life cycle of the BIM database, identifying possible update loops as indicated in the methodological section due to model uses changes.

2.5. OUTPUT

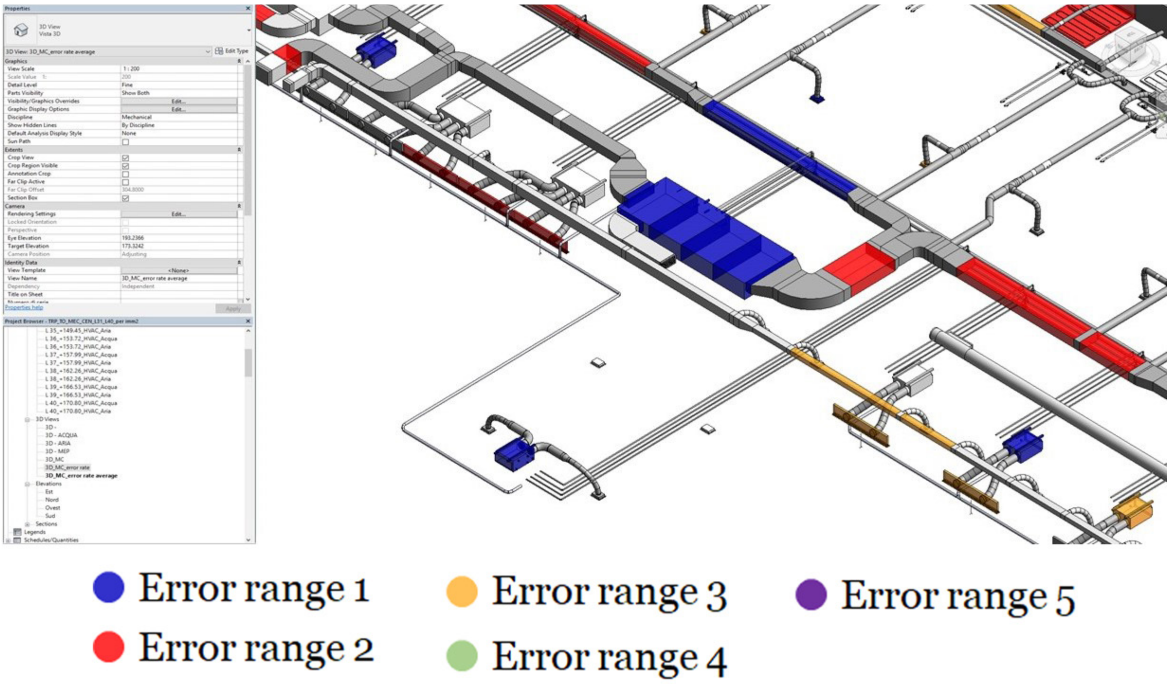
As mentioned before, the VPL approach was then used as a tool for data analysis, and the procedure was implemented to display the results obtained. Output within a VPL environment could be implemented in different ways: in this research, creating an algorithmic automatic generation of tabular and graphical reports was developed, managing a large amount of data generating PDF reports automatically. Output strategies can be followed directly from the control script to visualize the results of the checks, as shown in the image below. The outputs obtained are characterized by the unique identifier of every single disciplinary object that results as non-compliant with the syntactic structuring.

Figure 6 highlights a meaningful overview of the Data Validation activity, summarizing the alphanumeric content distribution according to a specific category


REPORT OF ERRORS

Parameter 1		Parameter 2		Parameter 3		Parameter n	
Error Rate	1,26%	Error Rate	7,12%	Error Rate	9,29%	Error Rate	%
Category	Element ID	Category	Element ID	Category	Element ID	Category	Element ID
Walls	491698	Ceilings	689346	Windows	819318	x	xxxxxx
Walls	491699	Ceilings	689732	Windows	819427		
Walls	491700	Ceilings	689735	Windows	819899		
		Floors	712348	Windows	819900		
		Floors	712349	Windows	819912		
		Floors	712362	Doors	831298		
		Floors	712366	Floors	712366		
		Floors	712368	Floors	712368		

Fig. 6. *Example error report obtained as output.*



Properties

3D View
Vista 3D

3D View: 3D_MC_error rate average — Edit Type

Graphics	
View Scale	1 : 200
Scale Value 1:	200
Detail Level	Fine
Parts Visibility	Show Both
Visibility/Graphics Overrides	Edit...
Graphic Display Options	Edit...
Discipline	Mechanical
Show Hidden Lines	By Discipline
Default Analysis Display Style	None
Sun Path	☐
Extents	
Crop View	☑
Crop Region Visible	☑
Annotation Crop	☐
Far Clip Active	☐
Far Clip Offset	304.8000
Section Box	☑
Camera	
Rendering Settings	Edit...
Locked Orientation	☐
Perspective	☐
Eye Elevation	193.2366
Target Elevation	173.3242
Camera Position	Adjusting
Identity Data	
View Template	<None>
View Name	3D_MC_error rate average
Dependency	Independent
Title on Sheet	

Properties help — Apply

Project Browser - TRP_TO_MEC_CEN_L31_L40_per imm2

- L 35_+149.45_HVAC_Aria
- L 36_+153.72_HVAC_Acqua
- L 36_+153.72_HVAC_Aria
- L 37_+157.99_HVAC_Acqua
- L 37_+157.99_HVAC_Aria
- L 38_+162.26_HVAC_Acqua
- L 38_+162.26_HVAC_Aria
- L 39_+166.53_HVAC_Acqua
- L 39_+166.53_HVAC_Aria
- L 40_+170.80_HVAC_Acqua
- L 40_+170.80_HVAC_Aria
- 3D Views
  - 3D -
  - 3D - ACQUA
  - 3D - ARIA
  - 3D - MEP
  - 3D_MC
  - 3D_MC_error rate
  - **3D_MC_error rate average**
- Elevations
  - Est
  - Nord
  - Ovest
  - Sud
- Sections
- Legends
- Schedules/Quantities

Fig. 7. *Graphical representation of elements which belong to a range of errors.*

and disciplines in addition to the error rate. Besides, the possibility of thematizing objects in the BIM models was tested using the number of errors or ranges of errors (Fig. 7). This solution offers an easy understanding of the validation results through a user-friendly graphical representation of the whole set directly inside the parametric model.

3. RESULTS

The Data Validation concept, related to the alphanumeric content of BIM models, was introduced to enrich the definition of BIM Model Checking as the sum of Code Checking, Clash Detection, and Data Validation. This control becomes fundamental when data and its reliability are crucial to achieving particular BIM model uses based on specific internal standards. For this reason, the Data Validation process is driven by compliance with the BEP, the guidelines, or the project's contractual documents.

This approach suggests a vertical investigation based on a specific disciplinary domain regarding the cross-cutting nature of Code Checking, founded on common standards and legislation requirements. It is based on the definition of semantic rules and syntactic aspects, which allow identifying, at the same time, the information that should be check and the checking type to be adopted. As discussed in this contribution, the Data Validation does not consider the geometry connected to the modeling activities but limits the scope only to informative alphanumeric content defined for the specific use of BIM models, and its outputs can be analyzed from multiple points of view. It is possible to have the percentage of error detected for each parameter analyzed, indicating the category and the unique identifier of the object. On the other hand, the different categories can be grouped according to the predominant discipline they belong to, thus obtaining the percentage of error of the single discipline.

The meaning of the error rate obtained after the proposed activity can have a double semantic or syntactic value. Suppose it is analyzed from a semantic point of view. In that case, its value allows the stakeholders to understand the coherence of the checked information concerning the examined object, allowing identifying possible errors in the structuring of the control itself. On the other side, the syntactic value, represented in the graphs of Figures 8 and 9, allows identifying the information that, properly controlled for an object type, does not comply with the defined rulesets. This type of error rate was considered the most frequent because it is connected to errors that occurred during the alphanumeric population, compared to

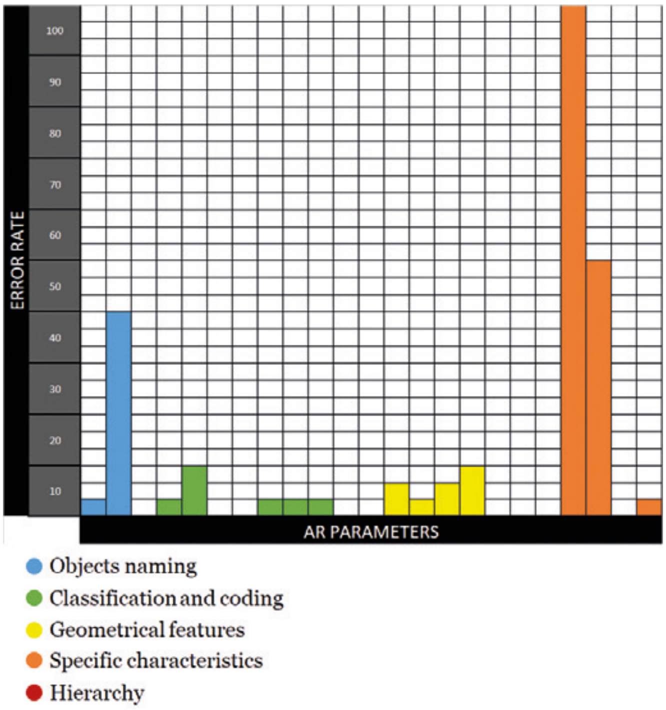


Fig. 8. *Error rate of AR parameters.*

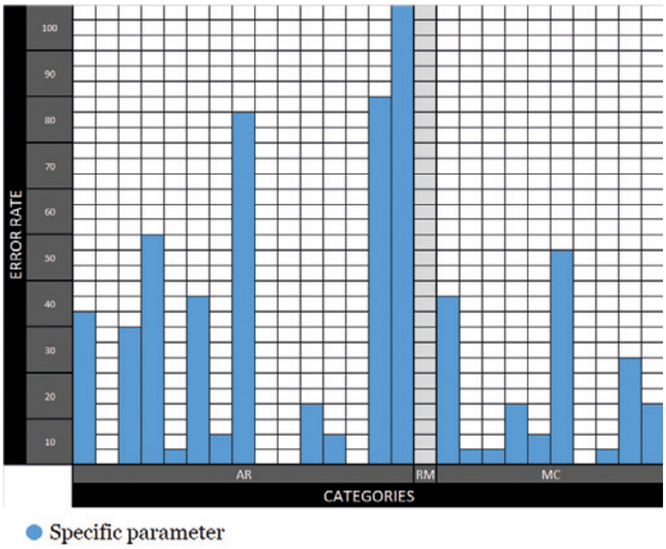


Fig. 9. *Error rate of a given parameter.*

the standards set out in the BIM procurement documents, which affect the quality of the data.

As an example, Figure 8 shows the percentage of error obtained while considering all the architectural objects present in the case study models concerning the overall parameters checked for this discipline. The used colors refer to the parameter clusters introduced for the checklist definition. The 100% error rate indicates that the parameter has not been compiled or has been filled in incorrectly for all the architectural objects and, consequently, the corresponding categories. On the contrary, 0% indicates that the parameter is compiled in compliance with the documents' dispositions for all objects with which it is associated. It can be observed that the adoption of rigorous standards leads to limit errors, given the considerable number of objects present in complex models such as those considered. Through this control mode, it is easier to detect systematic compilation errors and forgetfulness.

Compared to the distribution of Figure 8, the representation in Figure 9 provides a detail of the error rate of a given parameter with greater granularity, taking into account the different categories related to the disciplines analyzed. The percentage needs to be interpreted considering the number of objects present in the single model for each category, defining its importance.

Indeed, a certain number of parameter values compiled incorrectly generates a higher percentage for a limited number of objects related to a specific category (i.e., stairs). At the same time, the same produce a lower rate if evaluated on a considerable number of objects (i.e., walls). In Figure 9, for example, the "Type Name" attribute of the objects was chosen as it is considered one of the most significant for the case studies. This is not associated only with the room category as it is not applicable and therefore not checked in the validation.

Another significant result is the flexibility of the methodology proposed thanks to a VPL tool that allows Data validation to be performed for subsequent control steps, checking a selection of information from time to time. It also allows to manage and implement every control, and it could be configurable as an iterative process, as shown in the methodological diagram in Figure 2. The starting point of the procedure, which is composed of the project's meaningful information identification, the checklist, and the corresponding rulesets, evolved during the model life cycle depending on its uses, and the VPL tool can change in a flexible way to reflect these changes.

4. CONCLUSIONS

The study addresses the concept of Data Validation that will become increasingly important in the next few years also for the AECO sector. The process illustrated aims to ensure the correctness and meaningfulness of the input data using routines. In the age of big data, information acquires an economic value according to its availability and reliability. For this reason, this aspect should be managed from the beginning by drawing up detailed disposition in BIM procurement documents from which a rigorous Data Validation process can be activated.

In this way, the semantic and syntactic logic associated with the Data Validation process becomes part of the BIM procurement documents. Moreover, the introduction of this concept also allows professionals to achieve the UNI 11337 verification levels (LV1 and LV2) concerning the information content. The proposed methodology improves the quality and correctness of the relevant data according to the level of information need established by the specific BIM model uses. This process, based on a VPL tool, provides several advantages, including: (i) the possibility to control each validation step; (ii) replicability of control operations; (iii) full customization of the control algorithm structure based on the internal standards of the project; (iv) the creation of a validation process suitable for new and existing buildings; (v) the application for both integrated and federated models.

In addition, starting from the revision of the validation process mentioned above, a new research line could be launched to create BIM guidelines and, consequently, checklists and rulesets using a machine-encodable computer language. These can then be used as input for automatic data processing, optimizing the entire workflow illustrated during this research study.

5. AUTHOR CONTRIBUTIONS

Conceptualization, all authors; Abstract, all authors; Highlights, all authors; Introduction, F.M.; Background and literature review, R.V.; Case studies, M.D.G.; Method, all authors; Project's meaningful information, F.M.U.; Checklist, A.B; Rulesets, A.B.; Processing of rules, R.V.; Output, M.D.G; Results, F.M.U.; Conclusions, A.O.

6. REFERENCES

[1] Osello A (2015) Building Information Modelling, Geographic Information System e Augmented Reality per il Facility Management. Flaccovio, Palermo

[2] SIST EN ISO 19650-1:2019 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) - Information management using building information modelling

[3] Sydora C, Stroulia E (2019) Towards Rule-Based Model Checking of Building Information Models. 2019 Proceedings of the 36th ISARC, Banff, Canada. https://doi.org/10.22260/ISARC2019/0178

[4] Teicholz P (2013) BIM for Facility Managers. John John Wiley & sons, Hoboken, NJ

[5] Ghannad P, Lee Y-C, Dimyadi J, Solihin W (2019) Automated BIM data validation integrating open-standard schema with visual programming language. Advanced Engineering Informatics: 14–28

[6] Bolpagni M (2016) BIM Think Space, 19 July 2106. https://www.bimthinkspace.com/2016/07/the-many-faces-of-lod.html. Accessed 23 January 2020

[7] UNI 11337:4, 2017

[8] Hjelseth E (2015) BIM-based model checking (BMC). Building Information Modeling: Applications and Practices: 33–61

[9] McGraw-Hill (2014) The business value of BIM for Construction in major global markets. McGraw Hill Construction

[10] Hjelseth E (2016) Classification of BIM-based model checking concept. ITcon - Journal of Information Technology in Construction 21:354–369

[11] Hjelseth E (2010) Overview of concepts for model checking. Proceedings of the CIB W78 2010: 27th International Conference, Cairo, Egypt

[12] Nawari NO (2018) Building Information Modeling: Automated Code Checking and Compliance Processes. CRC Press, Boca Raton, FL

[13] Eastman C, Lee JK, Lee JM, Jeong Y-S (2009) Automatic rule-based checking of building designs. Automation in construction 18:1011–1033

[14] Pauwels P, Van Deursen D, Verstraeten R, De Roo J, De Meyer R, Van de Walle R, Van Campenhout J (2011) A semantic rule checking environment for building performance checking. Automation in Construction 20(5): 506–518

[15] Zhong BT, Ding LY, Luo HB, Zhou Y, Hu YZ, Hu HM (2012) Ontology-based semantic modeling of regulation constraint for automated construction quality compliance checking. Automation in Construction 28:58–70

[16] Dimyadi J, Amor R (2013) Automated Building Code Compliance Checking - Where is it at?. 19th International CIB World Building Congress

[17] Solihin W, Eastman C (2015) Classification of rules for automated BIM rule checking development. Automation in Construction 53:69–82

[18] Ciribini ALC, Ventura SM, Bolpagni M (2015) Informative content validation is the key to success in a BIM-based project. Territorio Italia 2

[19] Donato V, Lo Turco M, Bocconcino M (2018). BIM-QA/QC in the architectural design process. Architectural Engineering and Design Management 14(3): 239–254

[20] Nawari NO (2019) A Generalized Adaptive Framework (GAF) for Automating Code Compliance Checking. Buildings

[21] Eastman C, Solihin W, Lee Y (2019) The Mechanism and Challenges of Validating a Building Information Model regarding data exchange standards. Automation in Construction 100:118–128

[22] Preidel C, Borrmann A (2015) Automated Code Compliance Checking Based on a Visual Language and Building Information Modeling. ISARC 2015, Oulu, Finnland

[23] Preidel C, Borrmann A (2016) Towards code compliance checking on the basis of a visual programming language. Journal of Information Technology in Construction 21:402–421

[24] Preidel C, Borrmann A, Daum S (2017) Data retrieval from building information models based on visual programming. Visualization in Engineering 5(1)

[25] Reinhardt J, Mathews M (2017) The Automation of BIM for Compliance Checking: a Visual Programming Approach. Dublin, CITA BIM Gathering

[26] Kim H, Lee J, Shin J, Choi J (2019) Visual language approach to representing KBimCode-based Korea building code sentences for automated rule checking. Journal of Computational Design and Engineering 6(2): 143–148

[27] Amoruso G (2016) Handbook of Research on Visual Computing and Emerging Geometrical Design Tools. Information Science Reference, Hershey PA, USA

[28] Osello A, Ugliotti FM (2017) BIM: verso il catasto del futuro. Conoscere, digitalizzare, condividere. Il caso studio della Città di Torino. Gangemi, Roma