Technical report

Date : July 19, 2019
Contract : Contract SRSS/C2019/024
Project : Introducing a Building Information Model (BIM)-based process for building permits in Estonia
Created by : Future Insight Group

Funded by the Structural Reform Support Programme of the European Union and implemented by Future Insight Group B.V. in cooperation with the European Commission's Structural Reform Support Service (SRSS).

This Technical Report has been produced under a contract with the Union and the opinions expressed are those of Future Insight Group B.V. and do not represent the official position of the European Commission.

Index

1 Part A: State of BIM in Estonia **3**
1.1 Introduction 3
1.2 Preparatory developments 3
1.3 BIM Quickscan analyses 4
1.3.1 Introduction BIM Quickscan 4
1.3.2 BIM Quickscans in Estonia 5
1.3.3 Chapters 5
1.3.4 Aspects 8
1.4 BIM analytics 9
1.4.1 Geometry 13
1.4.2 Classifications 15
1.4.3 Materials 17
1.4.4 Properties 18
1.5 Interviews summary 22

2 Part B: State of the art outside Estonia **25**
2.1 Rule checking versus compliance checking 25
2.2 History of compliance checking 27
2.3 Review per country 29
2.3.1 New Zealand 30
2.3.2 Singapore 30
2.3.3 Norway 31
2.3.4 Australia 32
2.3.5 USA 32
2.3.6 Korea 33
2.3.7 The Netherlands 34
2.4 Summary 35
2.5 Conclusion and recommendations 37

3 Part C: Conclusions and recommendations **38**
3.1 Data standards 38
3.2 Process implementation 40
3.3 Technical approach 43
3.3.1 Proposed solution 44
3.3.2 Proof of concept 45
3.3.3 Types of checks 46
3.4 Requirements for use of the Prototype 48
3.5 Reflection 48

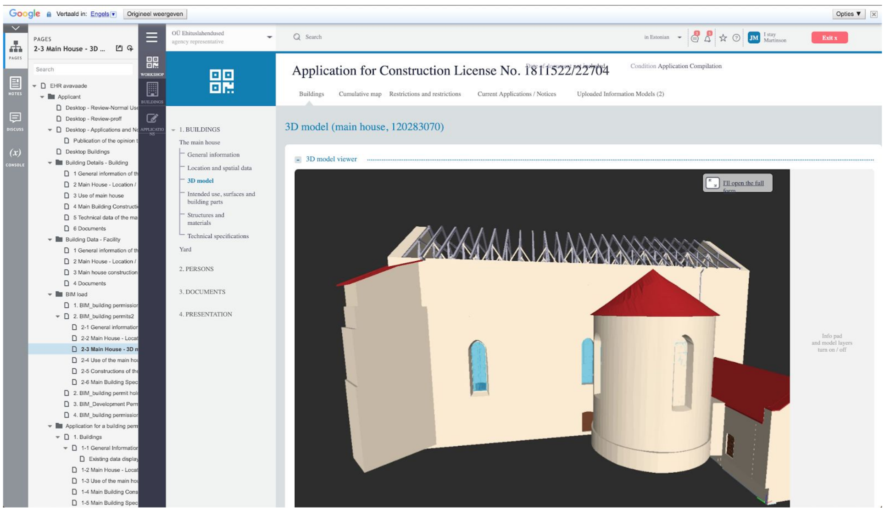
1 Part A: State of BIM in Estonia

1.1 Introduction

To be able to give good advice about an approach for the BIM-based model checking solution it is important to have a good understanding of the current state of BIM in Estonia. Not necessarily to lower the ambitions as discussed during the kick-off meeting, these are quite high. But to be able to give realistic advice including some steps to develop in that direction. To get a good overview four activities were executed as a preparation for writing the technical report.

1.2 Preparatory developments

The BIM-based process for building and usage permits project is part of a larger program 'Defragmenting the Construction Industry'. In preceding projects like the 'redesign of the EHR interface' a lot of work was already executed and reports were delivered which should be taken into account. Especially the PWC report which describes the optimization of the EHR and the EHR prototype which was setup gave a good impression on the chosen path. Based on the information we could extract from the PWC report, the Building Registry (EHR) prototype, Annex I from the tender specification and the overviews created during the graduation research by Chris Raitviir we got a better understanding of the checks currently executed for building permits. This was especially important for defining the desired abstraction level of the process description.

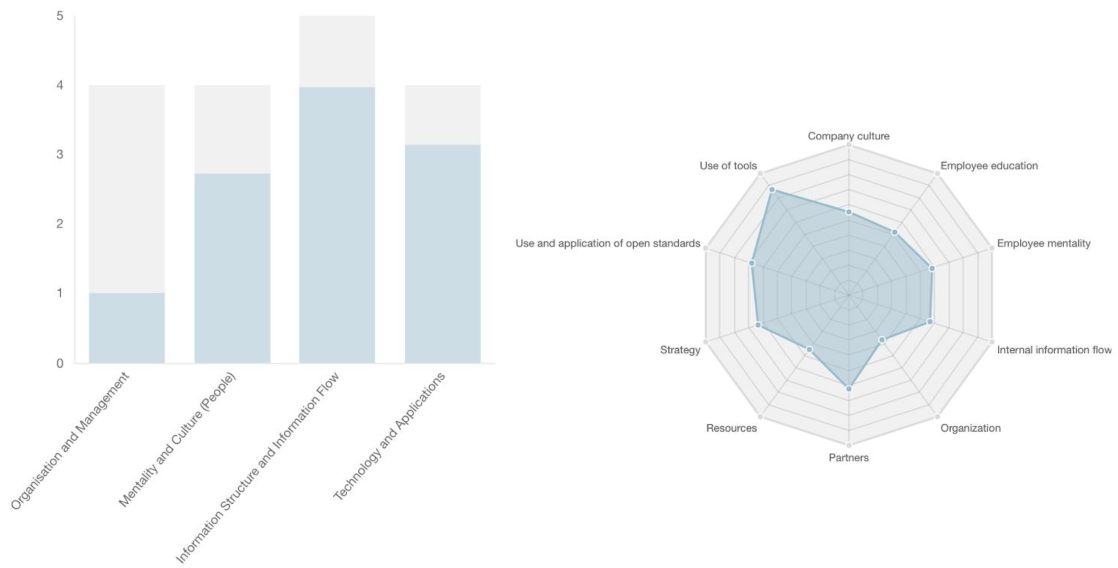


1.3 BIM Quickscan analyses

The first thing which was organized was the BIM quickscan analyses. An extensive questionnaire was sent to `around 50 people from the BIM industry in Estonia.

1.3.1 Introduction BIM Quickscan

The BIM Quickscan® is an instrument for common benchmarking performance for firms that are applying BIM.
The instrument aims to provide insight into the current BIM performance level of firms using BIM. The purpose is to raise awareness and establish a strategy for innovation with BIM, as well as to justify the qualification of the parties to be commissioned for projects.
The benchmarking instrument is based on a quick scan method. It combines quantitative and qualitative assessments of the 'hard' and 'soft' aspects of BIM.



An example of the results from the BIM Quickscan.

The BIM Quickscan® is intended to be used to scan an organisation over four main chapters that represent both 'hard' and 'soft' aspects of BIM, namely:

- Section 1: organisation and management;
- Section 2: mentality and culture;
- Section 3: information structure and informationflow; and
- Section 4: tools and applications.

Each chapter contains a number of KPIs in the form of a multiple-choice questionnaire. The total number of criteria is limited to 50 in order to keep an in-depth scan that can be performed with reasonable speed.

Within the first chapter (corporate management),the following KPIs are addressed: vision and strategy, distribution of roles and tasks, organisation structure, quality assurance, financial resources and partnership on corporate and project level.

The second chapter (organisational culture) focuses on BIM acceptance among the staff and workers, group and individual motivation, presence and influence of the BIM coordinator, knowledge and skills, knowledge management and training.

The following KPIs are composed in the third chapter (data-structure and information flow): use of modelling, open data standards, object libraries, internal and external information flow, type of data exchange and type of data in each project phase.

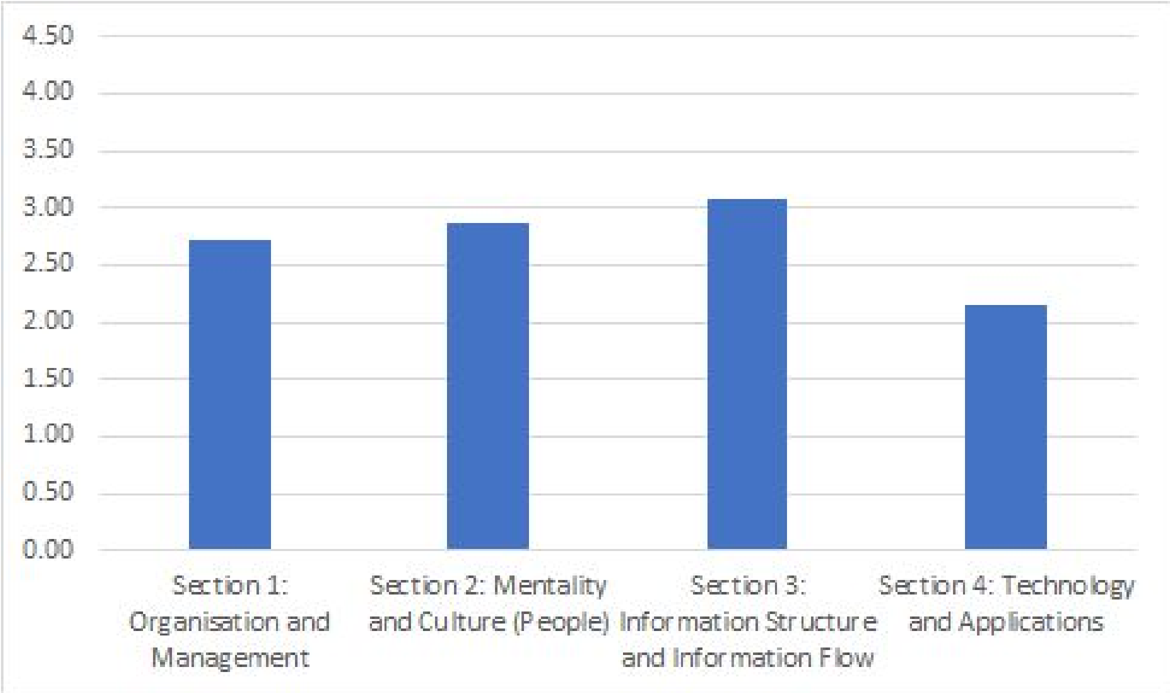
The hardware- and software-related KPIs are pulled together in the last chapter (technology platforms and tools): use of model server, type and capacity of model server, type of software package, advanced BIM tools, model view definitions and supporting rules

1.3.2 BIM Quickscans in Estonia

In total 15 BIM Quickscans have been filled out by 15 different users. The 15 quickscans covered 12 different companies. The Quickscan measures Chapters and Aspects. The Quickscans have been filled in between June 4th and June 20th.

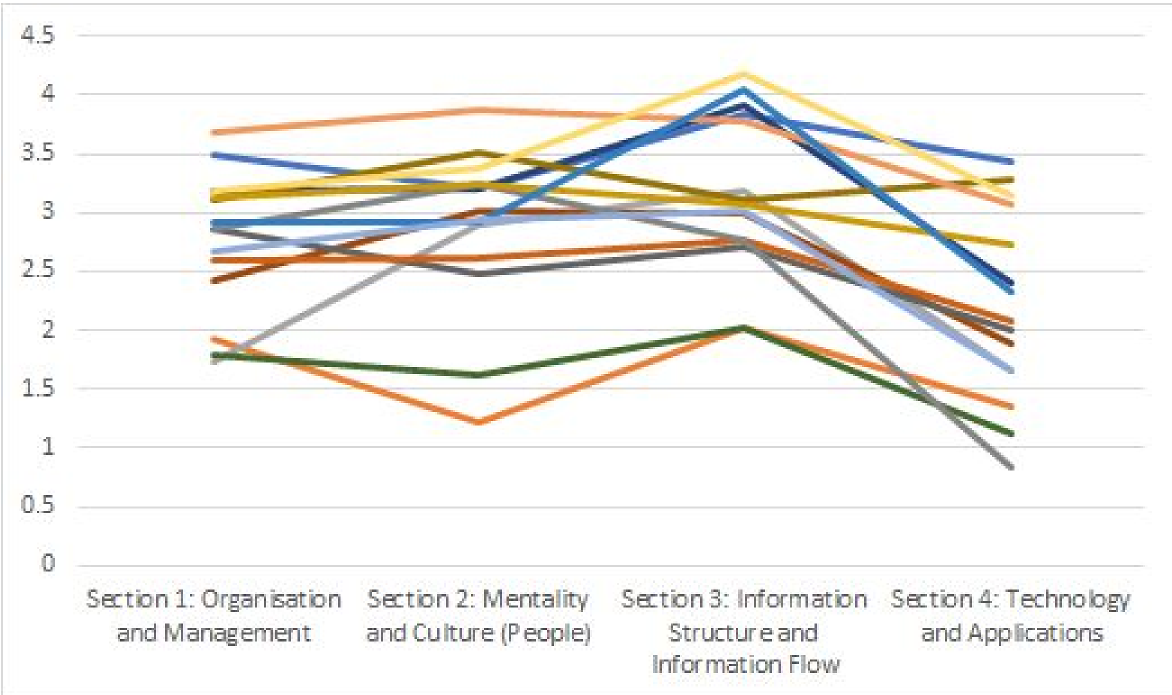
1.3.3 Chapters

The results of the BIM Quickscans have been very scattered. The average result per chapter is shown in the picture below:



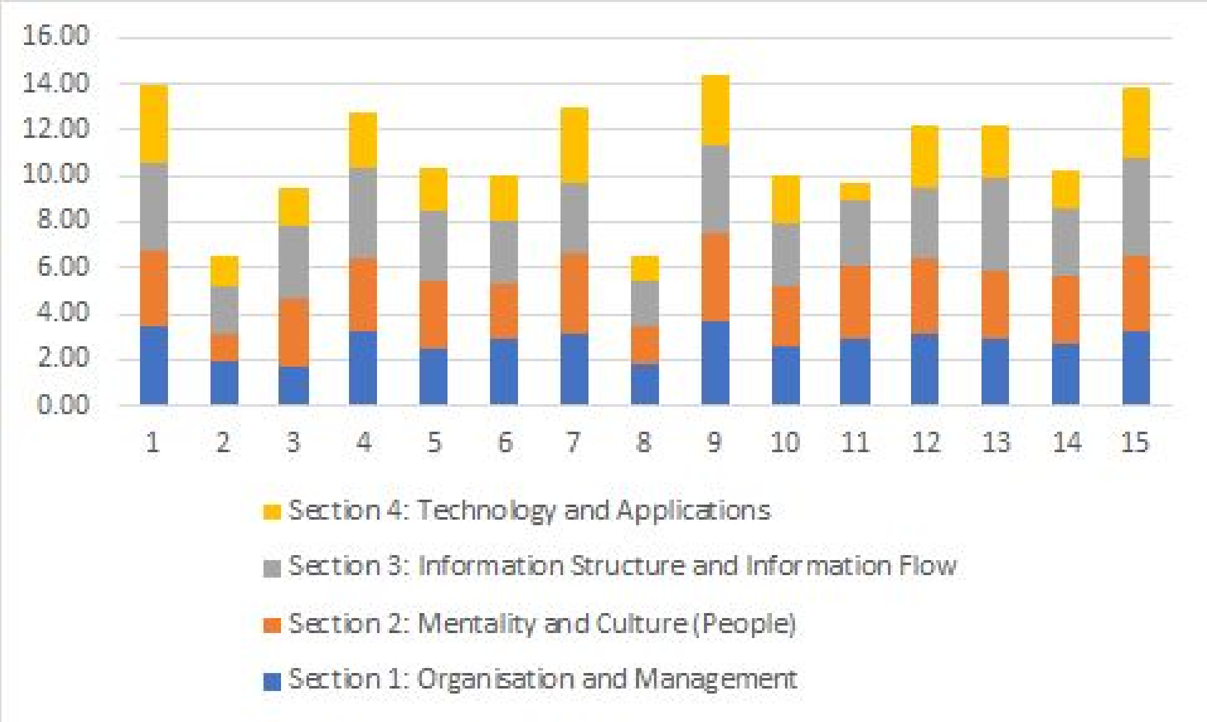
This average result is in line with the average results in other countries. This average shows that there is a strong focus on information structures and information flow. This chapter is seen as the most important to develop first when pursuing a high BIM maturity level.

When plotting the results from all 15 scans, the graph is showing more insights:



It is clear there is a wide variety of BIM Maturity levels between the 15 different participants. Results vary from a very low maturity to a very high maturity. It is

noteworthy that the level across all four chapters is consistent for every participant. This means the development of BIM Maturity is happening over the full extend of the digitalisation challenges.
This is very clear when showing the 15 Quickscans in a stacked way:



The results from the Quickscans are in line with international developments and results. The big differences between participant / organisation BIM Maturity is very common since the field is still developing.
It is encouraging to see that there are some organisations that are scoring above average on many levels. Specifically the group of organisations that is scoring between 3.7 and 4.2 on section 3 (information structure and information flow) can be seen as the trailblazers of BIM Maturity in Estonia. The main group scores between 2.5 and 3.2, and there are only two participants that are lacking behind a bit.

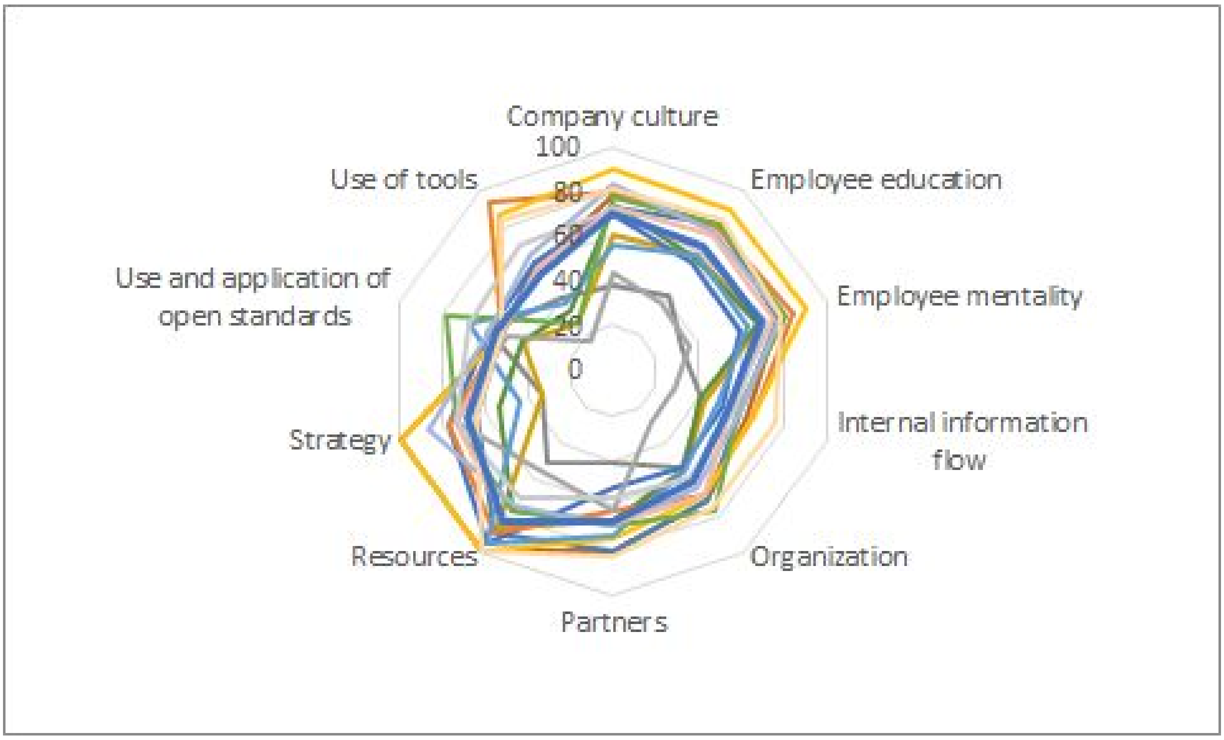
This ration between lead group, peloton and laggards is actually quite positive. Usually the lead group is smaller compared to the peloton, and the laggards are more. The ratio of 30% lead group, 60% peloton and 10% laggards is showing a very healthy and progressive outlook for BIM Maturity in Estonia.

Of course, these results are only based on 15 participants, that have been invited to participate in this research. Analyses of a broader group might show a different result.

1.3.4 Aspects

Besides the four chapters, the BIM Quickscan also provides results about ten aspects. The maximum score of the aspects is 100%. The results of the 15 Quickscans on these aspects is as follows:

On average there are many resources available, and there is a healthy focus on company culture and employee education.
Since we already saw a large difference between the BIM Maturity of the participants, this is also represented in the score on the aspects:



As shown in the graph, some organisations score almost 100% on some topics, while others score close to zero% maturity.
Again, this is in line with international results and paints a healthy picture of a developing industry.

The strong focus on information structures & flows (Section 3 of the BIM Quickscan) and employee education, combined with a very high score on resources, creates the ideal breeding ground for BIM Maturity to accelerate and flourish in the coming years.

1.4 BIM analytics

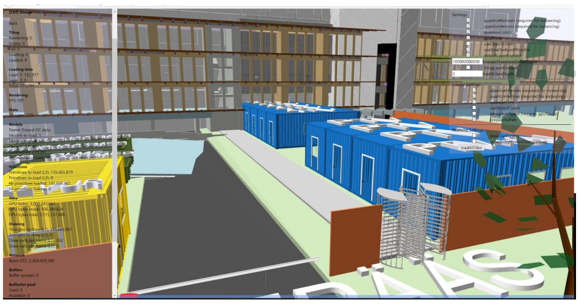
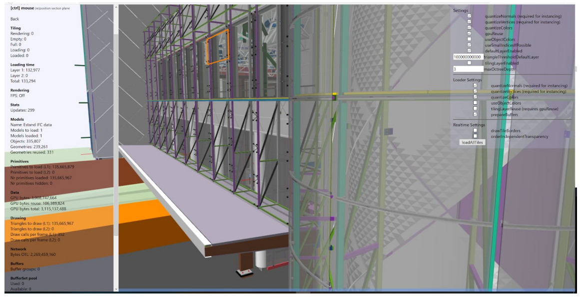
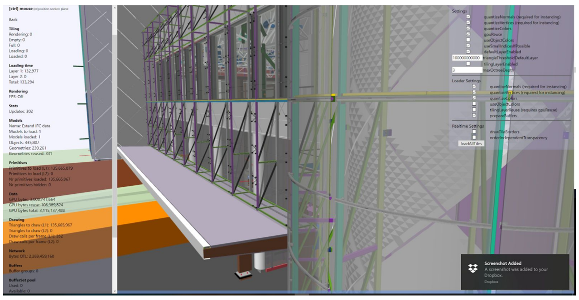
Based on recommendations by the ministry of Economics, In total four participants provided seven projects with a total of 40 separate IFC files. Four of these files contained invalid GUIDs. In practise this is not a big problem, but in the analyses tool that is being used for this research the data from these files could not be imported. Therefore these files have not been analysed. In total 36 files have been analysed.
These datasets have been provided by the Estonian users. It has to be noted the data is not from a design phase, as it would be in the case of BIM Based permit checking.

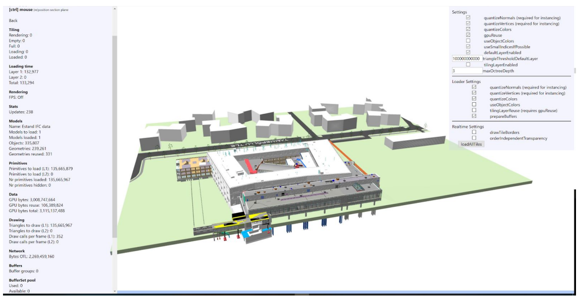
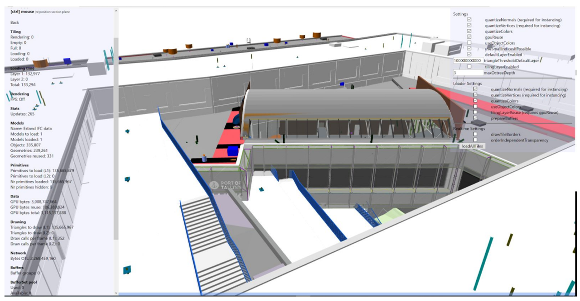
All of the files have been exported as IFC 2x3 data. Typical authoring tools that have been used are:

- Tekla
- MagiCAD
- Graphisoft
- AutoCAD
- Vectorworks
- CADS

Use of the Solibri Optimizer was also seen in the files that have been provided. The files have been generated between January 2017 and May 2019. Two files have been generated in 2017; 17 files in 2018 and 17 files in 2019.

Some screenshots of the IFC datasets to provide an impression of the geometrical detailness:

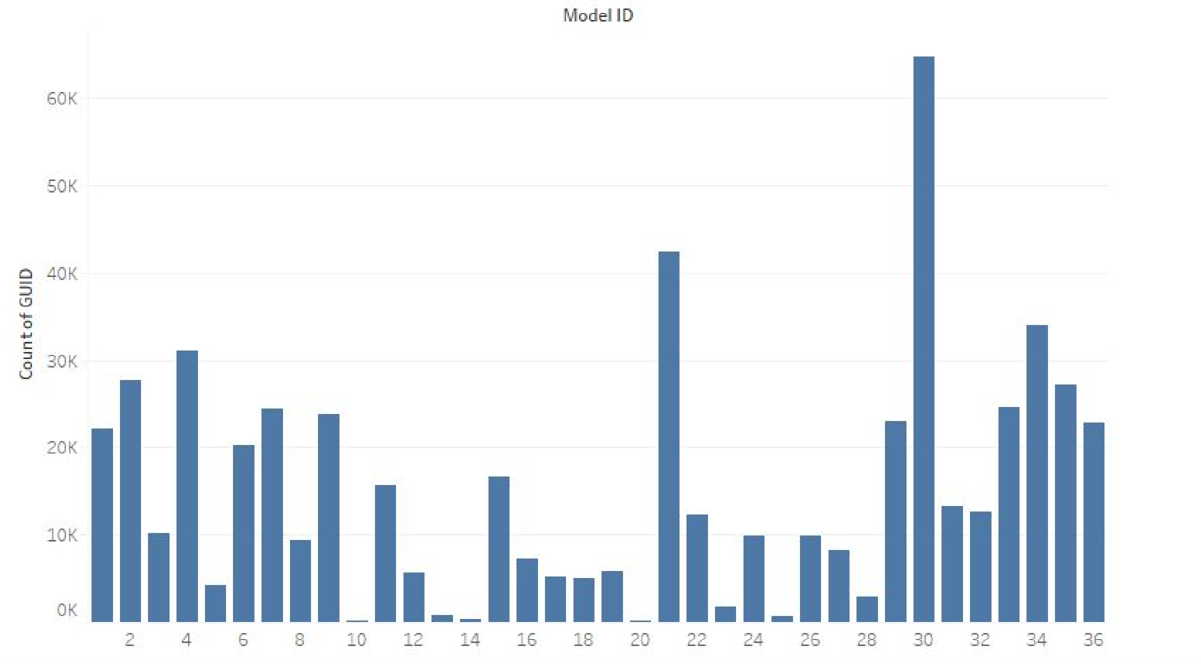



The following table provides an overview of the number of Entities, Relations and Products in each IFC dataset:

Model ID	Number of objects (IFC entities)	Number of relations (IfcRelationship)	Number of objects (IfcProduct)
1	399239	7837	22198
2	548473	3353	27650
3	230015	3026	10216
4	542813	762	31105
5	3750841	25372	4226
6	303099	19951	20210
7	719404	24074	24412
8	4508433	49164	9402
9	721441	31922	23780
10	555057	201	154
11	1292323	49362	15739
12	5670843	8194	5614
13	132176	7	853
14	43188	8	258
15	2941445	100288	16662
16	5174150	16488	7296
17	446893	14503	5107
18	373667	6868	5092
19	387259	7686	5825
20	133572	314	98
21	1785727	45758	44648
22	1556884	8056	12318
23	436291	10581	1867

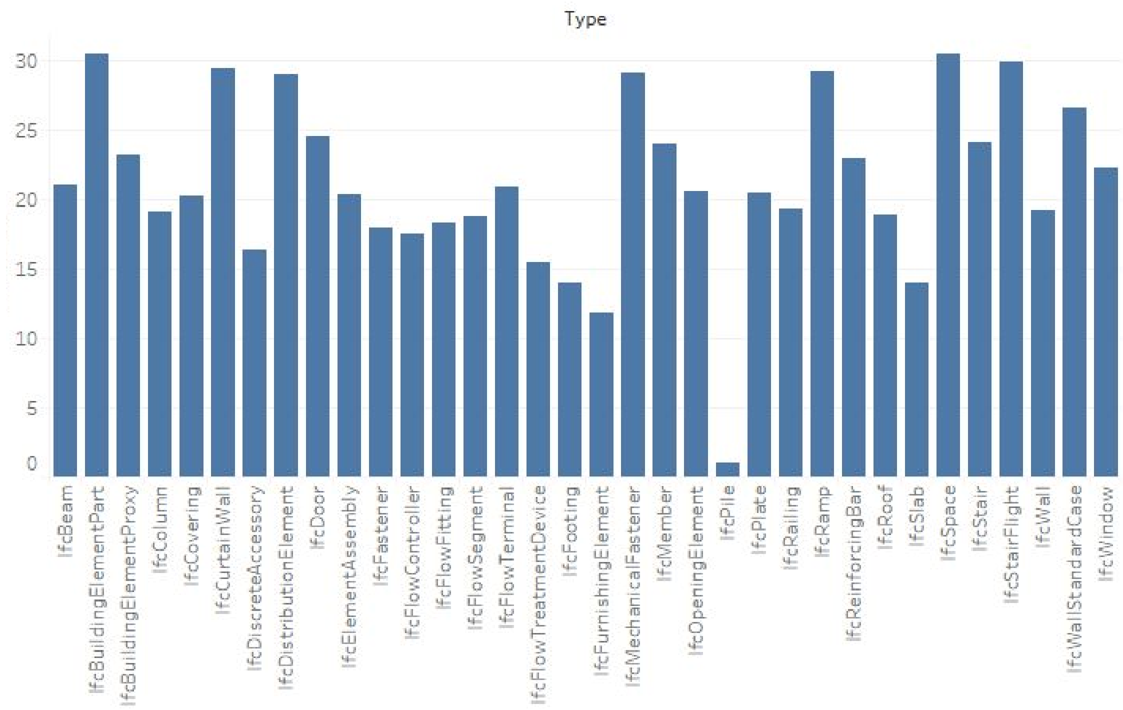
24	761028	10267	9912
25	443464	2860	1218
26	1272113	48258	28363
27	448535	12062	8194
28	276604	17018	8119
29	4058751	136058	23057
30	1017940	63778	64712
31	579815	17301	13368
32	647355	17379	12586
33	1206253	36005	24519
34	2676122	69231	33947
35	1551971	39984	27170
36	435524	19504	22855

Most interesting part of this statistics are the number of IfcProducts (the objects with a GUID). The number of IfcProducts per file is also represented in the following graph:



As shown, there are files with very little objects, and files with very much (>60 thousand). This does not say anything about the quality, but more about the richness of the datasets.

To get an indication of the type of IfcProducts in the different files, we created an overview of the average number of Types of IfcProducts specialisations used in the different files:



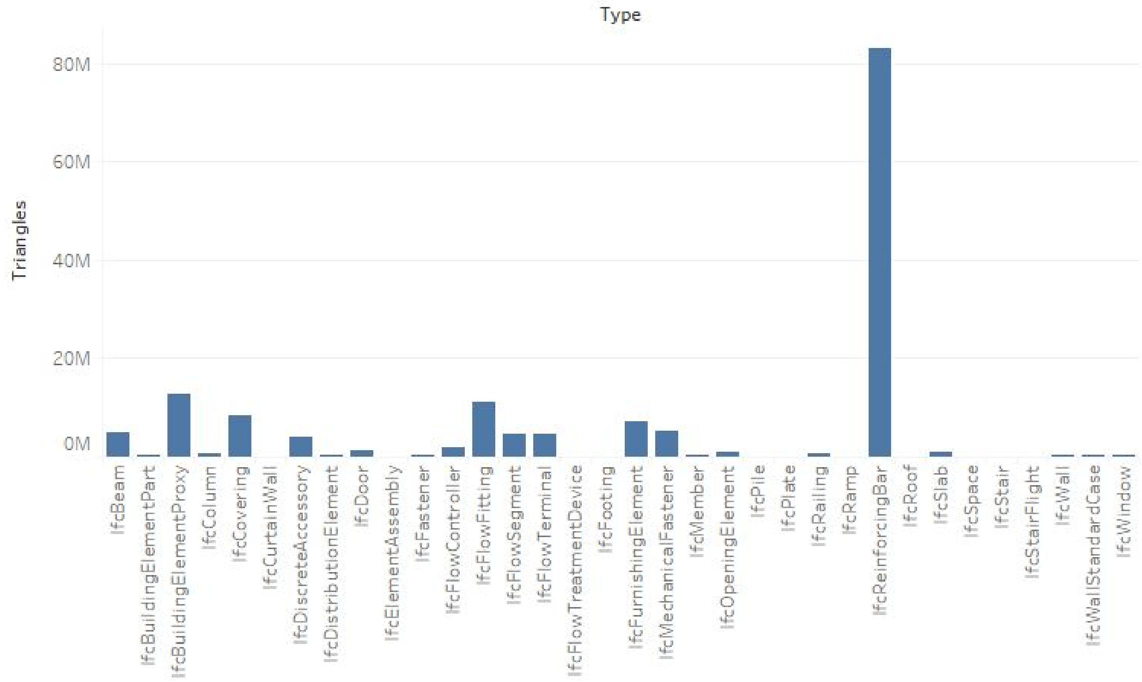
This shows a normal and healthy distribution of average usage of the different IFC types.

1.4.1 Geometry

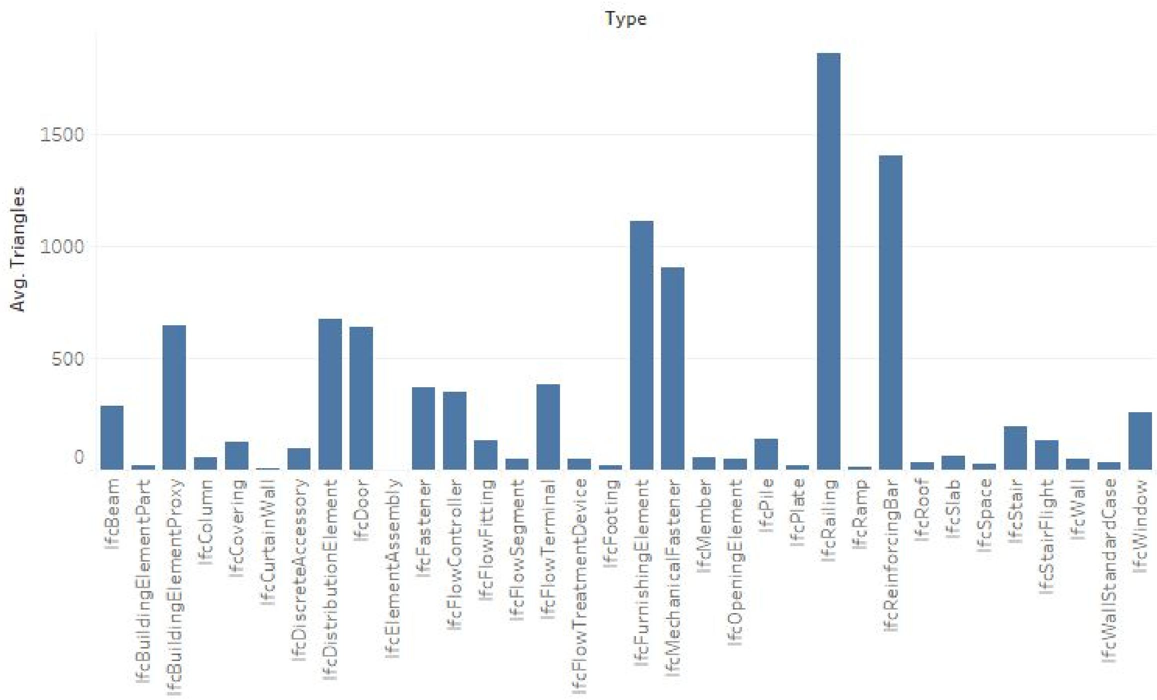
For privacy reasons, the files have been analysed as one large dataset. Further reporting in this chapter will be based on the full set of IFC data.

In this chapter we focus on the richness of the properties (number of properties per objects) and the geometry (number of geometrical triangles per object).

In the following graph the sum of the tessellated triangles is shown per IFC Type (again, over all 36 files):



Since the geometry richness is measured with triangles, this is of course depending on the total size of the objects. Therefore we focus on the average number of triangles per IFC Type:



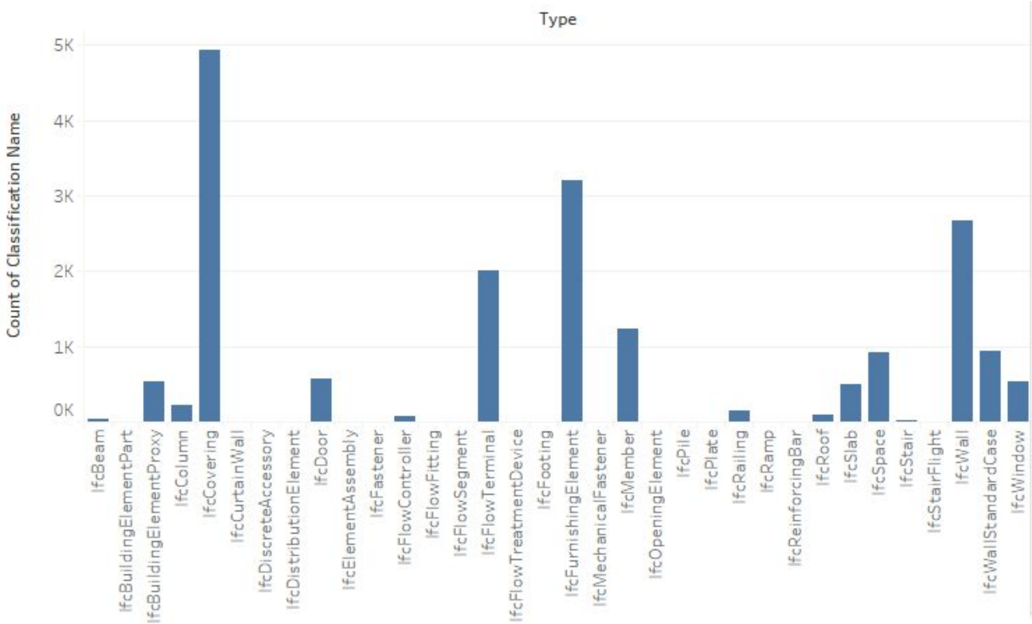
This provides insight that the average geometry richness per object is highest for the IfcRailing and IfcReinforcementBar.
This is seen often in IFC Analyses, and can be explained by the fact that most of these objects have many round elements that result in many tessellated triangles. This also explains the other peaks in the graph: IfcMechanicalFastener, IfcFurnishingElement and IfcDistributionElement.

Notable is the relative high average geometry richness for IfcDoor and IfcWindow.

1.4.2 Classifications

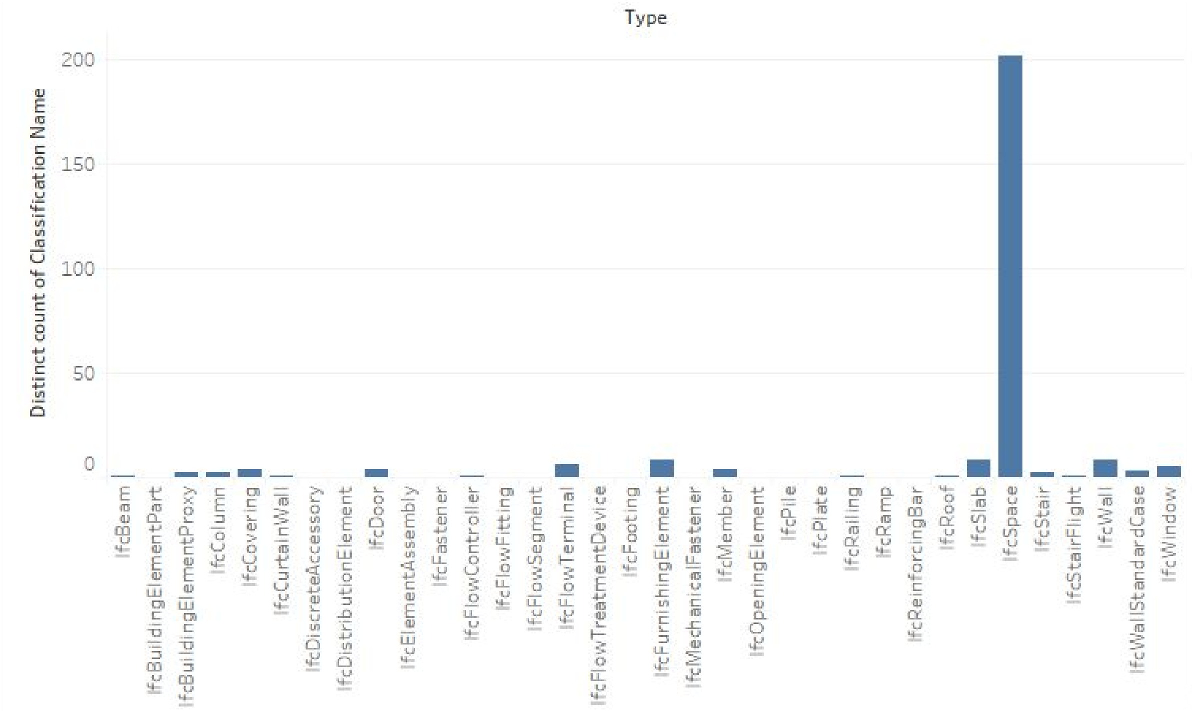
During the interviews and the kick-off we found that the use of a national classification system for building elements is not broadly accepted when using BIM exchanges. This is also what we found in the IFC datasets.
The use of IfcClassifcations and IfcClassificationReference is not used as much as seen in other international IFC datasets.
"Only" about 18 thousand IFC Products had classification names, on the total of over 570 thousand.
This is much lower than the international average where classifications are seen as a major contribution to the semantic richness of IFC data.

This next section will elaborate a bit on the classifications that have been found.
The following graph shows the number of classification names used for different IFC types in the provided files. Again, this is only for the 18 thousand objects that have classifications.



Specifically the classification of IfcCovering stands out and is remarkable. Further research revealed that the classification has the value 'www.rakennustieto.fi'. We can assume these classifications came along with the use of an object library in the authoring tool. Diving into the other types revealed a generic classification of 'Valvekaamerad, audio- ja videoseadmed, reaktiivvõimsuse kompensaatorid ja muud elektri seadmed' for the IfcFlowTermial, the same rakennustieto URL for furniture and 'Seinad' for walls. Another often used classification was the text 'Not defined'.

Looking at the data in a different way does reveal something else though:



A 'Distinct count' of classification names per type shows a spike in for 'IfcSpace' objects. This graph shows the number of different values used in the classification of the types. Again, this is only for the 18 thousand objects that actually have classifications.

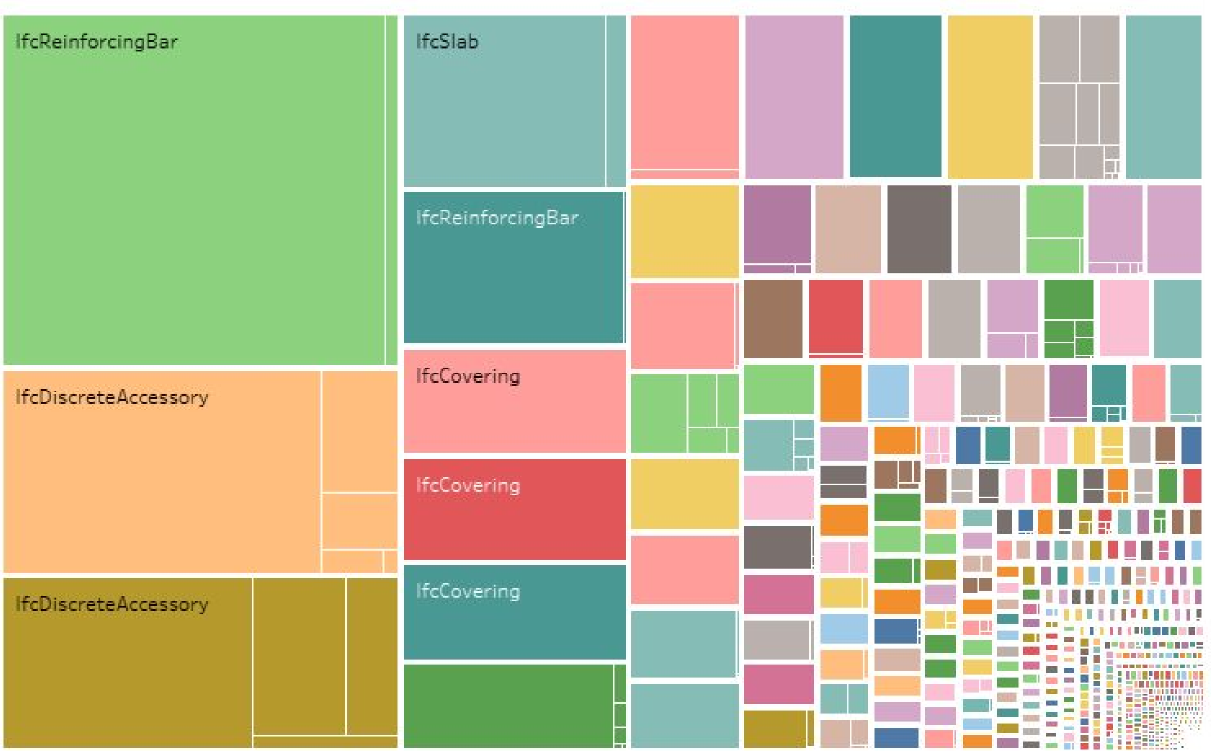
Looking at the data, there are some files that seem to be using an Estonian classification system to classify spaces. An example:

02.02.06	Abiruum
02.01.04	Abiruum
12.11.01	Abiruum
11.06.01	Ajutine tööruum

Although this is only found in a small number of files, it shows around 200 different classifications used for 914 different IfcSpaces.
This shows that the knowledge to classify spaces, and export it in the correct way to IfcClassifications is available in Estonia.

1.4.3 Materials

Another indicator for the richness of a dataset is the use of materialisation in the dataset.



In the chart above each block (each color) represents a material. The deviation between the blocks shows the types that have the material associated. The size of the colored blocks show the amount of times the material is used (also per type). This overview tells that the material Armatuuriteras (anchor steel) is mostly used for IfcReinforcingBar objects.
Another view of the same data, shows the use different materials used per type:

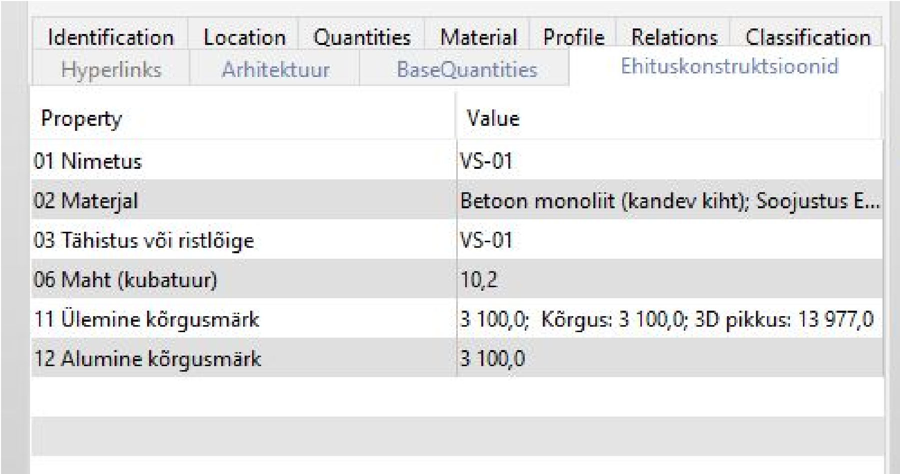
In this case the blocks (separated by the thicker white lines) are showing the Types, with a specialisation of the materials. In this case we see the Armatuuriteras again in the lower left corner for the IfcReinforcingBar. The reason why the Coverings are now in the upper left is because there are more IfcCovering objects then IfcReinforcingBar objects.
In the previous graph the IfcReinforcingBar was in the top left because there are more Armatuuriteras objects then any other objects.

This second overview shows the dominant materials in darker blue. As we see, the Armatuuriteras, STEEL, STEEL/Undefined and Terasplekk are the most used materials.

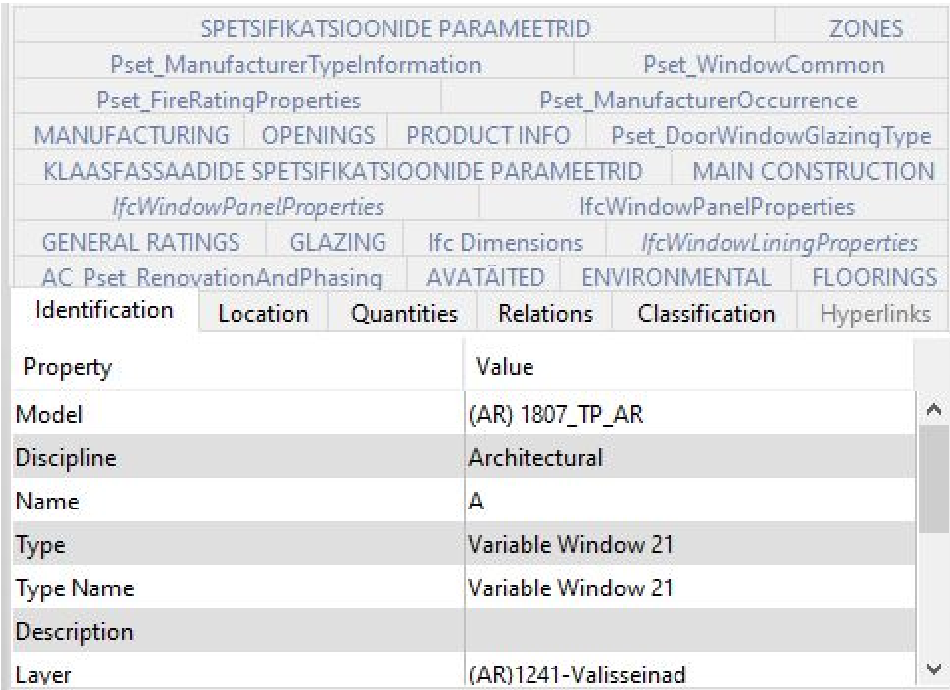
The total times a material is linked to an object is 273.947 times.

1.4.4 Properties

The final characteristic that is looked at, are the properties, propertySets and Psets in the data.
A manual look at the data revealed that there is a notable use of user defined property sets.



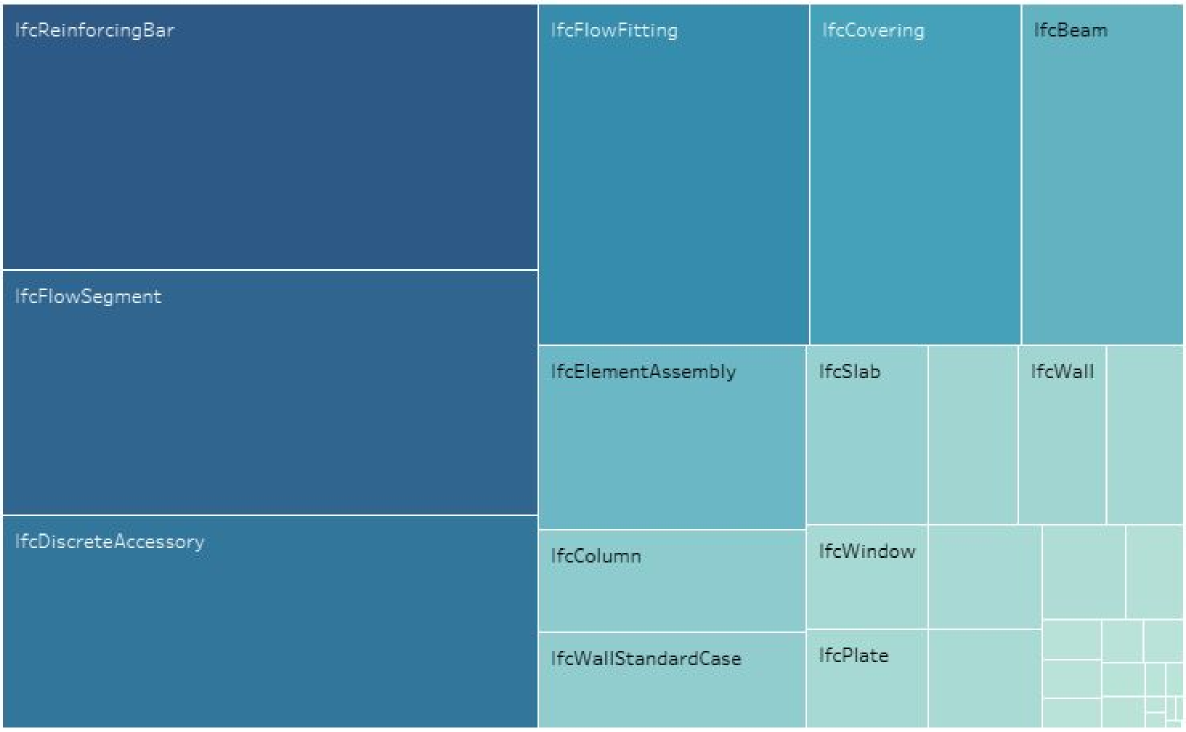
Property	Value
01 Nimetus	VS-01
02 Materjal	Betoon monoliit (kandev kiht); Soojustus E...
03 Tähistus või ristlõige	VS-01
06 Maht (kubatuur)	10,2
11 Ülemine kõrgusmärk	3 100,0; Kõrgus: 3 100,0; 3D pikkus: 13 977,0
12 Alumine kõrgusmärk	3 100,0



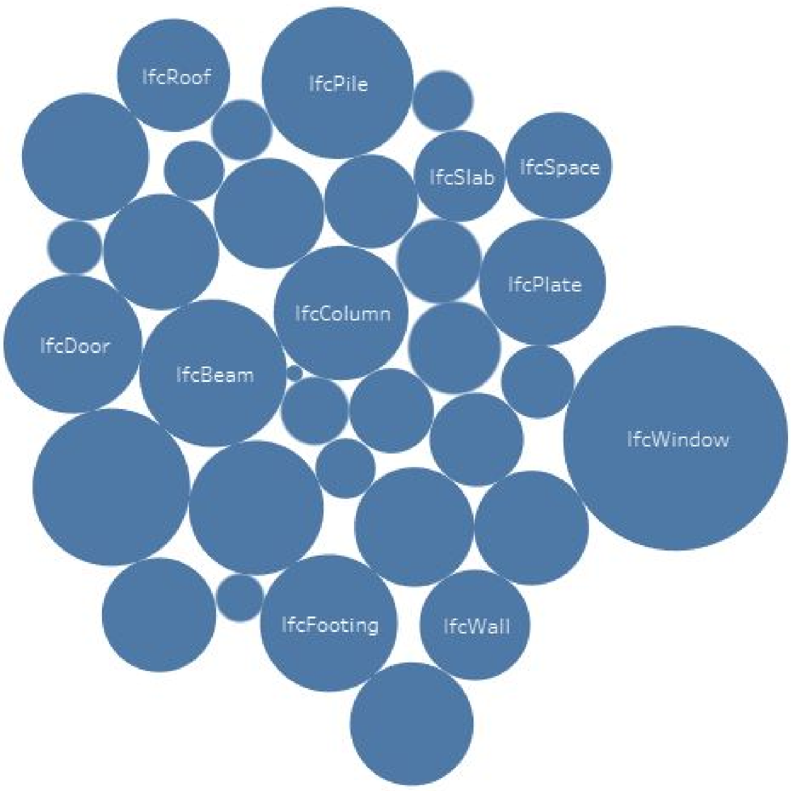
Property	Value
Model	(AR) 1807_TP_AR
Discipline	Architectural
Name	A
Type	Variable Window 21
Type Name	Variable Window 21
Description	
Layer	(AR)1241-Valisseinad

The use of property sets in general was quite rich. The active use of user defined (special named) property Sets shows the focus on information exchange, creating high quality data sets and shows a pragmatic approach to dealing with the lack of classifications.
It can be concluded from our analyses that the level of experience and education of the participants in this research is quite high. This proves that the industry in Estonia is on the right track to start a network effect of quality improvements throughout the supply chain.

The total number of properties in the 36 files was 7.1 million. It is no surprise that most properties are in the IfcReinforcingBar, and the MEP elements of the IFC structure:

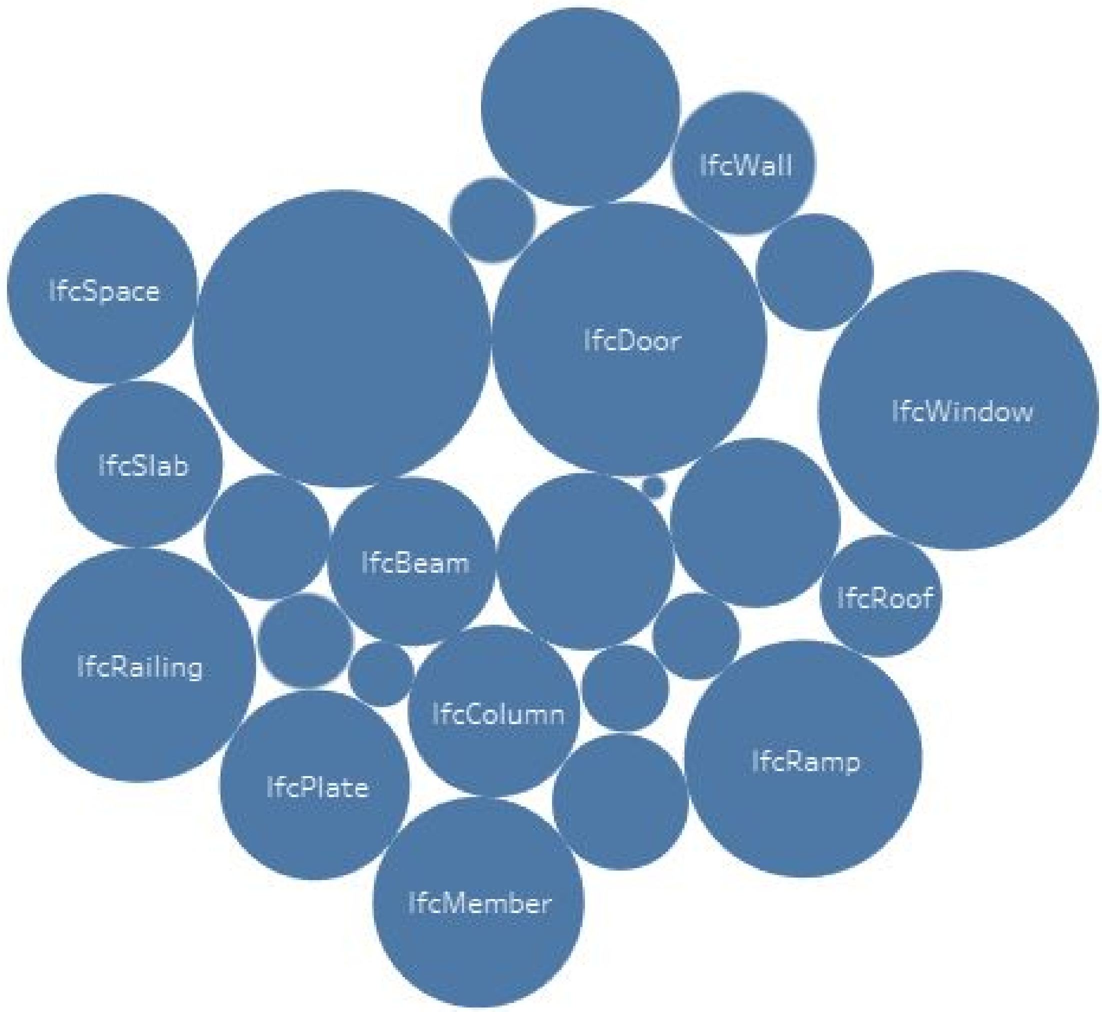


When we look at the average number of properties per IFC Type, we see the following graph:



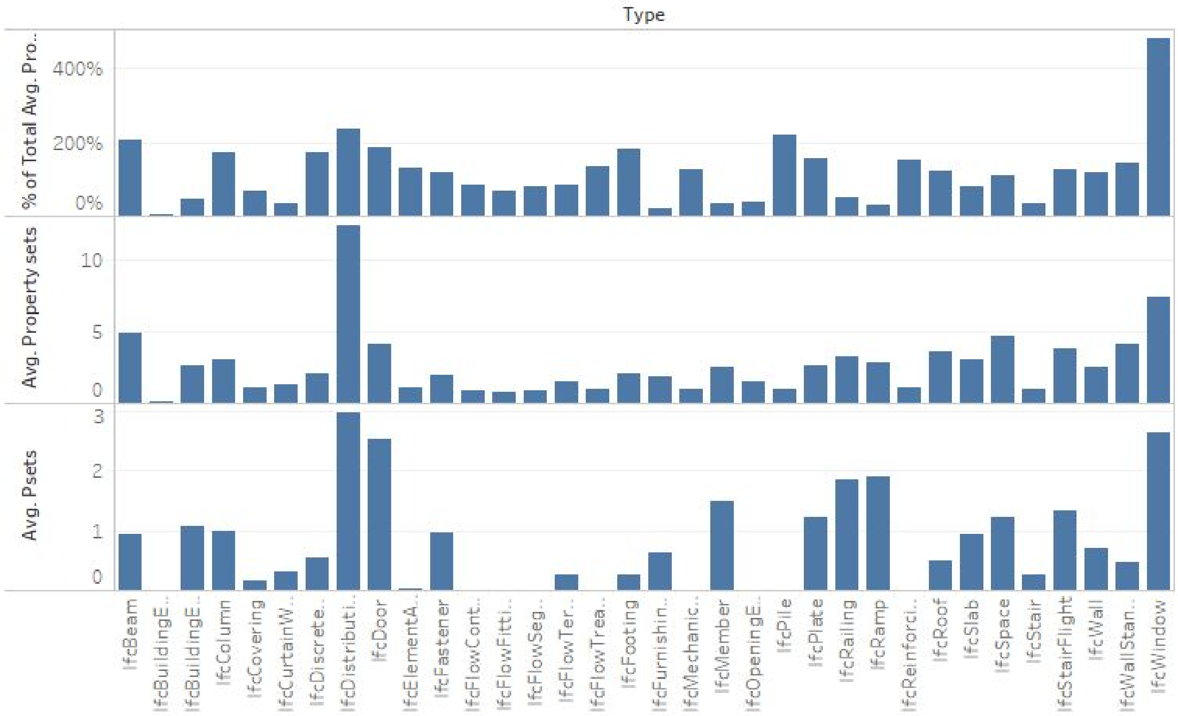
The common IFC Products like windows, footings, walls, beams, columns, doors, spaces, etc. have a high number of properties per object (on average, compared to other types in the dataset).

This is an overview of all properties. When we have a look at the standardised property sets in IFC (called PSets) we see this graph:



This result is very comparable with the overall property view. This means most of the properties that have been defined, are also stored in the official PSet structure from IFC. This is not so common and shows the high quality of IFC exports.

The last graph shows us the comparison between the average number of properties, the average number of property sets, and average number of Psets per IFC Type in this dataset:



This confirms our earlier statement that, based on this dataset, with these participants, we could conclude that the quality of IFC exports is higher than the average international benchmark for this.

1.5 Interviews summary

For the research, interviews were executed with several people working at different organizations and functions in BIM from Estonia. Most of them also sent the IFC data as preparation which have been analyzed. This gave a good overview of their level of BIM usage and the way BIM works in Estonia. The following interviews were executed:

Organisation	Function
Estonian Ministry of Economic Affairs and Communications	Head of Digital Construction Product owner EHR
Esplan	CEO
Novarc group	BIM Manager
Geospatial.ee	Owner
Nordecon AS	Development manager

Riigi Kinnisvara	BIM Manager

All interviews had the same structure. After a short introduction, usually the results of the Quickscan and BIM analytics were a good starting point to start the interview. This was a good way to discuss both technical and organizational topics. By going through these results topics like the daily work of the organization, the role within projects, collaboration with other organizations using BIM and the ambitions of the organization were discussed. Finally we also discussed the opportunities, risks and support for the 'BIM-based model check' project.

The first notable thing we would like to mention is the open and enthusiastic attitude of the people we interviewed. The BIM files for the analyses were easily shared and discussions about the results were very constructive. Both good and bad experiences were openly shared giving a lot of insight.

Regarding the technical part of BIM most of the commonly used BIM software like Revit, Tekla, Archicad but also MagiCad are used by the interviewees. IFC is widely used as shared BIM format to collaborate between parties working together on construction projects. They find it logical to work with this open standard so that everyone can make their own software choice and still collaborate. Also model checking, mostly using Solibri is commonly used by different parties. Although there is no national Classification Standard yet, organizations have set up their own, and really use them to optimize the standardization. Adopting a future national standard Classification system, for example based on CoClass, should not be a big challenge for them. Currently the state real estate company, Riigi Kinnisvara already develops all kind of standard requirements, which are already used a lot by other organizations. Also the Digital Building Clustre is mentioned often as platform which drives the BIM development to a higher level.

The initiative to setup a BIM-based permit checking facility by the government could count on a warm welcome. The main opportunities mentioned here were to speed up the permitting processes and to increase transparency in decision-making. Being able to test designs against the building code yourself in advance of the formal permit application was also found a great added value. Hereby it is assumed that the service is available online and that the results can easily be used in their own design software.

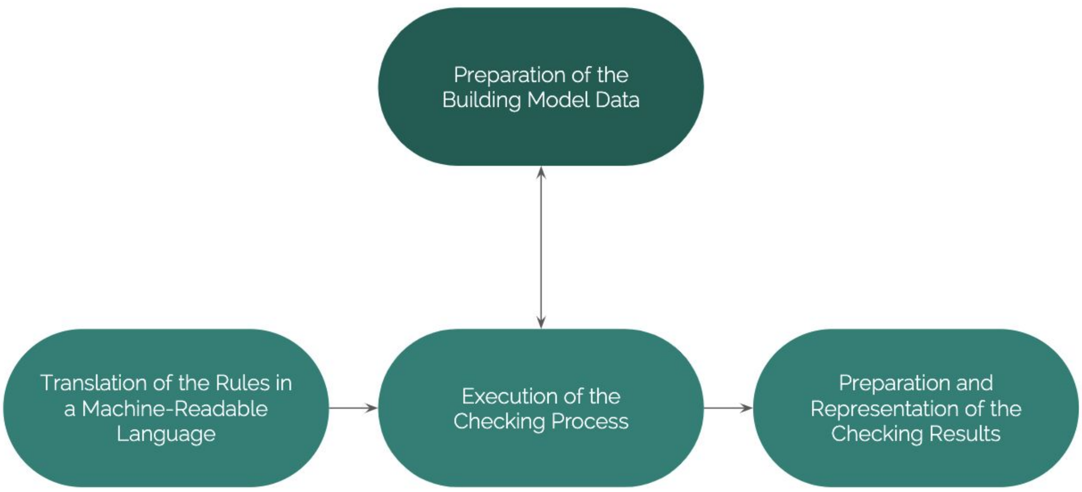
A lot of time can be saved automating 'the stupid checks' both at companies and at municipalities. Providing municipalities with an easy to use web based model checking service will also enable them to start using BIM. However, there is also a fear that so much extra information must be added to the design, that the extra time it takes to add it to the design no longer outweighs the time savings from the model checking. Finally the ambitions and possibilities for such a system are high, whereby it is alright to look at the experiences of others only with the desire to do better

2 Part B: State of the art outside Estonia

2.1 Rule checking versus compliance checking

Borrmann and Preidel discussed the major challenges of Automated Code Compliance Checking. Based on Eastman, they presented the common structure and basic components of the process. In general, the overall process is divided into four components:

1. Translation of the Rules in a Machine-Readable Language,
2. Preparation of the Building Model Data,
3. Execution of the Checking Process and, finally,
4. Preparation and Representation of the Checking Results.



The translation of the contents of the codes and guidelines into a machine readable language represents the starting point and is therefore the core task of an Automated Code Compliance Checking. Two essentially different approaches can be distinguished here:

The significantly easier way of translation is based on the direct transfer of the checking process in hard-coded program routines or methods. This means that the digitization of the contents of a code or guideline focuses on the definition of machine-readable algorithms, which are usually hidden from the user. Therefore, the readability of the translated rules for the user is limited and an involvement of a user in the encoding process is disregarded. As a result, the execution of the checking process is a hidden procedure, in which the user does not have an insight. Also, extensions and modifications are only possible by incorporating the software vendor. Such a process, which makes only the ingoing and outgoing information visible, but not the processing procedure itself, is called the Black-Box method. The major advantage of this method is the comparatively low error rate of the overall process because of the closedness and the direct access to the internal data structures of the code checking system.

In contrast to these hidden procedures, White-Box methods make the internal processing steps visible and therefore comprehensible for the user. To achieve this transparency, the contents of the translated guideline or code must be readable not only by the machine but also by the user. The rules must be translated based on a code representation system (a language), which is a system of symbols and rules. These elements can be used for the sufficient description of objects, methods and relationships. So the major target is not only to cover all kinds of information a code or guideline could contain, but also to enable the user to understand and retrace information at any time and follow the progress step of the checking procedure. Although the development and implementation of such a system requires significantly more effort compared to the closed checking approach, it has major advantages for the execution of a checking task.

The accuracy, correctness and consistency of the building model is a basic prerequisite for the following checking process and therefore a basic condition in order to produce resilient results. Although there is a continuous development of non-proprietary and open data standards, especially the IFC standard, many research findings point out that a complete correctness of the data standard can only be achieved by providing a formal rigid data structure. Therefore, the generally valid formulation of a checking process for a specific data standard is quite difficult and can only be realized by a preprocessing step, which checks and prepares the data model.
Since a rule usually applies only to a certain subset of data, it is recommended to create and prepare this subset before the rule is checked. Different subsets and derivations of a model for different purposes lead to a high workload for the modeller.
As a last step, the results of a checking process must be reported so that the responsible person can understand the intended meaning of the detected problem to be able to initiate the correct post-process, i.e. solve the detected noncompliance. Therefore, the detected problems should be presented as a written report or, better, digitally communicated to the responsible person, e.g. using the BIM Collaboration Format (BCF).

2.2 History of compliance checking

Dimyadi and Amor have provided an overview of compliance checking applications pre-2000 and post-2000. This chapter provides a summary of their publication.

The successful implementation of AISC Specifications in 1969 as a network of decision tables motivated a number of developments well into the 1980's. Examples include an advanced 3D graphical CAD system known as STEEL-3D for the design of steel frames to AISC Specifications, a software tool developed at Carnegie Mellon University for the design of reinforced concrete beams, an automated compliance checking system developed at University of Austin, and computerised building standards research at VTT Finland.

Following the successful implementation of SASE, two compliance checking applications were developed. SICAD (Standards Interface for Computer Aided Design) incorporates one way mapping functionality to assist the user to navigate, evaluate and extract required information from standards. Missing or incomplete data is managed by additional inputs from the user. This was implemented successfully as a design tool with AASHTO Bridge Design System and used for several years. The Standards Processing Expert (SPEX) was another software application developed in 1986 based on SASE as a knowledge-based system to determine conformance of component materials, structural and geometric properties with the design standards.

The application of computational techniques were explored by researchers in Australia in the mid 1980's. They came up with a prototype expert system called BUILD as a proof of concept. Some of these techniques were later used in the development of BCAider and DesignCheck.

In New Zealand, research into the application of expert systems in this domain saw the development of FireCode in 1987, which was used to check design conformance with a draft prescriptive Fire Safety Code. Other related software applications developed included "Seismic" for checking building design against earthquake and wind loading requirements, WallBrace to assess compliance with light timber-framed building standards, an object-oriented system "ThermalDesigner" for checking conformance of a residential building with the Thermal Insulation Code in 1992, and the ALF spreadsheet tool for conformance checking with thermal insulation standards for residential buildings.

In the US, Life Safety Code (LSC) Advisor (1988), a rule-based compliance checking tool was developed for auditing architectural plans against the prescriptive requirements of LSC (known as NFPA 101) that regulates building design for life safety and fire protection. LSC Advisor was later extended and developed into an expert-system Fire Code Analyzer (FCA), at Massachusetts University around 1991. FCA is closely related to SICAD and uses a frame-based architectural model representation, a set of rules as well as some geometric algorithms. EXPOSURE, an expert system version of NFPA80a was also developed around this time for fire protection design of building exteriors.

iCADS (Intelligent Computer-assisted Design System) was another expert-system example developed around 1990 with an extended knowledge-base covering space layout, structural system selection, day-lighting, artificial lighting, noise insulation, climate control and energy conservation, and construction costs. It incorporated a CAD system, a geometry interpreter, a relational database and an expert design advisor.

Since the emergence of the IFC open data model, we have seen the development of several important tools that are being used today, namely Express Data Manager (EDM) Suite (now incorporating EDMmodelChecker), Solibri Model Checker (SMC), Fornax plan checking tool, Avolve plans review, Design Data System (DDS), etc.

In 1995, the Building Construction Authority (BCA) of Singapore initiated the CORENET (Construction and Real Estate Network) electronic consent submission system incorporating an in-house developed Building Plans (BP) Expert System to check 2D plans for compliance. The system was upgraded in 2002 to CORENET e-Plan Check replacing the 2D BP Expert System with the 3D IFC data model. Currently the BCA is looking into upgrading the system again, based on feedback from IFC usage in practise.

Express Data Manager (EDM) Suite was developed by Jotne EPM Technology in Norway in 1998 as an object database with tools to manage complex Product Data Models. It started out as a collaboration tool, but has since incorporated several additional modules including EDMmodelChecker that supports open development using the EXPRESS data modelling language (ISO 10303-11).

BCAider was an expert system released in 1991 by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Australia. It was commercially available for compliance checking against the Building Code of Australia (BCA) until 2005. In 2006, CSIRO announced DesignCheck, a new system that incorporated EDM as the core rule bases and compliance checking engine for the BCA. DesignCheck has not been used commercially and it appears that there is no plan for further development.

Solibri Model Checker (SMC) was developed in Finland in 2000 and started out as a BIM model quality assurance and validation tool, but has since developed into a stand-alone graphically-driven rule-based compliance checking and reporting application. SMC has a set of built-in rules that can be managed by a ruleset manager. A ruleset can be replicated, but the extent of user customisation is limited to changing parameters.

The US Department of Energy produced and published ResCheck (Residential Compliance) and ComCheck (Commercial Compliance) to allow anyone to check a building design against the applicable energy standards, e.g. IECC and ASHRAE Standards 90.1. Both of these compliance checking applications have all the standards criteria hard-coded into the tools, although managed by the government department that have control over any amendment to the standards.
Similarly, the US GSA (General Service Administration) Courts Design Guide automation project also incorporates an independent ruleset manually derived from the textual standards.

An expert system of the nineties that has survived the test of time is Design++. It has been developed into a knowledge-based design automation tool in conjunction with BIM. Design++ has been incorporated into a number of commercial products including Bluethink's House Designer. Apart from giving advice to designers based on the evaluative rules, this system can also incorporate a set of generative rules for creating objects automatically. Again, the rulesets are encapsulated into the application and can only be managed within the application.

Another project in the US is a collaboration between ICC, Solibri and Fiatech together with a few other software companies to develop AUTOCodes. All of the approaches discussed so far, including current commercial systems, appear to have one thing in common. They all use an independent regulatory data representation either directly or via other dependent systems, and the representation is hard-coded into the system and is subject to manual updates by the software developers.

2.3 Review per country

Approaches to developing automated building code-checking have been reported in the literature for a while now. This chapter cites from several sources like:

- Automated compliance checking using building information models, by Greenwood, Lockley, Malsane and Matthews
- Automated Building Code Compliance Checking, by Dimyadi and Amor
- BIM-Based Code Compliance Checking, by Preidel and Borrmann

2.3.1 New Zealand

Since 2013, the University of Auckland in New Zealand has undertaken major research on practical aspects of automated compliance audit processes. The research attracted some interests from local and national government bodies culminating in the commercialisation of the work in 2019 by Compliance Audit Systems Limited (CAS). Due to its practical approach and support of open standards, the New Zealand government has recently commissioned the translation of a core set of the New Zealand Building Code into computable rules encoded in the LegalRuleML emerging standard to serve the priority consenting requirements. The ACABIM approach is built upon the philosophy of human-guided automation, which alleviates the need for human experts to perform trivial tasks that they find tedious to do. The impact of the approach is to minimise human errors and expedite processing by machines.
ACABIM supports three primary external input components, namely ISO-standard BIM model exchanged in IFC, BPMN-compliant workflow model, and LKM, as well as secondary supplementary human input and interfacing with simulation tools.
The IFC and LKM input components are treated as external data sources upon which the workflow model can query and extract data.

2.3.2 Singapore

The BP-Expert system had been available in Singapore from as early as 1995 for checking 2D drawings. In 2000 it was replaced by e-PlanCheck as part of the Construction and Real Estate NETwork (CORENET) project. CORENET was one of the first initiatives in automated code-checking, and was funded by the Singapore Ministry of National Development and carried out by the Construction and Real Estate Network.

The basic intention of this platform is to collect all kinds of information related to a construction project and optimize the processes with the help of digital methods and tools. One of these tools is the application CORENET BP-Expert, which aims to check the compliance of digital 2D-based drawings with regulations regarding accessibility and fire safety. In 1998 CORENET was enabled to work with the IFC standard and therefore extended by 3D compliance checking. The current version of the tool was first published in 2002 as CORENET e-Plan Check and provides a code compliance checking feature of a digital

building model regarding a large extent of the Singaporean regulations in terms of building control, accessibility, fire safety as well as environmental healthcare.
The checking processes within CORENET are based on hard-coded routines and therefore the algorithms, process steps and methods are not transparent for the user. The overall process is structured into three basic phases. In a first step, the model information is checked for availability of the information in the required form to be processed. Subsequently, in a second step, the model is searched for the missing information in underlying information layers. If the missing information cannot be found here, it is created in a last step with the help of information derivation.
Currently BCA is re-evaluating the CORENET system and is in dialogue with the software industry to rebuild the system based on feedback from the AEC industry. While the use of BIM and IFC is adopted more widely in the AEC industry, there started to arise a mismatch between the data requirements for CORENET checking, and the modelling methods the industry uses to create BIM datasets.
In 2017 first experiments have been done in collaboration with divers software developers and researchers. In 2018 a grant call has been published for a BIM submission/collaboration system. Currently developments are on the way.

2.3.3 Norway

The CORENET work was developed and emulated in Norway with the ByggSok system. This is an e-Government system comprising three modules: an information system, a system for e-submission of building applications and a system for zoning proposals. Driven by the Norwegian Building and Construction industry and supported by Standards Norway and Norwegian buildingSMART it is heavily based on IFC standards. The work is ongoing and currently focussing on the issues of classification, terminology and standardising rule-checking in construction at an international level. Building upon their e-PlanCheck pilot projects Norwegian developers (Statsbygg) have experimented with multiple systems as part of their efforts to extend the use of IFC to the entire project life cycle in support of their mandate that by 2010 all properties will use IFC based BIM. The resulting systems have been piloted on real projects, with data being exchanged through a wide selection of software to suit the various stages / tasks of the project lifecycle. On the HITOS pilot, the code checking efforts have focused predominantly on accessible design. Here the building model data are stored and accessed through EDM Model Server in IFC format. The accessibility rules are parameterised, mapped to their associated building objects and executed using Solibri Model Checker's Constraint Set Manager. Solibri communicates directly with building model data in IFC format, but retrieves only the objects it needs – i.e. those mapped to the accessibility rules. The rules implemented to date focus predominantly on geometrical constraints and as such the objects and parameters are supported by the IFC data models produced by current BIM packages. The Statsbygg Solibri system does not support the enhancing of these data models or the export to IFC format, and so cannot currently be used for compliance checking of attributes not supported by the current BIM vendors. The Solibri Constraint Set Manager is implemented in java and ships with a library of built-in parameterised rules which can be configured by adjusting the parameters. New rules, however, must be custom made in collaboration with the Solibri software developers and as such are not easily adapted for other software. Solibri has the benefit of powerful 3D modelling engine which, in combination with the ability to directly read IFC files, allows for clear visual reporting of rule infringements for the user. Solibri's built-in rule library contains rules for validating a data model prior to rule checking which is useful.

2.3.4 Australia

Both the Solibri Model Checker and Express Data Manager were considered as possible platforms for automated code checking in Australia again focusing on accessible design regulations. The work was undertaken by Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the University of Sydney and was funded by Australia's Cooperative Research Centre for Construction Innovation. EDM was eventually selected and the resulting automated code checking system – DesignCheck has been on trial by the construction industry in Australia. Design Check uses object based rules, encoded using EDM. Building data models, in IFC format, are imported into the EDM database and transformed into the Design Check internal model. The Design Check model includes building code specific information not currently implemented by BIM vendors. A mapping schema, written in ExpressX translates the building data model from IFC format into the DesignCheck schema. The strategy is similar to that of e-PlanCheck in Singapore; however, DesignCheck has the advantage of supporting the ability to check for compliance at various stages in the design process, as it has a rule schema for early and detailed design stages as well as for specification. It is therefore targeted at Architects and Designers rather than just Building Control certifiers. DesignCheck did not have the ability to view 3D models and all reports are text based.

2.3.5 USA

Similar work on code-checking began in the United States around 2000, with the initial emphasis on health, safety and welfare. A major driver of BIM and validation of BIM models in the United States is the US General Services Administration (GSA). The GSA issued BIM-guidelines in late 2006 (GSA, 2006) and in 2007 proposed that all planners seeking funding for their spatial planning projects would need to produce BIM models for validation as an open standard (GSA, 2007). SmartCodes was a project driven by the International Code Council, in conjunction with AEC3 and Digital Alchemy. This project has focused largely on addressing the problem of transforming paper-based codes (of which there are thousands) into machine-interpretable rules; generally a lengthy process requiring many iterations between Building Code officials and software developers. In order to streamline this process the SmartCodes project developed a methodology for applying tags to electronic copies of Building Codes using a 'tag dictionary', or ontology. The rules are then automatically extracted, following a strict mathematical pattern, into an IFC constraints schema. The resulting IFC constraints schema is mapped to the IFC building data model via the tag dictionary. The rules can currently be executed using either Solibri Model Checker, or AEC3 XABIO.

2.3.6 Korea

Taking advantage of BIM, the Korean government is establishing automated building permit system. As a part of the Korea government-led initiative for enhancing building design productivity as well as overall quality and performance using BIM, automated checking of Korea building permit requirements and rules is the motivation of this paper. The project can be summarized as follows:

1) development of an open BIM based quality verification system,

2) development of an open BIM based construction document optimization standard and applied technology,

3) development of an integrated cooperation work system during the public administrative process.

Several software tools that assist in the automated review process are also under development and will soon be publicly available.

KBimCode is a script language. It is developed to represent the Korean Building Act as explicit computable rules. The application of KBimCode can be summarized as:

1) Preparation of natural language rule,

2) representation of rule,

3) definition of rule,

4) execution of rule, and

5) execution result and report.

Natural language rules are translated into KBimCode. KBimCode is an ongoing project and will contribute to establish automated design quality assessment system for building permit in Korea.

2.3.7 The Netherlands

The Ministry of internal affairs of the Netherlands, in collaboration with the Netherlands organisation for applied scientific Research TNO, are currently developing a proof of concept online tool for automated code compliance checking. This is done based on the concept of BIM Bots, to facilitate the industry with an online tool that can be triggered from any other workflow or tool that is used by the industry. Results are being returned in the buildingSMART BIM Collaboration Format (BCF).

Core principle of the Proof of Concept in the Netherlands is that the tool should work on 'any regular IFC dataset currently used in practise'. This means that they are trying not to put many requirements to the industry on how to model the data, and where to put different properties, just for the code compliance check.

The quality of IFC usage in practise is quite good in the Netherlands, and the adoption of the "BIM Base Information Delivery Specification" is high. The team developing the tool in the Netherlands is taking that practical base as the starting point for checking. This increases productivity because the industry does not have to create a special dataset just for code compliance checking.

This tool can be seen as a 'black box' approach, with sophisticated algorithms to check the data in a fast way. Another reason why this approach is chosen is to eliminate 'false positives' that may arise in rule based checking. Two typical examples are mentioned in the project:

1) The checking of the slope of a ramp: Rule based checking would search of IfcRamp objects and check the slope based on the requirements. But when users generate an IfcSlab, IfcProxyElement, or any other IFC Type, it will not check the rules for slope, because the rule is only set on ramps. Lower quality IFC exports could lead to false positive results in a rule based code compliance check.
2) The check on fall protection: The building decree says that a person may not fall more than one meter, and that any place where this might happen in a building needs a fall protection of at least x centimeters. This often leads to a checking rule that checks all the walls, fences and railings to make sure they are x centimeters. But the places where people can fall are not always guarded by walls or fences. There could be another object that is valid (a concrete flower box) or an object that does not comply to the rules, but passes the test (a door). Furthermore, the building decree also states that people have the ability to climb elements between 20 and 70

centimeters. An object between this range could be close to the fence, and therefore the fence of x centimeters from the floor is too low to actually meet the requirement, because a person can stand on the element and be higher. Separate rules to check separate parts of the building decree could lead to false positive and false negative results in a rule based code compliance check.

The reason why the Netherlands ministry of internal affairs is experimenting with this approach is to increase the productivity of the construction industry. The base point of this initiative is to make code compliance checking possible on an IFC dataset that is generated by the industry anyway. The ambition is to eliminate the need for a specific information guideline, template or requirement to generate a specific IFC structure just for the reason of code compliance checking.

2.4 Summary

There seem to be a couple of returning factors in the development of code compliance checking based on BIM:

- White box versus Black box
- Rule based versus algorithm based
- Validity of the input data versus validity of the rules and tools
- Focus on workflow and process support versus focus on automated checking

These factors are related and need to be looked at in an integrated approach.

From a general perspective, it is important to distinguish white-box solutions, which implement an open, transparent approach for rule representation based on an accessible rule repository, from black-box solutions, which typically rely on coded implementations of specific regulations. Whereas white-box solutions are open to verification, modification and extension by domain experts, black box solutions are currently more powerful as they can make direct use of internal algorithms and data structures of the code checking system.
To use the best of both worlds, we suggest to use an open source solution, where the algorithms are published and transparent for users to evaluate. This means the solution will have the powerful algorithms, but still used based on open data standards and the open accountability of a white box solution.

The results of a checking process are highly dependent on the correctness and availability of the information in the underlying BIM dataset.

Since a process cannot produce correct results based on incorrect information, this correctness of the digital building model is often an essential prerequisite for following code checking processes. Although the overall correctness of a model is often subsumed under the general term data quality, it can be divided in two parts: the formal part and the content related correctness of a BIM model.

First of all, the information provided by a BIM model must fulfill formal criteria, which means that the information follows defined "grammar rules." Usually these rules are defined by the syntax of the data model, e.g., IFC, which is used for the representation of information. These requirements can be extended by further project-wide requirements, which the project participants contractually agree on in the beginning of a building project. Such requirements can contain individual constraints which information should fulfill in order that all the stakeholders have a unified interpretation of the contents of a BIM model, such that the information is interpreted in the same way. Templates for project-wide requirements can be found in several guidelines, such as in the GSA BIM Guide[1]. Usually these requirements are written down in modeling requirements and provided as part of the BIM Execution Plan. Since formal criteria are straight-forward rules (e.g., checking the availability or the data type of certain attributes), these conditions are easy to check. Unlike the formal criteria, the content-related criteria are significantly more complicated to check, since they require the interpretation of the information.

Content-related checks include, e.g., the compliance with reasonable boundaries and the consistency of the provided information.

Next to the correctness of single BIM models, the validity of multiple composed models must also be taken into account. According to the Federated Model Approach, each stakeholder, who is responsible for discipline-specific model contents according to the requirement specifications, has to submit a BIM model when reaching specific milestones. These submitted models result in a composed overall model, which finally describes a comprehensive description of the building to be constructed. Required information for model checking often refers to different discipline models and so not only the quality of a single model, but the quality of the overall model must be taken into account. This applies particularly to intersecting building components, redundant or contradictory information.

---

1 https://www.gsa.gov/real-estate/design-construction/3d4d-building-information-modeling/bim-guides

Besides the above mentioned factors that shine through in the international initiatives, there are also other factors to take into account when creating a new system for BIM Based code compliance checking:

- Short term and long term potential
- Potential Productivity boost
- Technical requirements and risks
- System Maintainability
- Investments needed
- Caveats (how users could cheat)

2.5 Conclusion and recommendations

Most international initiatives focus on checking of IFC data in a descriptive way.
In many cases a pre-processing of the BIM model is therefore required, as information needed for checking some regulations (e.g. excavation routes) are not
directly provided by the dataset and must be computed or derived beforehand.

This leads to a situation where the authorities might have a productivity gain, but only when the industry delivers a dataset that needs to be specifically made for that single purpose of code compliance checking.

Currently the open data standard IFC is the primary input for a BIM Based permit checking system. The IFC data will be checked on an algorithm based solution. This is possible in a scalable way because the building decree is (generally) the same every time, on every location. A check against the zoning requirements (also called 'design requirements') is different every time. These checks should be based on a rule based engine. Prerequisite of the rule based zoning check is that the design requirements for every location need to be available in a a machine readable form. At the moment LandXML and CityGML don't seem to be focussed on this use-case.

Algorithms based solutions are currently more powerful as they can make direct use of internal algorithms of the code checking system. To provide accountability and transparency, the algorithms should be published. To make sure the system is available to a broad number of users, it should be based on open standards.

Singapore is searching for solutions to decrease the burden of creating specific data just for a compliance check. The examples of Korea and The Netherlands follow the principle of checking based on an IFC dataset that is already available in the regular process of the AEC industry. Checks can be done on multiple moments during the design and engineering process, to increase the predictability of the outcome of the check. A side effect of the chosen approach is the higher rate of validity of the checks because false positives are eliminated.

Ultimately, the decision for a specific solution or technology approach will be based on the situation and the goal of the initiative. In Estonia, there is a clear goal to increase productivity of the AEC industry, and increased predictability of the outcome of the checks.

Therefore we recommend to create several online tools (microservices) that can be triggered from one (or potentially multiple) entry point. The data that is being sent to the system should be IFC, the output BCF.

The microservices are an implementation strategy, because of the overall system needs a combination of multiple technologies to check several requirements from the building decree.
This makes it difficult to list the technical requirements and risks, because they are different per check.
Both the short term and long term potential look bright with this approach.
The system Maintainability is relatively easy because of the centralized nature of the online solution.
The approach means the white box rule checking issues that give users caveats to generate false positives are not an issue.

3 Part C: Conclusions and recommendations

In this chapter the conclusions and recommendations on data standards, process implementation and technical approach are described as requested in the tender specifications.

3.1 Data standards

To provide an e-construction platform with the right information, several types of 3D information need to be stored. For these 3D information types different open standards are applicable. It is important to conform to open standards, in this way you are not dependent on specific software vendors. Below we present an overview of the most relevant open standards. This is based upon our experiences and accumulated knowledge in relevant comparable projects.

First of all for BIM designs of buildings the most obvious format is IFC, this is an ISO standard and used worldwide. From the Quickscan and interviews we derived that within Estonia the acceptance grade regarding IFC is quite high and is used a lot. The concept of using object libraries to agree on the specific description of objects within designs, is used a lot as well. There are no agreements on a national object library yet, although the ideas to connect to the CoClass initiative are there. Since companies within Estonia are already used to work together in a shared object library, the step to connect to a National object library should not be a big challenge from a technical point of view.

Besides the building designs there are civil designs as well, in particular roads and areas etc. Where for building designs IFC is worldwide accepted as the open standard, standards for civil designs are on a much lower level of maturity. Another standard is LandXML, but this is not used in practise that much. The latest developments on IFC for infra by buildingSMART seem to be more promising for the future. It might be a bit too early to make the final decision on this standard, but for the proof of concept, we recommend the use of IFC.

Within the future E-construction platform the 3D Digital Twin of Estonia has an important role to play. The 3D Digital Twin could be used to place designs in their surroundings and check all kind of restrictions in the area like zoning plans, existing underground infrastructure or noise barriers. For the description of the existing build environment in 3D, CityGML is the best open standard to apply. Although CityGML is still less mature as IFC since it's over ten years younger, it's getting adopted by more and more cities and countries worldwide. Currently most of the additional 3D information like underground infrastructure and noise barriers does not fit into the basic CityGML model yet. There are several additional 'Application Domain Extensions' (ADE's) available for CityGML, to offer a standardized data model for this extra information. A list of these extensions can be found at: http://www.citygmlwiki.org/index.php/CityGML-ADEs. Most of these extensions are not yet widely used, but are still in a research phase. To really adopt such an extension, close collaboration should be sought with other cities or countries who are already working on this. For most of the ADE's these frontrunners are known and will be happy with support and ideas from new users.

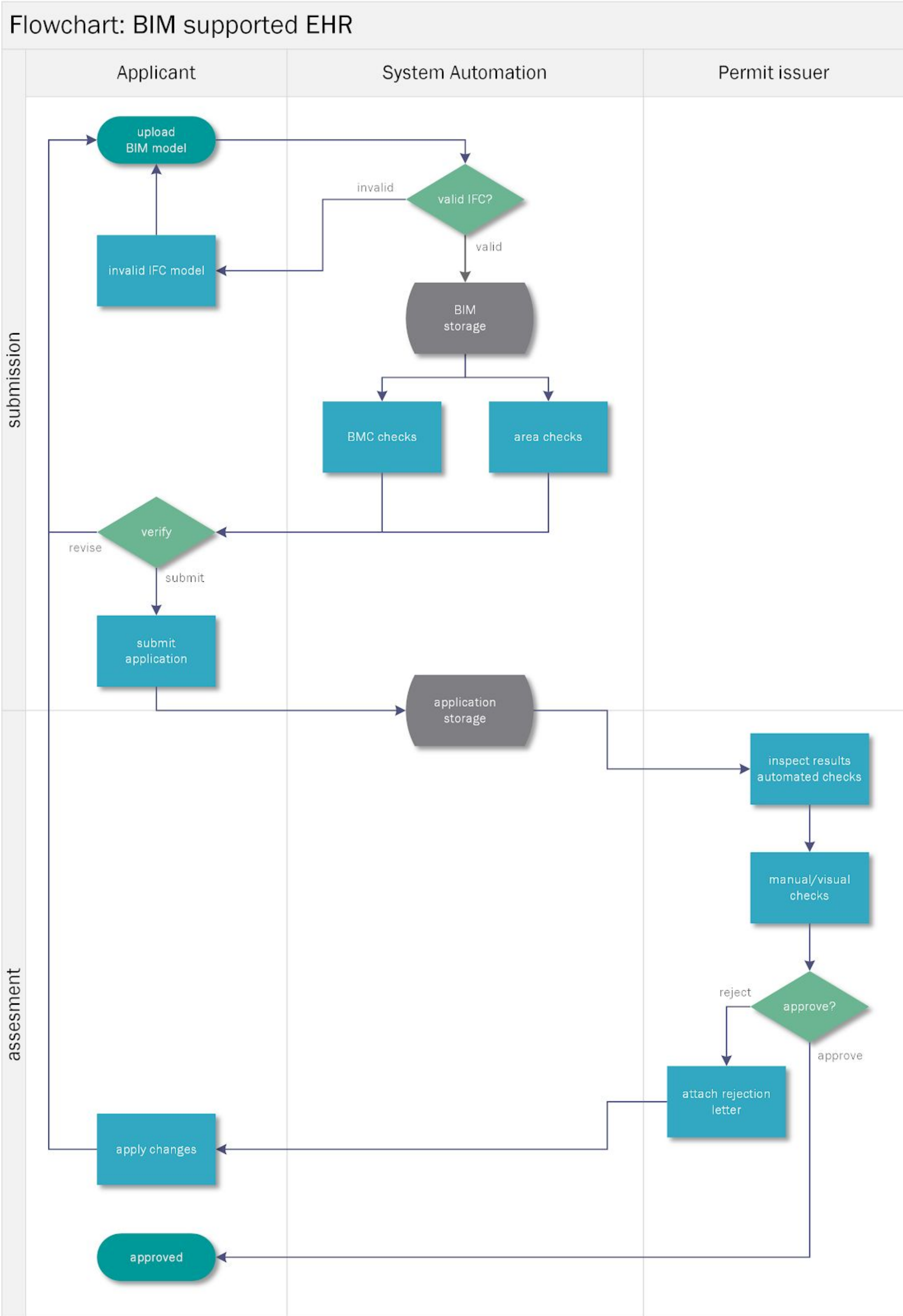
3.2 Process implementation

From the tender, a to-be business and technical process for the 'building' and 'use and occupancy' permit are requested. It should connect to the earlier performed research by PWC about optimizing the EHR and the draft of the new user-interface. Since this project specifically handles the BIM-based model checking opportunities we will specify a generic way the technology could be positioned within the EHR ecosystem. Depending on the process, different automatic checks should be prepared, which together should perform all the necessary checks for specific permit. The more detailed technical description for the BIM-based model checking are described in paragraph 3.3 Technical approach.

Below the description of the components corresponding to the process scheme on the next page.

1. **upload BIM model:** Depending on the permit type the applicant is requested to upload his BIM model via EHR.
2. **valid IFC?:** The BIM data are checked against the schema whether it's a valid IFC 2x3 or IFC 4 data format.
3. **BIM storage:** The IFC dataset is stored in a server
4. **invalid IFC model:** If the uploaded dataset is not a valid IFC format, or does not match one of the supported versions, the applicant will be sent an email stating that their model didn't pass the first checks, and needs to be fixed and resubmitted.
5. **BMC checks**: If the model did pass the schema validation, the process continues with the BIM based model checks. In this stage the predefined checks are executed on the BIM model and the results are stored. The specific checks that will be executed are depending on the type of permit requested.
6. **area checks:** The area checks are computed. Area checks are related to the living environment like noise and zoning plans. The BIM model will be placed in his environment and the checks are executed. Just like with the BMC checks, the checks that need to be executed will be based on the requested permit type. For certain checks, like a visual check against the zoning plan, an up to date 3D Digital Twin is required.
7. **verify:** All check results are made available through the verification interface. The applicant can consult the results and zoom in to specific issues or remarks. Additionally, the verification interface can provide information about what checks will be manually performed by the permit issuer once the application has been submitted. Thereby the verification interface provides the applicant a transparent overview on all permit checks that are executed. The more checks are available here, the better the applicant will know in advance if his design meets the requirements.
8. **submit application**: When the applicant is ready to formally submit his application they can do this via the submit button. They can add additional information prior to submission or if the applicant want to adjust his BIM model he can do this as well. The process will proceed to the first process step where the applicant can upload his revised BIM model.
9. **application storage:** Any additional information the applicant has left, will be stored. And the application is now released for manual review by public authority.

10. **Inspect results automated checks:** The permit issuer receives the application and scans over the results of the automated checks.
11. **manual/visual checks**: For now it will still be required to have manual or visual checks complementary to the automated checks. The manual checks should be described as objectively as possible. The next step could be to automate these manual checks eventually resulting in an automated system.
12. **approve**? Based on the results the issuer decides if the application is approved.
13. **attach rejection letter**: When the permit issuer decides to reject the application he is required to provide an additional explanation by including a rejection letter. The rejection letter will be send to the applicant, explaining why the application has been rejected, in such a way that the applicant can re-submit with minimum effort.
14. **apply changes**: The applicant receives the notification that he needs to apply changes in his application in order to get an approval.
15. **approved**: The official permit is send to the applicant.



Flowchart: BIM supported EHR
Applicant
System Automation
Permit issuer
submission
assesment
upload BIM model
invalid
valid IFC?
valid
invalid IFC model
BIM storage
BMC checks
area checks
verify
revise
submit
submit application
application storage
inspect results automated checks
manual/visual checks
reject
approve?
approve
attach rejection letter
apply changes
approved

3.3 Technical approach

Based on the results from the Quickscan and the interviews the market in Estonia seems to be ready for BIM Based code compliance checking.

In many international examples, a pre-processing of the BIM model is therefore required, as information needed for checking some regulations (e.g. excavation routes) are not directly provided by the dataset and must be computed or derived beforehand. This leads to a situation where the authorities might have a productivity gain, but only when the industry delivers a dataset that needs to be specifically made for that single purpose of code compliance checking.

Black box solutions are currently more powerful as they can make direct use of internal algorithms and data structures of the code checking system. Singapore is searching for solutions to decrease the burden of creating specific data just for a compliance check. The examples of Korea and The Netherlands follow the principle of checking based on an IFC dataset that is already available in the regular process of the AEC industry. Checks can be done on multiple moments during the design and engineering process, to increase the predictability of the outcome of the check. A side effect of the chosen approach is the higher rate of validity of the checks because false positives are eliminated.

Ultimately, the decision for a specific solution or technology approach will be based on the situation and the goal of the initiative. In Estonia, there is a clear goal to increase productivity of the AEC industry, and increased predictability of the outcome of the checks.

Therefore we recommend to create several online tools (microservices) that can be triggered from one (or potentially multiple) entry point. The data that is being sent to the system should be IFC, the output BCF.

Besides the BMC checks 'area' checks can be distinguished. These are checks where environmental factors relevant to the permit can be checked using the 3D Digital Twin. Off course these environmental factors need to be known in the Digital Twin first before they can be checked. In the short term, a facility has to be established to place the designs in the right place on the map, since usually the location is not available in the BIM design. On the longer term agreements should be made on how to store location correctly to make this step unnecessary.

To effectively perform the area checks it is probably necessary to convert the BIM designs to CityGML LOD1 or LOD2. There are several solutions available who can perform a first version of such a conversion available. Also the solution needs to hold the IFC data in a database, with the option to serialize/publish the data as glTF in a Cesium environment. This makes it possible in the future to view the original submission (not the simplified converted CityGML) in the digital twin environment of Estonia.

The microservices are an implementation strategy, because of the overall system needs a combination of multiple technologies to check several requirements from the building decree.
This makes it difficult to list the technical requirements and risks, because they are different per check.
Both the short term and long term potential look bright with this approach.
The system maintainability is relatively easy because of the centralized nature of the online solution.
The approach means the white box rule checking issues that give users caveats to generate false positives are not an issue.

3.3.1 Proposed solution

The proposed solution is to create an online web service for users to upload IFC data. This online service will perform a Code compliance check and return results to the user. The solution is strongly relying on buildingSMART open standards for data (IFC), results (BCF) and API (openCDE API)

The whole solution will work online using an IFC model server, a webGL based viewer and a React Javascript based user interface. No additional plugins need to be installed on the computer of the user, to effectively use the proposed system.

There are multiple ways to interact with the final system:

- upload IFC via a graphical user interface
- upload IFC through a plugin in an authoring tool (Revit, Tekla, ArchiCAD, etc)
- Upload IFC through an open API (for example the openCDE API)

This gives users the option to constantly evaluate the data during design and engineering. Every time a significant change is detected by the authoring tool the online code compliance check web service is triggered and evaluates the data and the design. Results will be presented to the users in the BIM authoring tool. This will create immediate feedback to users on code compliance of their design in a very early stage, creating efficiency and productivity for the industry. The triggering of the online service will be based on the BIM Bots concept (www.bimbots.org) This technology will aid in the model validation process so that issues in the model are detected and addressed before the actual compliance check at the authority.

All data will be stored on the online system so relevant authorities are able to consult the design (online) in an early stage and possibly provide wavers.

3.3.2 Proof of concept

The source code of the solution will be available as open source and the online service will be available to users in Estonia for at least x years.

During the preamble of this project, there was a starting point of two proof of concept tools. Since the basic technological challenges of code compliance checking are the same for any purpose, we recommend to create óne tool that functions as the proof of concept for multiple checks.

We intent to use proven robust and scalable technologies for this system. The open source BIMserver.org platform will be the core to store and check the IFC data online. Related products like IfcOpenShell and BIM Surfer will handle the geometry part and visualization.

The code compliance rules will be written in BIMserver in a customizable way, although we will not focus on the configuration of the rules during the proof of concept.

The geometry engine IfcOpenShell has a strong relation with BIM server and is used for any geometric analyses. The online viewer will also use IfcOpenShell which is an open source IFC render engine. The accuracy and integrity of the combination BIMserver and IfcOpenShell will ensure that all errors are correctly identified and results will be reliable and reflect the true errors and compliances in the data. The viewer will be based on the newest BIM Surfer (the V3 branch).

The BIMserver solution has the option to publish IFC as glTF in a Cesium environment. This makes it possible in the future to view the IFC buildings in the digital twin environment of Estonia.
The graphical user interface that users use to interact with the system will be based on the React framework to fit into the ecosystem of Estonian tools.

The system will be able to handle valid IFC2x3 and/or IFC4 data.
We will provide hosting for system for x years, and make it available for free to Estonian users.

The open source license of the tools we use requires this work to be open source as well. The license provides no obligation to publish the source code online, but we will provide the source code for free with this project.

3.3.3 Types of checks

In the review of BIM based code compliance checks from other countries, a few different categories of checks can be derived:

Type	Example of the check from the building decree	BIM Characteristics
Analyses	Accessibility: "Minimum headroom of 2m along circulation routes and minimum corridor clear width of 1.5m for barrier free access route"	Info not in a typical IFC/BIM model; check cannot be done with rule checking
Entity based checks	Changes in level: "The barrier free access route must not have any sudden change in level. Ramps that mitigate level changes has to comply with a table"	Not following Modelling guidelines will create a false positive result. (Using IfcProxyElement instead of IfcRamp)
Multiple entities based check	Ventilation Ratio: The size of the window on the exterior facade of the building should be at least 5% of the size of the room space/ space it is ventilating	The model needs to have correct semantic relations between multiple objects in the datasets
Simulation	Safety barriers: "Where there is a drop of more than 1m, a safety barrier with a minimum height of 1.0m should be provided"	Info not in a typical IFC/BIM model; check cannot be done with rule checking; integrated approach (agents) is needed.
Property based	"For every 10 car park lots provided, there should be at least 1 accessible	This is forcing the industry to add

checks	car park lot provided for wheelchair users"	properties to IfcSpace objects to pass the check, while they would normally not add this information this way.

This overview confirms, and shows the added value for an algorithm based approach.

We propose to implement the following checks in the proof of concept, with the technology on which the algorithm will be based to check the data:

Check	Algorithm
"Where there is a drop of more than 1m, a safety barrier with a minimum height of 1.0m should be provided"	Using Voxel technology to analyse the geometry independent from the IFC Type that is used.
"There should not be cracks and holes in the building where vermin can enter"	Using advanced algorithms to check the density of the geometry.
"Ramps that mitigate level changes has to comply with a table"	Checking *'all'* objects to see if the walkable surface complies with the preset requirements
"Doors should have a minimum height"	Basic check on property of an IfcDoor object. Optional: Check if the geometry is consistent with the property.
"All objects should have a material"	Check if all objects have a material (either IfcMaterial or Pset, or optional a material in IfcName attribute)

This list of different types of checks and solutions will show the potential of the proposed solution, and highlight the differences between a rule based check and an algorithms based check.

3.4 Requirements for use of the Prototype

To stay in the concept of productivity enhancement, the proposed solution requires as little as possible from the end users. There are some requirements to the system of the users though:

- The computer of a user needs to support webGL;
- We only test the system in chrome/firefox;
- The connection of the user needs to support data streams through a websocket;
- We recommend the use of a large screen (high number of pixels) for a good user experience;

As stated many times in this document, the system will focus on using IFC data that is generated by current practitioners. We found the level of IFC quality in Estonia quite high and feel confident that we can assume the following data will be available in the IFC datasets:

- Valid IFC
- IfcSpaces have to have a closed 3D geometry

This list might grow during the development of the prototype.

3.5 Reflection

There are some things to take into consideration when adopting this approach.
The building decree might be updated and rules might change. The black box approach means that the algorithms need to be adjusted as well. In a rule based approach this would be easier (simply edit the rule), but in the algorithm based approach a programmer needs to change the code or algorithm. When the system needs to support older versions of the building decree, version management of the software needs to be in place.

Although a black box approach based on algorithms does not make that much demands to the IFC data structure, it can still only check on the data that are in the dataset. Missing parts can not be recognized and will not be checked. This is the same with rule based checking, but good to pay attention to.

Creating a government driven BIM Based checking tool might cause a disruption in the market of BIM checking tools. Some users that buy a BIM checking software tool, might not buy this anymore because the government has one in place. This might lead to unwanted situations.

During the development we will use the IFC datasets we obtained from the interviews to test the system. We assume this is allowed. The data will not be published on any public repository and we will adhere to the privacy guidelines.