GEOMATICS AND ENVIRONMENTAL ENGINEERING • Volume 19 • Number 2 • 2025

https://doi.org/10.7494/geom.2025.19.2.99

Jędrzej Pasalski[1], Tomasz Owerko[2], Dariusz Kasznia[3], Maciej Wrzosek[4], Maciej Rochel[5], Kamil Czerw[6]

Efficient File-Naming Convention for Data Exchange in Construction Industry

Abstract: This paper aims to find common and universal naming convention for files to allow for the efficient exchange of data within the construction industry. Since building information management (BIM) has become a norm on the market and new technologies are rapidly driving the concept, there is a noticeable lack of unification in the deliveries of projects. The analysis focuses on exploring the most common naming conventions that have been mostly defined in the national standards and explores the widely adopted British National Annex 2 to BS EN ISO 19650-2:2018. Part of this research is a survey that was undertaken among professionals to understand the needs and habits that are related to file-naming. Furthermore, the ergonomics, comfort of use, and information capacity of the naming system were examined. Based on all of these, the authors suggest an optimal file-naming convention for building construction projects, which should help the industry in efficiency gains. In the future, there is the potential of developing a form to define project needs at the very start or tools for checking naming compliance with a proposed structure.

Keywords: naming convention, standardization, file-naming, BIM, BS EN ISO 19650, nomenclature

Received: December 23, 2024; accepted: April 7, 2025

[1] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, email: jpasalski@agh.edu.pl (corresponding author), https://orcid.org/0009-0005-4430-1248

[2] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, email: owerko@agh.edu.pl, https://orcid.org/0000-0002-2873-3535

[3] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, email: kasznia@agh.edu.pl, https://orcid.org/0009-0006-0942-7854

[4] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, email: mwrzosek@agh.edu.pl, https://orcid.org/0009-0000-6317-5631

[5] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, email: mrochel@agh.edu.pl, https://orcid.org/0000-0002-8109-973X

[6] independent researcher, Poland, email: kamilczerw@gmail.com, https://orcid.org/0009-0000-5792-4122

1. Introduction

1.1. Overview of Problem

In the context of dynamic technological progress, automation, and artificial intelligence, the construction sector seems to lag behind the general trend of growth and prosperity. This was proven in the research in [1] (represented with Figure 1), where it is clear that the total economy and (especially) manufacturing are growing much faster than the construction industry.

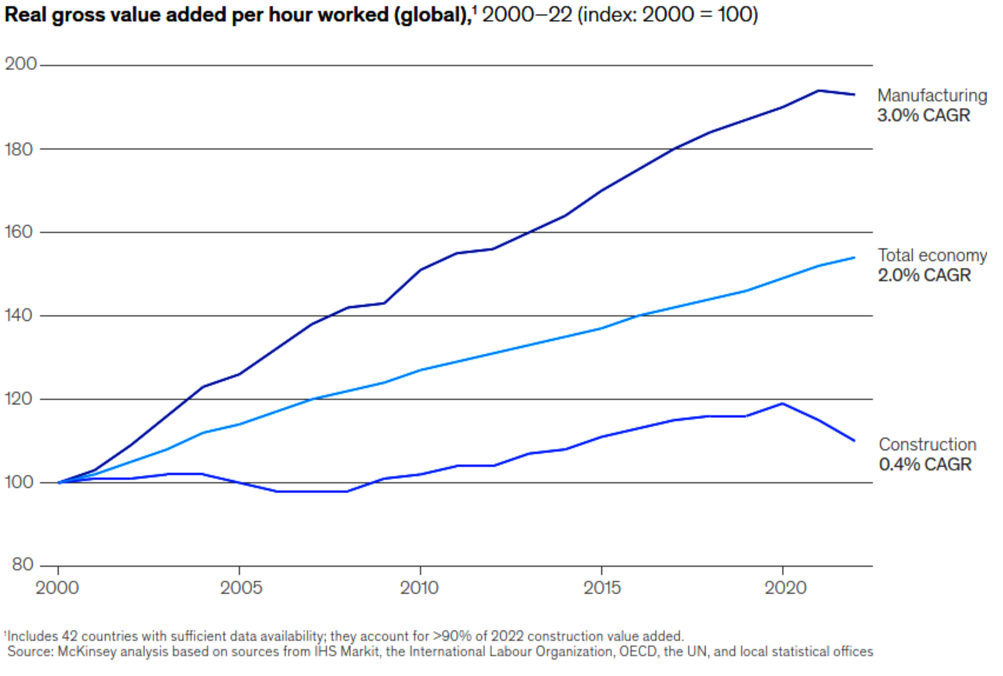

Fig. 1. Global productivity growth trends

Source: [1]

The industry's specificity, which is mainly based on producing unique construction and infrastructure objects and providing services at various stages of their lifecycles, makes it difficult to introduce automation (at least not at the same scale as in manufacturing, where the repetitiveness is on a much higher level). The lack of repeatability limits automation possibilities, contradicting the fundamental goal of construction investments (which is to generate profits).

It has been successfully proven in many fields that the introduction of some degree of unification in the industry could enable standardization; this, in turn, would pave the way for savings through automation, thus contributing to increased efficiency. Although this seems to be challenging in the context of the architecture or infrastructure itself, such possibilities exist in the realm of processes that accompany the executions of construction projects. Considering each stage of a construction project, an object's lifecycle allows for identifying repeatable mechanisms. Their standardization and support with modern technologies such as building information management (BIM) can significantly contribute to project efficiency.

Despite the intense technological development in recent decades, standardization – crucial for effective automation – has not evolved as dynamically. This led the authors of this article to look for patterns, as the lack of them hampers the efficient use of any available tools; this poses a challenge for the industry to identify regularities and introduce the necessary unification. This has been proven by engineers, who claim that efficient data-management is the most challenging aspect that affects time, cost, and quality. As per the research results from [2] (represented in Figure 2), the most problematic action that they must undertake in the design process is related to the searching and retrieval of data; this is purely about data-management.

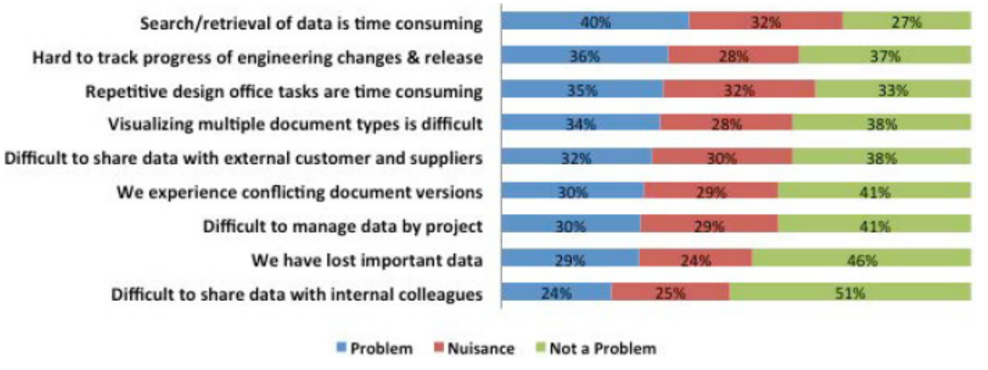

Fig. 2. Severity of design-data-management challenges
Source: [2]

One of the areas where attempts have been made to introduce standardization is the level of the advancements of models or component nomenclature [3, 4]. This first aspect of the level of advancement is covered within the level of detail, development, or definition [5]. The second aspect of the component nomenclature is also well-explained in many different ways, including with Uniclass, Omniclass, and other classifications [6, 7]. A lot is already been said about the approach for the naming of objects considering the volumetric approach [8]. There are also national proposals for object naming, like in Portugal [9]. What is not completely and unequivocally resolved, however, is very basic aspect of the naming procedure (or nomenclature) for files. These are referred to in BIM as information containers in [10]. Despite various attempts to adapt naming standards for files, there has been no globally accepted solution. The vast majority of large organizations work according to some standards in order to allow for unified processes; this makes sense for investors who can expect and impose the use of their naming conventions by all of the appointed parties that are involved in a project. For other parties (like designers and contractors), this means adjustments of internal standardizations for every single client whom they serve. Theoretically, this is not that problematic; when familiarizing oneself deeper into the technical side of things; however, there is a requirement to change parts of template documents and models and to teach people who are new to a scheme about all of the methods and procedures and new rules that have been applied. This may lead to unintended mistakes and errors.

The simplest example from the design phase is the necessity to change title blocks for drawings in the documentation. Figure 3 shows such a title block, where the naming of a document (which then becomes the filename) is based on the National Annex (NA) [11]. Each field represents information of a document's content and is created as separate metadata that is provided while creating a drawing as the values of the parameters (their names are provided above the values). Separators (hyphens in this case) are embedded in a title block, while all of the other documents in a project follow the same rules of naming – the order of fields, the number of characters for each of them, and any information behind the coding.

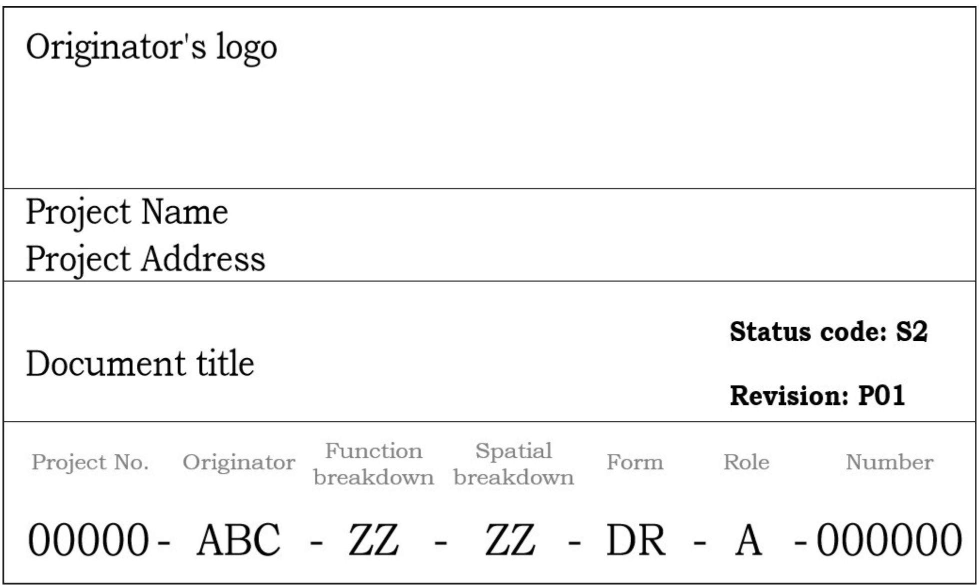

Fig. 3. Example of title block in British project

Thanks to such a unified approach, data can be managed in an efficient way on the common data environment (CDE) platform; this should be provided by the appointing party (as stated in [11]). Together with CDE, an appointing party provides exchange information requirements – a document that contains the standards, methods, and procedures for a project (which may include the nomenclature for the files). If this differs from the above rules, then each party that is involved in the project must change their internal settings; this is time-consuming and requires many additional actions.

An example of the same title block with some other rules for naming is presented in Figure 4; this is specifically related to the number of characters that has changed; the order of the fields and the information that they encode is the same. This is one of the most frequent scenarios, as the length of the fields is not specified in [11]. Nevertheless, it must be stressed that other scenarios (including more-significant changes) are also noticeable on the market. This results not only in the aesthetics (as shown below) but, most importantly, in the discomfort and inefficiency of use, the need for setting up or adjusting CDE platforms when starting each project, and the changes in the automating scripts that are used by the involved parties for each project individually (or internal document templates).

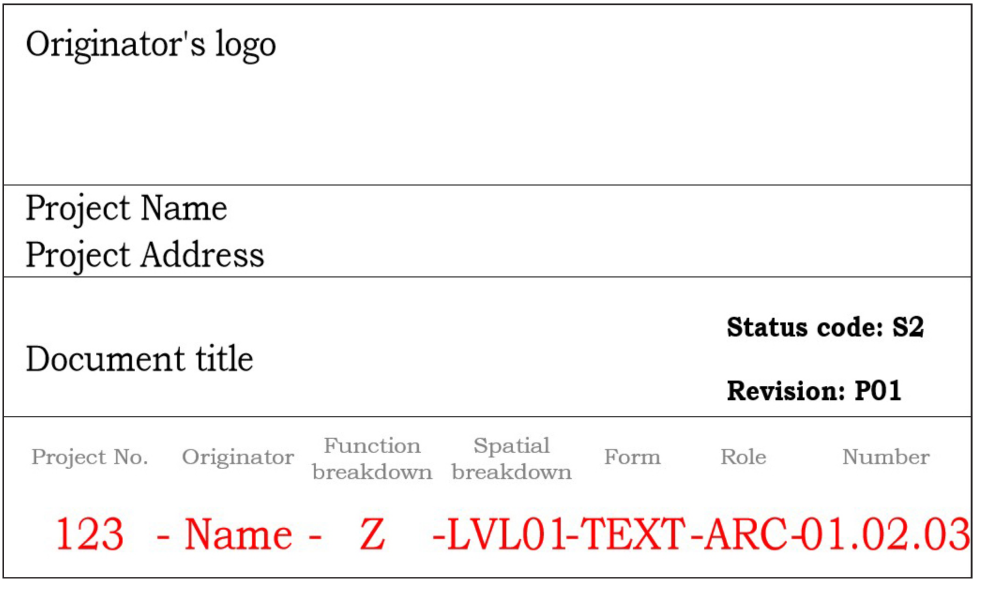

Fig. 4. Example of title block with changed nomenclature

Introducing a globally standardized naming convention for files is a novelty that offers numerous advantages that contribute to the efficiency, consistency, and professionalism of a development project. One significant benefit is the provision of additional metadata through identifier names, which helps clarify the purpose and context of each file. This practice formalizes expectations within a team, thus promoting consistency and reducing misunderstandings. A well-implemented naming convention also facilitates the use of automated tools for refactoring or searching for and replacing terms, thus minimizing the risk of errors. Moreover, it enhances the clarity by preventing ambiguity, thus improving the overall readability of the project. From an aesthetic perspective, a naming convention ensures that filenames are professional, avoiding overly long, comical, or vague names. This is particularly important in collaborative environments where different organizations might combine their work, as it helps to avoid naming collisions by maintaining unique and meaningful identifiers. Additionally, a clear naming convention supports smooth project handovers, providing future developers with meaningful data and better understandings of the code – especially in cases of reuse after long periods.

1.2. Approach to Research

To be able to identify naming protocols for building projects, it is worth assessing the present conditions and understand whether there are already any rules and regularities. It is equally beneficial to define the main features that the naming protocol should fulfill; this should be short and easy to remember. A character's number can make a significant difference when working with hyperlinks for paths and packing/unpacking files. On the other hand, it would be expected to recognize the content of a file quickly and without any doubts based only on its name. The more explanatory the file-naming is, the quicker and easier the participants will learn the logic behind it. All of these features are captured in Figure 5.

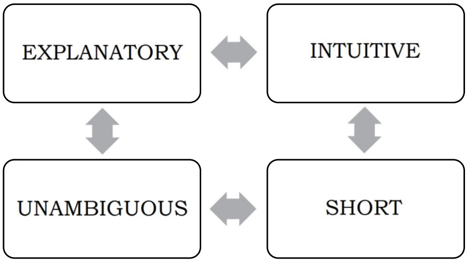

Fig. 5. Characteristics of good file-naming

To address these needs comprehensively, the authors employed several methods of study:

- Analyses and evaluations of national, corporate, and commercial standards – these provided a foundational understanding of existing practices and their effectiveness.
- The conducting of two surveys: the first one was intended to gain insights into the habits and behaviors of those individuals who work with standards, thus allowing the authors to assess the practical challenges and preferences in the real world; and the second one was meant to assess which were best based on the examples.
- An exploration of the psychological aspects of naming conventions – recognizing that standardization is closely tied to user experience (UX) and cognitive processes.

2. Research

2.1. Existing Overview of Standards

The diversity of the standards complicates processes and limits automation possibilities despite any potential benefits from BIM [12]. There are initiatives from the past that attempted to set up some international or national strategies [13], but the statements around file-naming were vague (like "naming convention should be used and documented" [14] or "consistently use agreed naming conventions..." [15]). Several national BIM standards do not even describe naming, stating only that this standard needs to be set up on a project basis [16, 17]. Some do not cover standards for naming but only for objects or layers [18, 19]. For projects that are based on these documents, whole strategizing needs to be done from scratch each time. There are manuals and guides that suggest naming protocols, but only for particular types of files (see Fig. 6); e.g., models or drawings in their native formats [20, 21].

- MODEL NAMING EXAMPLE
  - (Firm Proj #)-(Firm Acronym)- (VCU Bldg #)-(Project Acronym)- (Discipline)-(Subset).file type
    - 34567-BCWH-279-ALC-ARCH.rvt
    - 5678-KSS-279-ALC-ARCH-Interiors.rvt
    - 9876-HCYu-279-ALC-MEP-Electrical.rvt
    - **34567-BCWH-279-ALC-RECORD.rvt**
    - **9080-HOURIGAN-279-ALC-ASBUILT.nwd**

Fig. 6. VCU suggestion on model-naming
Source: [20]

This suggests that these were focused on the efficient use of technology rather than the standardization of all of the processes and holistic views of a project. The authors in [22] proposed different nomenclatures for models and drawings that were contrary to the logic of standardizing things as such; no matter what the format or content was, the files needed to be named in the very same manner for automation to work. For the values of particular fields, there were codes of different lengths that were proposed (see Fig. 7); these did not allow for easy automation, as CDE platforms are very often not ready to handle different numbers of characters in each field.

Some naming conventions that have been proposed in institutional guides have provided rather vague rules that referred to the whole titles of projects and spaces in names between words while not specifying any particular coding (see Fig. 8) [23].

Among the publicly available documents, there are also those with very specific and substantial naming conventions that have been defined for particular tasks or organizations [24, 25]. The rules that were proposed in these documents were narrow and, therefore, not suitable to be implemented worldwide. Principles and potential errors like the different numbers of characters for the same field were the same. Wider naming conventions have been covered within national guides [26–29]; these seem to have been better organized and prepared for various scenarios, yet none of them was popular enough to become officially a national naming convention (not to mention worldwide acceptance).

DISCIPLINE DESIGNATOR CODES

Discipline Name	Designator Code
Architectural	A
General	G
Title/Coversheet	T
Antenna	ANT
Hazardous Materials	H
Landscape	L
Civil	C
Mechanical	M
Electrical	E
Plumbing	P
Borings/Geotechnical	B
Builders Pavement Plan	BBP
Curb Cut	CC
Demolition	DM
Energy Code Compliance	EN
Construction Related Equipment	EQ
Fire Suppression Systems	F
Fire Alarms	FA
Foundations	FO
Fire Protection	FP
Structural	S
Stand Pipe	SD
Signs	SG
Excavation	SOE
Sprinkler	SP
Sprinkler & Standpipe	SP/SD
Site Safety	SSP
Other Disciplines	X
Zoning	Z

As an example, an Inter-Disciplinary Coordination File is being created for a Plumbing project with Project identification Number of LQ471BNA for its revision 7. The file should be named as follows:
LQ471BNA -P-07.nwd

Fig. 7. Discipline coding as per NYCD

Source: [22]

File Name

The file shall be named according to the GSFIC project number followed by the using agency abbreviation and project title. For example:

I-93 UGA Special Collections Library.ifc[14] or

TCSG-245 Griffin Tech Medical Technology Building.ifc

Fig. 8. GSFIC file-naming proposal
Source: [23]

There are better examples as well. One of them is Germany, where the obligation to use BIM on public infrastructural projects was introduced in 2020 by the Federal Ministry of Transport and Digital Infrastructure (BMVI) through the "Stufenplan Digitales Planen und Bauen" plan [30]. There was also the DIN SPEC 91391-1:2019-04 [31] that specified a naming convention that was similar to international norms [11]. In particular, DIN SPEC 91391-2 introduced the concept of "Information Containers," which are used for storing and exchanging data in BIM projects. Each such container includes a set of metadata points, such as the following:

- Identifier (ID): unique identifier for container;
- Name: human-readable name for container;
- Type: classification of container (e.g., "Drawing," "Model," or "Report");
- Description: additional description of container's contents;
- Created: date and time of container's creation;
- Creator: information about container's author;
- Recipients: list of intended recipients;
- Revision: container's revision number;
- Version: version number of container;
- Status: current status of container (e.g., "Open," "Rejected," or "Approved").

Although the standard did not impose specific file-naming conventions, it recommended using consistent and clear names in order to facilitate the identification and management of files within the CDE. Additionally, the standard emphasized the use of metadata for effective information-management in BIM projects. In practice, organizations often develop their own file-naming guidelines to ensure the consistency and ease of documentation-management in BIM projects.

In Spain, there is no single nationwide standard that has defined file-naming conventions in BIM projects; however, guidelines and recommendations for BIM implementation were developed within the Es.BIM initiative (funded by the Spanish Ministry of Public Works); these also covered aspects that were related to information-management (including file-naming conventions).

In Poland, there is book of good practice (BIM Standard PL [32]), where a naming convention was described as the activator for efficient BIM (for details, ISO 19650 is referenced). In the last couple of years, the National Committee for Norms (PKN) focused on translating the ISO 19650 series into the Polish language. Together with [32], this may enable the wider use of BIM in the construction industry.

2.2. National Annex 2

Similar to the approach in Poland, the majority of the other manuals and guides that have been created since 2018 refer to the nomenclature that is mentioned in National Annex 2 (NA.2) [11]; this concerns information regarding container coding, which encounters difficulties in universal applications despite its international significance. The nuances and possibilities for diversification in the standard show the limitations of the current approach to standardization. Such challenges have prompted a thorough analysis of the existing norms/provisions and considerations of proposed changes in order to more effectively implement standardization and automation in the construction industry.

To assess the naming standard that is presented in NA.2, a detailed analysis of its structure is required; this should cover aspects such as the following:
- parameter sequence,
- applied separator,
- lengths of individual fields,
- value suggestions and character type.

Key elements of the filename that assist in quickly orienting toward the context and content of the file include the following:
- project code to which file pertains;
- code of organization that is responsible for file;
- type of content code (to which functional part of project's content relates);
- location code (to which part of object does content pertain);
- information form code in file (which type of information – written, image, movie?);
- discipline or industry code;
- number that forms sequence number or number that, together with other name fields, creates unique identifier.

Container identification information per NA.2 BS EN ISO 19650-2:2018 [4] is as follows:
- Project Code,
- Oganization Code,
- Functional Breakdown,
- Spatial Breakdown,
- Form,
- Discipline,
- Number.

The designed parameter sequence is characterized by a hierarchical arrangement that starts from the most general elements and allows for a gradual transition to more-detailed data. This structure facilitates the efficient decoding of the presented data context and its precise classification.

In the naming-system-design process within NA.2, the selection issue of the naming-field separator highlights the importance of the 'en' (–) as the preferred delimiter. This choice is justified by a deductive analysis of the available separator options in terms of their universality, readability, and acceptance in use. As opposed to other potential separators such as a dot (used to demarcate a file extension), underscore (which can become unreadable if the entire name is underlined), comma (often interpreted as a numerical separator), or other special characters (which can be mistakenly interpreted as mathematical operators), an 'en' dash presents the highest level of readability and minimizes the risk of misunderstanding. Therefore, the 'en' dash is the recommended separator in the context of presentation clarity and the avoidance of interpretation errors, thus supporting its widespread acceptance in a project's data-management space.

For each of the fields that were mentioned above, NA.2 introduces more or less detailed recommendations:

- Project Code – lacks specific recommendations. Besides the project identifier, NA.2 indicates that the code might include additional components that identify any sub-projects that result from a project-breakdown strategy or other project-breakdown scopes. NA.2 does not specify the numbers of characters, but it can be assumed that these should range from two to six in order to apply the above recommendations.
- Organization Code – lacks detailed recommendations. The code of a given organization should be standardized and used in all of the projects in which the organization participates. NA.2 does not specify the numbers of characters; since the code should be easily associated with the organization, however, it can be assumed that this should also range from two to six.
- Functional Breakdown – identifies the functional part of the object or system to which the file data is related. NA.2 does not specify the numbers of characters; however, this is indicated with the following examples: ZZ – the file contains information that is related to various parts of an object or a group of systems; XX – the file does not contain any information that is related to parts of an object or a group of systems. Based on these two examples, it can be assumed that this may contain two characters.
- Spatial Breakdown – specifies the location (so, where the elements are in the object for which the data is specified in the file. This code can be used for standard locations (floor number, grid field number, reference point number, or a location that results from an adopted project division strategy); e.g., the left wing of a building. NA.2 again recommends some values: ZZ – the file contains information that is related to various locations; XX – the file does not contain information that is associated with a location. Therefore, it can be assumed that the location code should also contain two characters.

- Form – specifies the file content (drawing, 3D model, text, Excel sheet, etc.). NA.2 indicates a list of basic values of this code that should be included in the naming standard: D – drawing; G – diagram; I – image; L – various lists and tables; M – model; T – text; and V – video/audio. The field is coded with one character, but NA.2 allows for the use of a greater number of characters if necessary (this type of clause is a typical example that acts against unification).
- Discipline – specifies the industry to which the data that is recorded in a file belongs. NA.2 indicates a recommended list of 20 industries that should be included in the standard; the industries are coded with one character, but two are equally allowed.
- Number – NA.2 states that, if the use of the other name fields does not create a unique file identifier, a number that distinguishes files with the same set of values that other fields have should be used. The number of characters that the number should contain is not specified, but it seems that this number should depend on the size of the project and range from three to six characters. It is recommended to keep the length consistent.

The aforementioned seven fields unequivocally describe the contents of a file and distinctly identify it. However, NA.2 further discusses the proposed standardization of additional metadata:

- revision,
- status,
- information state.

This is because codes that contain the values of this metadata are often added to a filename. These fields include the following:

- revision number/code,
- common code for information state and suitability (status).

These two fields being included in the filename further aid in the efficient management of information but vary over the duration of the project. Therefore, they are not part of the file identifier but carry additional current information about a particular file revision.

Concluding this analysis, it can be stated that a record that uses the minimum number of characters in accordance with NA.2 contains a total of 15 characters:

P-O-FF-SS-T-D-N.

And, a record that uses the maximum number of characters contains a total of 32 characters:

PPPPPP-OOOOOO-FF-SS-TT-DD-NNNNNN.

The proposed rules and definitions allow for the standardization of filenames for both building and infrastructural projects; however, it is practically impossible to indicate a single universal naming standard that is optimal for all types of construction projects. This is best seen in the case of the location code, which would need to describe completely different methods of location that prevail in building projects (e.g., floors) and linear projects (e.g., mileage or sections from station to station).

NA.2 also indicates the possibility of applying solutions that increase the flexibility of the potential standard but may have different effects on the efficiency of using the standard:

- The possibility of building dependencies between different fields; e.g., depending on the value of the location field, the content field's value may carry a different informational value (NA.2, point NA.3.4, note 2). Such a solution is controversial in the context of the common data environment (CDE) system configuration, which should unambiguously verify the correctness of the coding. Of course, some solutions will allow for the "programming" of such dependencies, but these are not common functionalities. Typically, a specific code field is mapped to a specific metadata value that is recognized on the information-exchange platform. Therefore, it seems reasonable to use unique codes for individual values; in the cases of large numbers of potential values, it is worth considering to increase the number of characters in the code field – instead of using a single letter "X," one could propose "XX."
- The possibility of coding two metadata sets in one location field; e.g., B04 indicates the fourth floor of Wing B of a building (NA.2, point NA.3.5, Example 1). Here, however, it is worth using an additional separator; this was already discussed above in the context of the location field for infrastructural objects.

2.3. Survey 1

For the purposes of this article, a survey was conducted among 100 professionals who were employed in the construction industry but were not necessarily familiar with BIM. This was meant to understand whether naming standardization was meaningful and popular to the respondents and, if so, to what extent (and their preferences regarding how the nomenclature should work).

The first question concerned the level of BIM knowledge. On a scale of 1 to 6, the respondents were asked to rate their familiarity with the methodology. The respondents showed various levels of BIM experience and were grouped as below for further analysis:

- 28% had low-level knowledge (Group 1–2: novices),
- 15% had high-level knowledge (Group 5–6: experts).

As their preferred separator, 49% of the participants indicated their preferences for an underscore (_); however, 46% of them were partial to a hyphen (-); the other possible separators received marginal results. Interestingly, most of the respondents with high-level BIM knowledge (Group 5–6) indicated a hyphen (-) – 66.67% of this group. In Group 1–2, 32.14% of the respondents chose a hyphen (-), and 50% chose an underscore (_).

The next two questions concerned the length of fields and were open numerical questions. According to the respondents, the maximum length should be three characters (20%) or four (17%). Less frequently, there were suggestions of five (12%) and six (15%) characters. Values above ten characters (although different up to 80) were indicated by 22% of the respondents. In the expert group, the most common answer was six characters (26.67%); other popular responses were five, four, and three. The novices indicated answers of three and six at an equal level (21.43%).

According to respondents, the minimum field length should be two characters (41%); other popular answers were three (19%) and one (13%). Regardless of the level of BIM knowledge, the answers were very similar in the groups of experts and novices.

Regarding the character-type preferences, a decisive majority of 85% of the participants allowed for the interchangeable use of numbers and letters. Interestingly, slight deviations in the preferences occurred toward letters among the experts and toward numbers among the novices.

In the question about the number of fields in the naming protocol, the participants could choose answers from one to ten. The most common ones were eight (29%), seven (25%), six, and nine (12 and 9%, respectively). There were no deviations in either of the respondent groups.

In the last question, the respondents were asked to arrange the fields in order to form the naming rule in the number that was indicated in the previous question; the choices were Organization, Type/Form, Title, Project Name, Project Number, Discipline, Stage (state of investment), Individual Number (ID), Version, and Location.

Most often, the participants indicated Project Number (52%), Organization (19%), and Project Name (14%) as their first choices; this was not unequivocally dependent on the level of BIM knowledge. In the second position, 31% of the respondents placed Organization, Project Number (19% each), or Project Name (15%).

It was interesting that the field that is not considered in the British National Annex to [11] (i.e., Stage [State] of Investment) most often appeared in the respondents' third position (22%). Other suggestions for this position were Location and Discipline (15% each).

Considering that the respondents most often indicated seven or eight fields in their naming lengths, it was worth checking which fields were most often rejected and not taken into account when indicating the orders (see Table 1).

The highest total percentage shares in these positions were "Version," "Project Name," and "Title." In the first case, NA.2 clearly indicates that the version should not be part of the naming but added metadata. "Project Name" and "Title" are descriptive values that are not always suitable for encoding. The occurrence of the Location field (indicated in one of the last three positions by a total of 42% of the respondents) was surprising; this suggested that this group of users probably omitted this consciously or treated it as less important.

Table 1. Fields that were rejected from nomenclature

Value	First position from end [%]	Second position from end [%]	Third position from end [%]	Total [%]
Individual Number (ID)	9.0	17.0	11.0	37.0
Stage (State) of Investment	2.0	9.0	13.0	24.0
Project Number	4.0	2.0	1.0	7.0
Location	28.0	9.0	5.0	42.0
Organization	5.0	12.0	6.0	23.0
Project Name	19.0	18.0	7.0	44.0
Title	10.0	14.0	20.0	44.0
Type/Form	2.0	3.0	7.0	12.0
Discipline	1.0	0.0	4.0	5.0
Version	20.0	16.0	26.0	62.0

To sum up, the respondents largely understood and indicated the necessity of using both numbers and letters. The respondents with better knowledge of BIM tended to prefer the hyphen (-), while those engineers in the novice group preferred the underscore (_). This preference may have been due to the lack of familiarity with the standard and the awareness that not all operating systems and applications handle hyphens uniformly in filenames. Hyphens can be interpreted differently by various systems or scripts, which can lead to file-processing issues or errors. In programming and scripting (where filenames are often used as parameters), hyphens can be mistaken for operators or special characters, thus complicating the manipulation of filenames in scripts and code. It is important to consider readability – while hyphens can improve readability (especially in long filenames), some systems may handle other characters (like underscores) better, as they are not visually separated. In programming and scripting, underscores are often preferred because they are more "syntactically safe."

The number of fields varying between six and nine aligns with the recommendations of the NA; this suggests that the industry's informational needs are largely met. In other words, respondents mainly need seven to eight fields to convey information about the content of an information container.

Regarding field length, the knowledge of the NA among the experts was confirmed, as they indicated six characters as the maximum. Generally, the most common suggestion was three characters, which aligned with the ease of operation, memorization, and distinction of information. According to the majority, the minimum field length was two characters; when using both numbers and letters, this allows for 1296 unique values (even in large projects).

In arranging the order of the fields, the respondents correctly followed the principle from general and the most important information to specific and detailed information. There was no clear dependence of the data on the level of BIM knowledge, but specific information appeared in the order that was suggested by the NA (or close to it). The version and descriptive fields were also correctly rejected.

Survey 1 conclusions:

- The diversity of the results showed how much the industry lacks standardization in such a fundamental aspect as file-naming.
- The separator that is indicated in the NA contradicts the IT and technological approach and may cause problems; however, it is preferred.
- The field length in naming should fit two to six characters.
- The optimal name length is eight fields.
- Information in the name should be conveyed from the most general (concerning entire investment task) to the most specific (concerning particular documents).

2.4. Classifications

A naming convention can be perceived as a multiple classification. Each field in a filename represents some information that is coded with a number of characters; thus, it is crucial to explore existing classifications such as the most common ones: Uniclass [33], or the openBIM construction classification international collaboration (CCI) [34] (which has become more popular even though it is still pure and under development).

Uniclass is organized into a series of interconnected tables – each representing a different aspect of the built environment. These tables help classify information across various stages of a construction lifecycle, from design to management and maintenance. The tables are structured hierarchically, thus allowing for detailed breakdowns and the linking of related information. Each table contains a hierarchical code that starts with a prefix (e.g., Co, En, or Pr), followed by a numeric identifier; this allows for easy referencing and a standardized way to manage information across projects and sectors. The parts of the code are separated with underscores; therefore, using pure Uniclass coding within a naming convention might become overcomplicated. Where the fields are separated by one separator, one field should never be separated with another.

There is an aspect of Uniclass that is worth mentioning and implementing. One of the tables that describes "Form of Information" and is represented by the FI prefix also contains the optional coding of two characters as abbreviations of particular types of information; e.g., DG for drawing, M3 for Model 3d, SK for sketch, DS for data set, etc. Such a table can be directly incorporated into a file-naming convention; not only does it prove that two characters are potentially the best option for this field, but it also gives a set of values that should fully serve information exchange in the building industry.

A similar approach was introduced by the National Construction Information Classification for Lithuania [35], where coding is shown for many different aspects such as spaces, phases, and constructions. Based on the proposed values, it seems like the standard is dedicated more for design and build rather than for operational phases. Among these tables, there is one that is dedicated to documents where the coding is constructed with three characters; however, the first is always the same (D – document), so it can be omitted so as to avoid repeating obvious information. Another table covers participants that are coded with one or two characters.

2.5. Cognitive Psychology

Knowing the potential of NA.2 and the preferences that were expressed in the survey, it is worth remembering that there is another aspect that is related to the comfort of nomenclature usage. For the sake of efficiency, the standardization should be user-friendly; this means that it should be easy to remember, readable, and suggestive.

The main purpose of creating naming strategies is to encode information about a file's content; in other words, whomever sees a filename should easily and quickly understand what is inside. The process of encoding is inevitably the creations of interpretations of information; thus, we must rely on our memories of events in the future (encoded in our minds), which could be faulty [36]. When working on several projects, one might be affected by past experiences and standardizations; therefore, they can misinterpret a file's content based on its name when not familiar with the current nomenclature. This is also the reason for creating a common naming strategy for the whole building industry. Encoding usually filters out much of the richness of a sensory experience, focusing instead on capturing primarily the semantic content or meaning [37]. In data-management, this is exactly what is required – simplicity, a minimum amount of information, and efficiency. These can be achieved by searching for repetitive schemas.

When things are organized with schemas and applying heuristic encoding techniques, most typical situations do not require much strenuous processing. Individuals can absorb, arrange, and understand information quicker and with less effort [38] when they follow schemas that are known from the past. There are also two types of memory-retrieval, which may have an impact on information-encoding times and, therefore, comfort and efficiency when using naming conventions [39]. Recognition is the ability to identify whether a thing is already familiar or if anything is completely new. Recall requires more in-depth analysis to get details from one's memory; this is why recognition is far more preferable in user interfaces and coding mechanisms. What helps in easy recognition is chunking – a process where small individual pieces of information are grouped together to form a coherent and meaningful whole [40]. Those interconnected units can be measured by activation – how easily the chunk can be retrieved from one's memory. This can be affected by three factors – practice, recency, and context: the first one is about how often the chunk is seen and "worked with"; the second is about what was the last time the information was processed; and the third is related to situation(s) when the chunk was used (and with what it was associated).

Having been undertaken from cognitive psychology, the above theory suggests that the ease with which information can be retrieved from one's memory depends on the frequency of one's encountering of the information, the recency of its use, and its relevance to the current context. In global naming, standard types of information and example values should be strictly defined as well as the numbers and orders of fields and their lengths and characters types.

For the number of fields definition, there is a theory that was represented by Miller [41] that the memory span of young adults is approximately seven items (±2). Equally, the number depends on the complexity of these pieces (chunks).

A field's sequence is basically the structuring of an information architecture (IA) – the systematic process of organizing and presenting the various elements of an entity (whether physical or digital) to users in the most straightforward manner. IA can be applied to a wide range of contexts, including naming standardization. To construct effective IA, it is essential to understand some of its key components.

Organization Systems: these classify information into categories, making it easier to navigate. They can be further broken down into the following:

- hierarchical: organizes information by importance;
- sequential: organizes information logically or in steps from start to finish;
- metrical: organizes information based on individual user preferences.

Labeling Systems: these involve the consistent use of terminology to describe and represent different pieces of content, ensuring clarity and ease of understanding for users.

The other two components, which are navigation systems and searching systems, are not applicable in the context of naming standardization.

One of the very important aspects is a separator – a character that divides fields from each other. The vast majority of those that are represented on computer keyboards are dedicated to mathematical operations or computer system syntax. Those that are most comfortable in perception (vivid and easy to distinguish) are underscores (_) and hyphens (-).

In the literature that is dedicated to cognitive psychology in architecture [42], there is evidence for human short-term memory that is characterized by a limited capacity for processing information, effectively encompassing up to six characters. In this process, information is subconsciously organized into groups, with four-character sequences being perceived as continuous and five- and six-character configurations being divided into smaller subsets (into two groups [2 + 3, or 1 + 4] or two three-element subsets). Such a grouping mechanism is applied, among others, when providing phone numbers.

In terms of character types, it is worth remembering that fields that contain letters can form a word that can be easily associated on the one hand but misleading on the other. The values for particular fields will not always form words, so a naming may look inconsistent. Using only digits allows for easier pronunciation as a whole number – but only up to some point. When creating fields that are longer than two digits, users will most likely memorize and read them by digits; e.g., 25 – "twenty-five"; 257 – two five seven (rather than "two hundred fifty-seven"). To allow for the maximum number of options within one field, having both digits and letters seems to be the best choice.

Analyzing the length of each field, it is necessary to consider the potential amount of information that it can carry and, thus, the number of scenarios that can be realized through the diversity and type of the characters. Numeric formats (10 digits) and alphabetical formats (26 letters of the alphabet) are available. Depending on the choice of one or both of these formats and the number of characters that are contained in a field, it is possible to generate a specific number of variants with repetitions (which is indicated in Table 2). According to the NA.2 recommendations, the maximum field length is six characters.

Table 2. Number of combinations depending on number and type of characters

Number of characters	Numbers only	Letters only	Numbers and letters
One character	10	26	36
Two characters	100	676	1,296
Three characters	1,000	17,576	46,656
Four characters	10,000	456,976	1,679,616
Five characters	100,000	11,881,376	60,466,176
Six characters	1,000,000	308,915,776	2,176,782,336

2.6. Survey 2

After concluding the previous stages of the research, how a naming convention may appear became clearer. The vast majority of the rules can be implemented directly from NA.2, but the numbers of characters for each field are still questionable.

Therefore, another survey was undertaken to understand how fast and comfortable the nomenclature should be. The form was completed by 60 participants – students of post-Masters classes in BIM.

The survey aimed to assess the efficiency of file-naming protocols in accordance with NA.2 considering the diversity in the numbers and types of characters in the protocol; it investigated how different naming scenarios affected the identifiability and clarity of a file. The introduction outlined the structure of the naming convention, including the numbers and types of fields and separators. The participants were asked to decode two sample filenames over three stages – once for a building project, and once for an infrastructure project. Along with the filenames, the participants received a brief situational context to help decode the fields. The questions focused on determining which project a file belonged to, its author, its represented discipline, the specific part, and the type of document. For each question, "it is hard to say definitively" was an available response. The participants were encouraged not to hesitate to choose this response, as preliminary tests indicated a psychological tendency to view this choice as a failure to decode a name (whereas it is actually a valid response).

The survey stages differed in the levels of coding and amounts of information that were provided:

- Part One: minimal number of characters. The participants evaluated filenames with the minimal number of characters that were necessary to identify the content of each file. Each field contained the smallest possible number of characters.
- Part Two: maximum number of characters. In this part, the filenames contained more characters, which made it theoretically sufficient for the easy decoding of the codes.
- Part Three: standard BIM documentation. The participants had access to filenames with short fields along with the naming protocol that was used in the BIM methodology.

The analysis of the results from the first survey indicated that the majority of the respondents found it difficult to unambiguously determine what each field represented; the average for this response was 54.94%. It should be noted that, where the instruction was intentionally ambiguous and the names were similar, the response "it is hard to say definitively" was chosen by up to 85% of the respondents. Conversely, the respondents confidently identified the correct ones where the values were distinctly different. However, it should be noted that, in such cases, there were no similar values or they were not specified. For example, only the architectural discipline was listed without automation or archaeology. Given the code "A," it was naturally assumed to represent architecture; however, there are typically many disciplines that could be coded the same way in large projects. With single-character coding, the first available letter in the alphabet is often sought, which may not intuitively suggest the value of the field.

In the second survey (where the code fields were represented by longer values), the distribution of the responses was significantly different. The response "it is hard to say definitively" was indicated in an average of 46.33% of the cases, but correct answers were much more frequent; the average for the incorrect answers (or errors) fell to below 10%. Thus, the respondents either expressed their uncertainty cautiously or provided the correct answer. The time that was taken to complete this survey was also the shortest (2.4 minutes, as compared to 3.2 and 3.74 minutes. This suggested that longer code field values left less room for doubt, thus enabling the users to make decisions more quickly. Despite any potential similarities or closeness in names, longer values allowed for greater diversity; for example, the architecture, automation, and archaeology disciplines could be represented by the codes AR, AU, and AL, respectively; this makes them easily identifiable.

For the third survey, a supplementary naming protocol was prepared; this was similar to what is usually included as part (or an annex of) a BIM execution plan. This protocol contained a decoding legend for all of the fields. As a result, the number of responses that indicated "it is hard to say definitively" dropped to 12.04% in this survey; also, incorrect responses were often nonexistent.

3. Results

In the context of the above analysis of NA.2 to [11] and other naming standards (while also taking the preferences of naming use among the respondents of the survey and the cognitive psychology aspects under consideration), the authors of this thesis came up with a universal naming convention that is dedicated to buildings; this covers the following:

- separator type,
- number of fields,
- order of fields,
- number of characters for each field,
- type of characters in each field.

Field order is as follows:

- Project Number,
- Originator,
- Stage,
- Volume,
- Part,
- Form,
- Discipline,
- Number.

The convention suggests the inclusion of eight distinct fields – one more than is proposed in NA.2, with the additional field accounting for the Stage (as was indicated by the majority of the survey respondents). This structure was also the most frequently selected option in the survey. The sequence of fields followed a logical progression from the most general information (such as project identification and party involvement) to specific document identifiers. It was aligned with the proposal that was mentioned in National Annex (plus a field that indicated "Stage"). The number of fields went along with the rule of the magical number "seven" (mentioned earlier).

It is questionable, however, if the order of the fields could have been represented in a more logical way. Currently, it was a mixture of information about the projects, the involved parties, the parts of the project, and the disciplines that these parties represented. By grouping the fields that characterized similar information, more clarity could have been achieved. This was why it was proposed to reorder the fields, so parts of the name answered questions about where, who, and what (see Table 3).

Table 3. Final naming proposal

WHERE			WHO		WHAT		
Project Number	Volume	Part	Discipline	Originator	Form	Stage	Number
PPPPP	VVV	RRR	DD	OOO	FF	SS	NN
5	3	3	2	3	2	2	6
D(+L)	D+L	D+L	L	L	L	D+L	D(+L)

The first three fields described the information container affiliation – which project, which volume, and which part. The next two fields described who the document belonged to – the discipline and originator. The final three fields were about the document itself – its form, stage, and unique number.

The proposed naming convention featured a total of 26 characters, including separators. This sat below the maximum allowable limit as is delineated in the National Annex and outlined at the beginning of this research; this was despite the introduction of another field ("Stage"). To enhance the readability and maintain consistency, it was recommended to use a hyphen or minus sign (−) as the separator between the fields. This choice aligned with the guidelines that were established in the National Annex and significantly improved the visual distinction between the different metadata fields. Concerns regarding the hyphen being misinterpreted as a mathematical operator could be effectively addressed by enclosing the filename in quotation marks when used within programming scripts or formulae, thus ensuring that the name was treated as a string.

Each field's character limit was carefully calibrated to balance the need for brevity with the requirement to encapsulate sufficient information for a wide range of scenarios. The following breakdown elucidates the rationale for each field's character allocation:

– Project Number (five characters). This field can accommodate up to 10,000 projects using numerical identifiers alone. By incorporating both letters and digits, these five characters can encode additional layers of information; for example, the first two characters could represent the year the project was initiated, thus offering immediate insight into the project's timeline. Alternatively, these characters might denote a specific department within a public organization, thus allowing for the quick identification of the managing department. The five-character limit, thus, provides flexibility and clarity, making it suitable for the encoding of various scenarios.
– Volume (three characters). Retaining the terminology from BS1192:2007, this field denotes whether a file relates to a specific volume or if it is generic. Using both letters and digits increases the flexibility and recognizability of this field.
– Part (three characters). Similar to "Volume," the part field allows for detailed categorization – particularly in those projects where specific objects are associated with locations or levels. This three-character approach facilitates precise positioning within the building's hierarchy.
– Discipline (two characters). While a single-character field may suffice for smaller projects, larger ones often involve multiple disciplines that may share initials (e.g., Surveyor and Structural Engineer). Allowing for two characters helps to avoid ambiguity and aligns with the options provided by NA.2. Again, there are 676 options for naming the discipline of the stakeholder.
– Originator (three characters). This field represents the company or entity responsible for the file. In the architecture and construction industries, abbreviations of firm names (typically derived from the names of their founders or long-form company names) are common. An analysis of the top 100 architectural firms globally [43] showed that many used three- or four-character codes (e.g., Zaha Hadid Architects – ZHA, and Foster & Partners – FPA). By limiting the field to letters only, the convention allows 17,576 unique identifiers, minimizing the likelihood of duplication within a single project or organization.
– Form/Type of Document (two characters). Initially simplified to a single character in the February 2021 corrigendum to NA.2, this research recommends a two-character field to distinguish between closely related document types (e.g., 'M3' for 3D models, 'MB' for marketing brochures). This differentiation enhances the clarity and supports a broader range of document types. Two characters for this filed are also indicated in Uniclass; therefore, there is already a classification that is available on the market to represent the information about the form or type of document. Allowing for two letters only gives 676 different options; this should be sufficient for all phases of a construction project.

- Stage (two characters). The survey respondents highlighted the importance of including a stage (phase) field early in the naming convention to indicate the project stage. Different countries have varying numbers of defined project stages, ranging from five in the United States (AIA) to nine in India (COA) and Germany (HOAI). A two-character field permits the representation of the primary stage and any necessary variations (e.g., multiple submissions to planning authorities), thus enhancing the precision of the file categorization.
- Individual Number (six characters). This field accommodates unique identifiers for each document; these are commonly used in large projects where drawings are organized into packages. The use of two digits to denote packages (reflecting categories like substructures or windows) and three digits for document numbers within each package aligns with international practices. An additional digit can denote the document's information status (e.g., existing, demolished, proposed, or temporary).

The example of information container naming which follows proposed convention is provided below in Table 4.

Table 4. Example naming

Project Number	Volume	Part	Discipline	Originator	Form	Stage	Number
12345	B01	LGF	AR	XYZ	VI	R2	022099

The explanation of meaning for each field is as follows:
- Project Number (12345) – the individual project number that is assigned by the party (this is to be considered if a project does not have its number assigned by the appointing parties for all of the stakeholders).
- Volume (B01) – building 01, which is easily recognizable and allows for a maximum of 100 buildings for one project.
- Part (LGF) – again, three characters allow for more-recognizable naming; in this case, the name refers to the lower ground floor.
- Discipline (AR) – architecture; the number of the available combinations of two characters should be sufficient to find individual representations for all disciplines (AV – Audio Video, AC – Archeologist, etc.).
- Originator (XYZ) – the company name that provides the information.
- Form (VI) – visualization; the logic is similar to Discipline.
- Stage (R2) – RIBA (Stage 2 as an example).
- Number (022099) – this can be a combination of the package number (02); then, "2" may refer to the proposed phase, and 099 is an individual number of the drawing (referring to the floor). In this case, 100 would be the ground floor, and the numbering goes up and down.

4. Conclusions

The proposed naming convention has been designed to balance complexity and usability, thus offering a structured approach that supports efficient data-management across various project scales. It can be used for whole construction industry and building projects (however, not for infrastructural [linear] investments). The naming is not only applicable to the drawings (as is very often the case nowadays) but also to all kind of files (including marketing materials [visuals, animations], legal documents, schedules, reports, models, and any other digitalized forms of information).

Unfortunately, infrastructural projects differ in terms of the representation of the localization; therefore, the Volume and Part fields should be described with more characters. The rest of the convention (the number and order of the fields) can be exactly the same.

By adhering to these guidelines, organizations can achieve greater consistency, clarity, and ease of use in managing project data. The inclusion of specific fields and character limits reflects a comprehensive understanding of industry needs, promoting a standard that can be widely adopted for the benefit of the building industry. Future work should focus on refining these conventions further and developing supportive tools and tables that assist practitioners in implementing these standards across diverse project types. In the next steps, it will be required to create tables of example values for each field, as only what these are and how they are formed have been specified. Based on the example of a discipline, a list of all possible disciplines might be created within a first step. A similar approach regarding the document type has already been undertaken in Uniclass [34]. Together with the naming convention, these tables can form a future international version of NA.2 of [11].

5. Discussion

Regardless of the specific outcomes of this research, it is imperative to underscore that naming conventions play critical roles in software applications that are tailored for the building industry and CDE platforms. In these contexts, each component of a filename is treated as distinct metadata, thus encapsulating crucial information about the file and its contents. This structured approach to naming enables sophisticated functionalities for filtering, sorting, and grouping files based on specific parameters and their corresponding values. The proposed convention is applicable to all kinds of files that support information flow in the construction industry.

In contrast, widely used operating systems such as Windows, macOS, and Linux lack native support for such advanced naming conventions (thus placing the burden of data organization squarely on the end users). Typically, users employ a hierarchical folder structure that is organized by a project's phases, disciplines, stakeholders, and delivery dates. This conventional method does not support querying files by specific attributes across different phases or disciplines (such as retrieving all files that are related to a particular party). Consequently, this limitation results in suboptimal information-management and increased inefficiency.

Moreover, the default file-sorting mechanisms within these operating systems generally adhere to alphabetical ordering, which can exacerbate confusion and complicate data-management. For instance, naming conventions may be erroneously positioned at the end of a sorted list if alphanumeric codes are utilized to indicate underground levels (often labeled 'UG,' 'B1,' 'B2,' etc., where 'B' denotes 'Basement'). This raises the question of whether operating systems should evolve to incorporate more-complex naming conventions that facilitate the identification and extraction of specific metadata fields.

The naming convention that was proposed as a key outcome of this research builds on the guidelines that were suggested in [4]. As this document gains global traction, adherence to its standards may facilitate widespread implementation and standardization. The proposed naming convention offers a clear and unequivocal framework that is designed to accommodate a broad range of scenarios – potentially encompassing the most extensive and intricate projects. Future research could focus on the development of comprehensive tables that detail possible values for each naming field alongside rules and guidelines that are tailored to different project types. When combined with the nomenclature that was developed from this research, these resources could form the foundation of an international naming standard for the building industry.

An important consideration is whether the proposed naming convention is overly complex for smaller projects (such as single-family homes, which typically involve limited numbers of stakeholders and a single volume). Questions may arise as to whether the naming convention's length and flexibility might introduce inefficiencies by allowing too many potential scenarios. This is why the next step of creating the table of values for each field is needed. Alternatively, the standardized values could remain consistent throughout a project's lifecycle, reducing the need for frequent changes and simplifying the data-management. Further studies could explore the balance between the convention's complexity and its utility across various project scales, thus ensuring that it enhances efficiency without imposing unnecessary burdens on smaller projects.

Funding

The authors (1–5) would like to acknowledge the support from Europe's Rail Flagship Project 4 – Sustainable and green rail systems, which received cofunding from the Horizon Europe Framework Programme (HORIZON) under Grant Agreement Number 101101917 and the Polish Ministry of Science and Higher Education under Grant Agreement Number MEiN 5337/HE/2023/2 within the International Co-Financed Projects program. The project is being implemented at AGH University of Krakow during the period of 2022–2026. The FP4 Rail4EARTH project under Grant Agreement No. 101101917 is supported by Europe's Rail Joint Undertaking and its members. The views and opinions that were expressed in this paper are, however, those of the authors only and do not necessarily reflect those of the European Union nor Europe's Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.

CRediT Author Contribution

J. P.: conceptualization, methodology, visualization, formal analysis, investigation, resources, data curation, writing – original draft preparation.

T. O.: conceptualization, methodology, writing – review and editing, project administration, funding acquisition.

D. K.: writing – original draft preparation, methodology.

M. W.: writing – original draft preparation, methodology.

M. R.: writing – original draft preparation, methodology.

K. C.: writing – original draft preparation, methodology.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work that was reported in this paper.

Data Availability

The data that supports the findings of this study is available on request from the corresponding author [J.P., jpasalski@agh.edu.pl]. The data is not publicly available due to its sensitivity.

Use of Generative AI and AI-Assisted Technologies

During the preparation of this work, the authors used ChatGPT-4o in order to translate some parts of the text from Polish to English. After using this tool/service, the authors reviewed and edited the content as needed, and they take full responsibility for the content of the publication.

References

[1] Mischke J., Stokvis K., Vermeltfoort K., Biemans B.: *Delivering on construction productivity is no longer optional*. McKinsey & Company, August 9, 2024. https://www.mckinsey.com/capabilities/operations/our-insights/delivering-on-construction-productivity-is-no-longer-optional#/ [access: 27.10.2024].

[2] Brown J.: *Tech-Clarity Perspective: Best Practices for Managing Design Data: How Effective Data Management Fundamentals Enable World-Class Product Development*. Tech-Clarity, Cincinnati, OH, USA, 2012.

[3] Huang Y., Butler A., Gardner N., Haeusler M.H.: *Lost in translation: Achieving semantic consistency of name-identity in BIM*, [in:] Stojaković V., Tepavčević B. (eds.), *ECAADe 2021: Towards a New, Configurable Architecture: Proceedings: The 39th Conference on Education and Research in Computer Aided Architectural Design in Europe: Hybrid Conference, 8th–10th September 2021, Novi Sad, Serbia, Faculty of Technical Sciences, University of Novi Sad. Volume 2*, University of Novi Sad, pp. 9–20. https://doi.org/10.52842/conf.ecaade.2021.2.009.

[4] Chen K., Lu W., Wang H., Niu Y., Huang G.G.: *Naming objects in BIM: A convention and a semiautomatic approach*. Journal of Construction Engineering and Management, vol. 143(7), 2017, pp. 06017001–06017001-6. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001314.

[5] Abualdenien J., Borrmann A.: *Levels of detail, development, definition, and information need: A critical literature review*. Journal of Information Technology in Construction, vol. 27, 2022, pp. 363–392. https://doi.org/10.36680/j.itcon.2022.018.

[6] Royano V., Gibert V., Serrat C., Rapinski J.: *Analysis of classification systems for the built environment: Historical perspective, comprehensive review and discussion*. Journal of Building Engineering, vol. 67, 2023, 105911. https://doi.org/10.1016/j.jobe.2023.105911.

[7] Afsari K., Eastman C.M.: *A comparison of construction classification systems used for classifying building product models*, [in:] *52nd ASC Annual International Conference Proceedings*, Associated Schools of Construction, Provo, UT, USA, 2016. https://doi.org/10.13140/RG.2.2.20388.27529.

[8] Zhai Y., Zhang C., Wang B., Wang X., Liu K., Tang Z.: *The standardization method and application of the BIM model for interchanges*. Applied Sciences, vol. 12(17), 2022, 8787. https://doi.org/10.3390/app12178787.

[9] Salvado F., Silva M., Couto P., Azevedo Á.: *Standardization of BIM objects: Development of a proposal for Portugal*. Open Journal of Civil Engineering, vol. 6(3), 2016, pp. 469–474. https://doi.org/10.4236/ojce.2016.63039.

[10] British Standards Institution (BSI): *Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) – Information management using building information modelling – Part 1: Concepts and principles* (BS EN ISO 19650-1:2018). December 2018. https://www.iso.org/standard/68078.html [access: 27.10.2024].

[11] British Standards Institution (BSI): *Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) – Information management using building information modelling – Part 2: Delivery phase of the assets* (BS EN ISO 19650-2:2018). December 2018. https://www.iso.org/standard/68080.html [access: 27.10.2024].

[12] Ajayi S.O., Oyebiyi F., Alaka H.A.: *Facilitating compliance with BIM ISO 19650 naming convention through automation*. Journal of Engineering, Design and Technology, vol. 21(1), 2023, pp. 108–129. https://doi.org/10.1108/JEDT-03-2021-0138.

[13] Edirisinghe R., London K.: *Comparative analysis of international and national level BIM standardization efforts and BIM adoption*, [in:] *Proceedings of the 32nd International Conference of CIB W78, Eindhoven, The Netherlands, 27–29 October, 8–17, 2015*, CIB – International Council for Research and Innovation in Building and Construction, Ottawa, pp. 149–158. http://itc.scix.net/paper/w78-2015-paper-015.

[14] The Building Information Foundation RTS: *COBIM: Common BIM Requirements 2012*. https://wiki.buildingsmart.fi/en/04_Guidelines_and_Standards/COBIM_Requirements [access: 27.10.2024].

[15] Cooperative Research Centre for Construction Innovation: *National Guidelines for Digital Modelling*. Icon.Net Pty Ltd., Brisbane, QLD, Australia, 2009. https://eprints.qut.edu.au/30341/1/BIM_Guidelines_Book_191109_lores.pdf [access: 27.10.2024].

[16] Hong Kong Institute of Building Information Modelling: *BIM Project Specification (Rev 3.0)*. June 2011. https://www.scribd.com/document/377096701/HKIBIM-Specification-Rev3-0 [access: 27.10.2024].

[17] Resource Centre for Technologies and Innovation in Construction (CRTI-B): *Luxembourg BIM Application Guide. Version 1.0 EN*. June 2018. https://www.digitalbuilding.lu/wp-content/uploads/2021/10/Guide-BIM-Luxembourg__V1-0-EN-1.pdf [access: 27.10.2024].

[18] BIM Acceleration Committee: *New Zealand BIM Handbook: A Guide to Enabling BIM on Building Projects*. July 2014. https://www.biminnz.co.nz/resources/nz-bim-handbook [access: 27.10.2024].

[19] Norwegian Home Builders' Association: *BIM User Manual. Version 2.0*. November 2012. https://chcymru.org.uk/cms-assets/legacy/general/BIM_user_manual_2.0_WEB.pdf [access: 27.10.2024].

[20] Virginia Commonwealth University: *Building Information Modeling (BIM) Guidelines and Standards for Architects, Engineers, and Contractors*. September 2013. https://fmd.vcu.edu/media/fmd/documents/construction-planning-and-design/VCUBIMGuidelines-02072014.pdf [access: 27.10.2024].

[21] Kreider R.G., Messner J.I.: *The Uses of BIM: Classifying and Selecting BIM Uses. Version 0.9*. The Pennsylvania State University, University Park, PA, USA, 2013.

[22] Bloomberg M.R., Burney D.J., Resnick D.: *BIM Guidelines*. New York Department of Design and Construction, New York 2012.

[23] Georgia State Financing and Investment Commission: *GSFIC BIM Guide Series 01: Model Analysis and Validation: Design*. March 2013. https://gsfic.georgia.gov/document/publication/building-information-modeling-bim-guide/ [access: 27.10.2024].

[24] City and County of Denver, Denver International Airport: *Design Standards Manual 12, Engineering Data Systems (EDS) Submittal, CADD, GIS, BIM, Metadata & Survey. Volume 12 of 12*. 2012. City and County of Denver, Denver 2012. https://www.yumpu.com/en/document/view/50879199/12-dia-business-center-home-page-denver-international-airport [access: 27.10.2024].

[25] Massachusetts Institute of Technology, Department of Facilities: *MIT Design Standards: BIM and CAD Drawing Standards: T03 Thematic Folder*. 2022. https://web.mit.edu/facilities/maps/DesignStandards/T03 - BIM and CAD Drawing Standards 2022.pdf [access: 27.10.2024].

[26] Ente Italiano di Normazione (UNI*): Construction and civil engineering works – Digital management of construction information processes – Part 1: Models, documents and informative objects for products and processes* (UNI 11337-1:2017). January 2017. https://store.uni.com/en/uni-11337-1-2017 [access: 27.10.2024].

[27] British Standards Institution (BSI): *Specification for information management for the capital/delivery phase of construction projects using building information modelling* (PAS 1192-2:2013). February 2013. https://www.hfms.org.hu/web/images/stories/PAS/PAS1192-2-BIM.pdf [access: 27.10.2024].

[28] Building and Construction Authority: *BIM Essential Guide for BIM Adoption in an Organization*. August 2013. https://www.corenet.gov.sg/media/586143/Essential-Guide-Adoption.pdf [access: 27.10.2024].

[29] Statsbygg: *Statsbygg Building Information Modelling Manual – version 1.2*. October 2011. https://dok.statsbygg.no/wp-content/uploads/2020/06/statsbyggs-bim-manual-1-2-1_en_20131217.pdf [access: 27.10.2024].

[30] Bundesministerium für Digitales und Verkehr: *Stufenplan Digitales Planen und Bauen Einführung moderner, IT-gestützter Prozesse und Technologien bei Planung, Bau und Betrieb von Bauwerken*. December 2015. https://bmdv.bund.de/SharedDocs/DE/Publikationen/DG/stufenplan-digitales-bauen.pdf?__blob=publicationFile [access: 27.10.2024].

[31] DIN Media: *Common Data Environments (CDE) for BIM projects – Function sets and open data exchange between platforms of different vendors – Part 1: Components and function sets of a CDE; with digital attachment* (DIN SPEC 91391-1:2019-04). April 2019. https://www.dinmedia.de/en/technical-rule/din-spec-91391-1/302483139 [access: 27.10.2024].

[32] Piwkowski W., Styliński J. (oprac.): *BIM Standard PL: Projekt zasad przygotowania i realizacji inwestycji kubaturowych w Polsce zgodny z normą PN-EN ISO 19650 i krajowym prawem budowlanym. Wersja dokumentu nr 2.0*. Polski Związek Pracodawców Budownictwa, Warszawa 2020.

[33] Uniclass [delivered by National Building Specification (NBS)]. https://uniclass.thenbs.com/ [access: 27.10.2024].

[34] Construction Classification International Collaboration (CCIC). https://cci-collaboration.org/ [access: 27.10.2024].

[35] Nacionalinis Statybos Informacijos Klasifikatorius (NSIK). https://www.nsiklab.lt/nsik_ontologijos.html [access: 27.10.2024].

[36] Cornell D., Drew C.: *11 Encoding Examples* (*Psychology*). Helpful Professor, January 3, 2024. https://helpfulprofessor.com/encoding-examples/ [access: 27.10.2024].

[37] Fincher S.A., Robins A.V. (eds.): *The Cambridge Handbook of Computing Education Research*. Cambridge University Press, Cambridge 2019. https://doi.org/10.1017/9781108654555.

[38] Kleider H.M., Pezdek K., Goldinger S.D., Kirk A.: *Schema-driven source misattribution errors: Remembering the expected from a witnessed event*. Applied Cognitive Psychology, vol. 22(1), 2008, pp. 1–20. https://doi.org/10.1002/acp.1361.

[39] Gillund G., Shiffrin R.M.: *A retrieval model for both recognition and recall*. Psychological Review, vol. 91(1), 1984, pp. 1–67. https://doi.org/10.1037/0033-295X.91.1.1.

[40] Thalmann M., Souza A.S., Oberauer K.: *How does chunking help working memory?*. Journal of Experimental Psychology: Learning Memory and Cognition, vol. 45(1), 2019, pp. 37–55. https://doi.org/10.1037/xlm0000578.

[41] Miller G.A.: *The magical number seven, plus or minus two: Some limits on our capacity for processing information*. Psychological Review, vol. 63(2), 1956, pp. 81–97. https://doi.org/10.1037/h0043158.

[42] Żórawski J.: *O budowie formy architektonicznej*. Arkady, Warszawa 1962.

[43] Quirk V.: *The 100 Largest Architecture Firms in the World*. ArchDaily, February 11, 2013. https://www.archdaily.com/330759/the-100-largest-architecture-firms-in-the-world [access: 27.10.2024].