# KEYSTONE PLAYERS IN COLLABORATIVE BUILDING INFORMATION MODELING — FORM OF CONTRIBUTION IN JAPANESE LARGE-SCALE PROJECTS

*Tsukasa Ishizawa, Ph.D.;*
*Institute of Industrial Science, The University of Tokyo*
*Computational Design Group, Advanced Design Department, Takenaka Corporation;*
*ishizawa@iis.u-tokyo.ac.jp*

*SUMMARY: The stagnation of construction productivity is becoming increasingly serious in Japan with the decreasing construction workforce. Although BIM has attracted attention to overcome this problem, its adoption has not progressed among small organizations. Expanding the BIM use should be driven by the influence of large organizations. This paper stratifies users by cross-analysis using BIM log mining, a newly emerging analytics approach based on Autodesk Revit, combined with recorded software session times of other software to improve the shortcomings of the existing method. The target company, a Japanese general contractor, where external dispatched personnel accounted for most BIM activities, needed to recognize permanent employees who undertake the crucial role in promoting cooperative BIM projects termed the keystone BIM players. The machine learning-based clustering algorithm and visual analytics discovered a group of collaborative users whose intensity of software use was weaker than proficient users but who provided a substantial proportion of the team's workforce, including multiple applications. The semi-structured interviews as a verification process further clarified that they positively perceive collaboration with external BIM operators; while delegating most tasks, they strive to improve their own BIM knowledge to respect equal collaboration. The methodology provides an indispensable dashboard to improve the project BIM communication, which is the pivotal factor in influencing the further utilization of BIM in the whole industry. The contribution of the research is threefold; the extended BIM log mining technique, the discovery of keystone BIM players, and the exclusive focus on the cooperative relationship in the BIM project environment.*

*KEYWORDS: BIM log mining, visual analytics, semi-structured interview, keystone species, Japanese general contractors, large-scale projects*

# 1. INTRODUCTION

## 1.1. Foreseen problems in the Japanese AEC industry

The productivity growth in the construction industry internationally has been stagnant for decades (Teicholz, 2013). The index between 1990 to 2005 showed little growth in the states and France or even negative trends in Germany and Japan (Abdel-Wahab and Vogl, 2011). A decrease in construction workers has been happening much more rapidly than shrinking Japan's total population. Government statistics predict a shortfall of 470,000 to 930,000 construction workers by 2025 (Yamada and Isoyama, 2018). The aging workforce accounts for the principal factor. Thirty-five percent of the labor pool is over 55 years old in 2019, while less than 11 percent is under 30. The cohort has stayed older than the all-industry average, with the retiring population outnumbering the new entrants (Ministry of Land, Infrastructure, Transport and Tourism, 2017).

The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) has been striving to overcome the situation by reducing overtime, providing subsidies to employers, and reviewing the insurance system (Ministry of Land, Infrastructure, Transport and Tourism and Ministry of Health, Labor and Welfare, 2020). On top of these, Building Information Modeling (BIM) is imperative for improving overall productivity.

The role of large private enterprises is considered to play an extremely important role, particularly the general contractors providing design and build services, because they exert a large influence on industry trends with the potential as a hub of ordering and receiving relationships. The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) investigation in 2021 (Ministry of Land, Infrastructure, Transport and Tourism, 2021) depicted that 46.2% of the 813 target companies had already implemented BIM. Smaller organizations lagged in adoption; the implementation rate was reported as 87.8% for organizations larger than 5,000 employees, while it ranged from 20.0 to 36.7% for the firms smaller than 100 employees. 72.3% of BIM-ready firms answered about their motivation to anticipate future trends in the industry. Among the BIM companies that were encouraged by their competitors, 80.2% evaluated the effect of BIM positively. This tendency reflects the cooperative business environment in Japan.

General contractors represent the most typical of Japan's large private sector operators. They have long been known for providing inclusive and coordinative services in research, design, engineering, construction, and maintenance and management (Ku et al., 2008). Their actual BIM workflow is often considered an internal matter of one particular entity. Despite recognizing that challenges exist, scholarly contributions are scarce for overcoming them through data-based research and recommendations (Shide, 2015; Ishida et al., 2016; Yoshikai et al., 2018). Not many studies have been conducted on the nature and peculiarities of the Japanese general contractors themselves, with some attempting to clarify their practices during the bubble period of the late 1980s when they aggressively expanded into overseas markets (Bennett et al., 1987; Webster, 1993). In the 21st century, however, since they are no longer seen as an international threat, research on Japanese general contractors remains largely internalized by those involved in them (Ogasawara and Yashiro, 2018; Suzuki and Sui Pheng, 2019). Resolving issues of BIM workflow within Japanese general contractors is vital not only for general contractors but also to stimulate BIM use in the whole industry.

## 1.2. BIM log mining as an emerging field

While BIM maximizes its benefits when the designers use it, this may not be feasible in some cases. It is highly feasible when the project size is modest or the design period is sufficient. Conversely, large-scale or short-term projects inevitably require collaboration.

The BIM human resources employed are often external in Japanese general contractors. The project BIM Manager is often from the architecture design department, although having a full-time dedicated professional is not rare. In general, BIM technicians are underpaid compared to architects, but there appear to be reverse cases (Seow, 2014). Project-specific BIM team structures have not been fully studied.

BIM log mining is an emerging method to decipher the collaboration in BIM activities and extract useful knowledge to improve learning and management methods. This technique is applied to BIM-based process mining, which is widely used for web log analysis (Munk et al., 2010).

Autodesk Revit, one of the most common integrated environments worldwide, generates detailed session records called journals for troubleshooting purposes (Autodesk, 2017). Yarmohammadi et al. demonstrated that journal

files could be formatted to extract various insights (Yarmohammadi et al., 2016, 2017). The topic is followed by Pan et al. with the name BIM log mining, who proved that it could yield predictive information useful for project management combined with machine learning(Pan et al., 2020; Pan and Zhang, 2021). In particular, from the perspective of productivity, methods for contextual prediction of design work productivity (Pan and Zhang, 2020b) and methods for exploring typologies of design work have been proposed (Pan and Zhang, 2020a; Gao et al., 2021). Forcael et al. have published an attempt to measure the contributions of cross-disciplinary BIM users from BIM logs (Forcael et al., 2020).

One of the known challenges to BIM log mining is how the prerequisite productivity can be related to BIM activities -- the more active users are in BIM, the more they contribute; yet the reverse is not necessarily true. Another problem is the dependence of analysis environments on applications. Since the mining process relies on the log files produced by applications, it becomes more challenging to study with software that does not record log files to the same extent. One possible strategy is to mine common data formats such as IFC (Kouhestani and Nik-Bakht, 2020). However, such information, including editing processes, undone tasks, and settings necessary for collaboration, cannot be retrieved from model files alone.

## 1.3. Discovering keystone species in BIM ecosystem

BIM specialists are often appointed separately from the architects, such as BIM modelers and technicians. Prior research on BIM staffing focused on identifying the type of person required based on job descriptions (Barison and Santos, 2011; Mathews, 2015; Uhm et al., 2017).

In Japan, licensed architects usually hold responsible positions for client meetings, governmental consultations, et Cetra. The abovementioned government survey revealed that BIM practices are carried out by licensed architects and by supplementary personnel without licenses, particularly in the mechanical and electrical disciplines(Ministry of Land, Infrastructure, Transport and Tourism, 2021).

The project will not progress with only BIM operators in place. Their tasks need to be supervised and approved by qualified designers. Experienced architects may prefer to review the model in drawings to exercise their expertise. However, it only remains a partial optimization to check the model's progress through drawings. Decisions about the effectiveness of model work and information build-ups in the model are also necessary, which are deeply related to the project requirements. Non-BIM designers should also possess adequate BIM knowledge to facilitate this communication. Many practitioners' skills and knowledge of BIM generally fall behind due to the educational costs and the lack of awareness (Kaneta et al., 2017). Meanwhile, there exist some successful projects utilizing BIM. Social network research of BIM projects showed that only a limited number of personnel are engaged in modeling (Zhang and Ashuri, 2018).

Interpreting the concept of keystone species from its origin and applying it to the broader context of the AEC ecosystem is helpful to analogically understanding the situation. A keystone species is an entity that has a disproportionately large influence on an ecosystem compared to its abundance. Similar phenomena can be witnessed in the environment like the internet and software systems in the natural world. This concept has been applied in non-ecology fields like internet services (Ejima et al., 2019; Hileman et al., 2020). Ejima et al. define the keystone species as a set of species that significantly impacts the ecosystem if removed from the system, irrespective of its small biomass. From the perspective of social activities, the stability of ecosystems is realized by keystone species' high level of activity and engagement; thus, identifying them may work to increase the community's activity. An attempt to identify key players in BIM projects who silently yield a great positive influence to project BIM progress can be interpreted as the method to discover the keystone species in an organization's BIM ecosystem, which can be termed as keystone BIM players.

## 1.4. Aim of research and the expected contribution

BIM training programs mostly emphasize modeling work on the presumption that architects do most modeling operations by themselves. In actuality, however, the task of producing digital architectural data has been widely outsourced to the dedicated operator as highly specialized work since the era of 2D-CAD (Nakamura and Ito, 2018). Even though BIM can streamline the building and construction process, it is not automatically evident that BIM enables federate tasks formerly outsourced into the in-house workflow.

BIM software is a conglomerate of existing modeling, viewing, rendering, engineering, coordinating, and data management applications. Obtaining fluent BIM software skills requires significantly more time, thus the

acquisition cost is considered high (Park and Lee, 2010; Sharag-Eldin and Nawari, 2012). Eventually, many practitioners find it nearly impossible to master the ongoing project tasks. It is unrealistic to expect everyone in the construction industry to become a BIM user capable of modeling. On the other hand, insufficient BIM knowledge poses a significant risk for project collaboration and management (Chien et al., 2014). The practitioners should acquire a practical BIM skill balanced between creation and collaboration.

BIM personnel acting as the keystone species in the project environment is considered to lead the large-scale projects to success. Suppose such BIM talent in projects and organizations can be identified from a broad perspective, including non-Revit users. In that case, optimization of BIM-related investment allocation becomes possible and better collaboration at the project level.

This paper aims to recognize the keystone species among BIM-enabled staff in collaborative projects. The challenge lies in detecting the characteristics of those who play a promoting role in promoting BIM projects from actual data. To answer this question, the authors employ BIM log mining as a data-driven approach to analyze how the staff, regardless of their discipline and affiliation, contribute to the project BIM in a large-scale BIM project promoted by a Japanese general contractor.

## 2. METHODOLOGY

### 2.1. Overview

This article is a part of the author's doctoral thesis research. The research flow is based on the standard design science research method (DSRM). The framework for this paper is presented in Fig. 1.

The structure of this paper is as follows. In this section, we explain the data acquisition and analysis methods. A major corporation that suits the research background is invited to collect log files and relevant datasets for a limited period. The corporation also consented to use the collected data for the research purpose. The method reported by Ishizawa et al. (Ishizawa and Ikeda, 2021) is employed for BIM log mining in this paper, reducing the dimensionality by clustering with machine learning.

In Section 4, visual analytics techniques enable us to discover patterns empirically. Command-based, organization-based, and project-based analyses interpret the clustering results heuristically. Because the existing methods are biased towards Revit content, cross-analysis was conducted, including each user's attribute information and the usage time data of software other than Revit. To test the patterns obtained, the authors conducted semi-structured interviews to understand the actual state of BIM use. the findings are summarized per topic, and the commonly shared responses are explored. The subsequent sections provide the discussion, limitations of this paper, and conclusions.
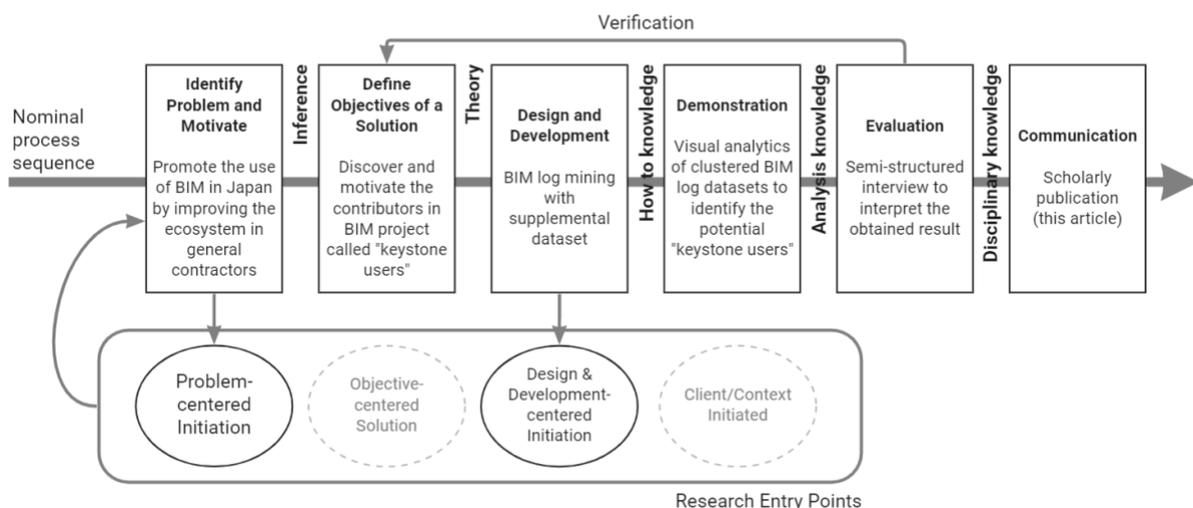


*FIG. 1: Research workflow following Design Science Research Method (DSRM)*

## 2.2. Data collection

### 2.2.1. BIM log

A major Japanese general contractor and its subsidiary company have supplied the BIM log data for the subject dataset. This company is one of the largest general contractors in Japan that handles many design-build projects. The affiliated employees are either Permanent staff, as lifetime employees mostly hired right after graduation, or short-term temporary ones (External). In many established Japanese companies, the hiring, employment, and evaluation systems differ between Permanent and External specialist. External members, either dispatched from temporary staff service or working as a freelance, are appointed on-demand and seldom become the Permanent. Some functionality like CAD operators have been largely outsourced to External members to control the design and development expense; this tendency also applies to BIM. The exclusive focus on the contribution per affiliation is immense to interpret the BIM ability and capability at the organization level. The typical deployment in design-build project is illustrated in Fig. 2.
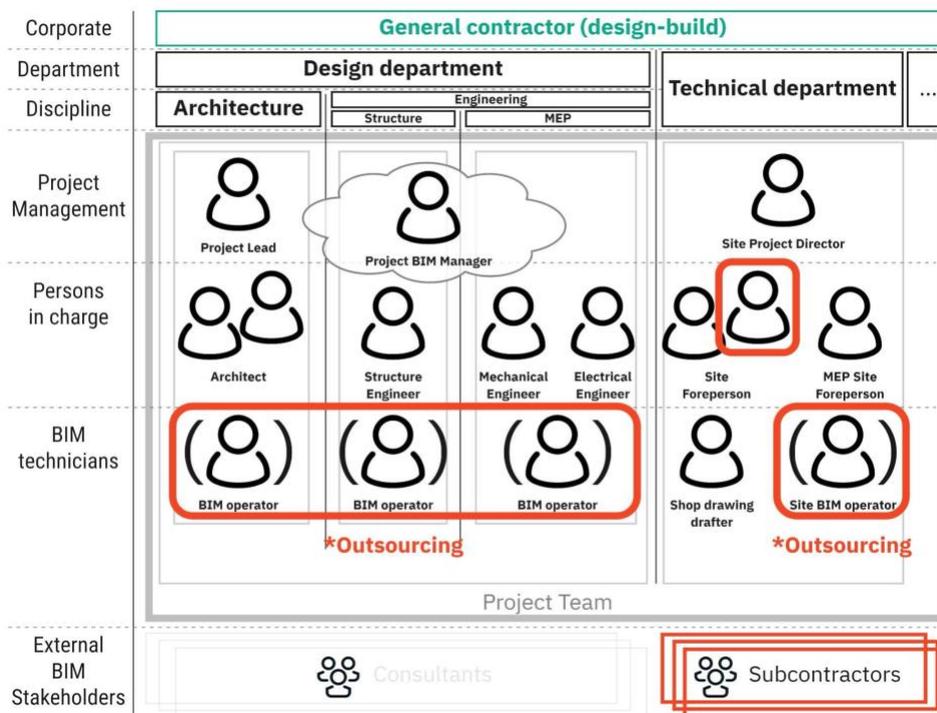


*FIG. 2. The deployment of External staff in schematic project collaboration in a design-build general contractor.*

The data was collected from Permanent and External staff working at the Tokyo office of the firm above and employees engaged in BIM-related work at its subsidiary company that provides drafting services. In traditional Japanese companies, the hiring, employment, and evaluation systems differ greatly between Permanent and External staff. To date, functions like CAD operators have been mainly outsourced to External personnel. This tendency also applies to BIM. Therefore, we differentiated between the two by using the attribute "affiliation" to indicate the type of staff (Permanent or External personnel).

The author extracted employees who had used Revit from the software activity records. Only those who agreed to the data collection method and use were eligible for the data collection. To minimize the burden of data transmission and prevent missing data, we installed a batch program for data transmission remotely and automatically acquired the log files. The program automatically stopped after the scheduled data capture period. From 238 targeted users, 182 persons had provided the data. The rest either declined to cooperate or failed to transmit their logs due to telecommunication problems. The overview of the collected dataset is as listed in Table 1. Separately, the demographic information including the data donor affiliation and department is detailed in Table 2.

*TABLE 1. Overview of the collected datasets.*

|  | Dataset |
| --- | --- |
| Number of data donors | 182 |
| Number of collected log files | 8,432 |
| Issued commands (effective) | 565,289 |
| Issued commands (overall) | 1,113,977 |
| Average issued commands per log | 132 |
| Collection Period | April 2019 – August 2019 |
| Method | Automatically transferred |

*TABLE 2. Number of data donors per department and affiliation*

|  |  | Permanent staff | External staff |
| --- | --- | --- | --- |
| Design department | Architectural design | 14 | 17 |
|  | Structure engineering | 15 | 7 |
|  | MEP engineering | 8 | 2 |
|  | Other (computational design, BIM promotion) | 9 | 2 |
| Construction department | Production detail design | 11 | 9 |
|  | Quantity survey | 1 | 0 |
|  | Construction | 14 | 27 |
|  | Construction, MEP | 8 | 0 |
| Administration |  | 4 | 4 |
| (Subsidiary company) |  | 0 | 30 |
| Total |  | 84 | 98 |

### 2.2.2. Aggregated session time per software

Large-scale BIM project collaborations usually require the combined use of software besides Revit. For example, ARCHICAD is used in Japan to a similar extent as Revit. Rebro, a software product developed in Japan, is extensively used for mechanical and electrical design. Also, managers usually prefer to use integrated reviewing software, including Solibri and Navisworks, for the audit and approval process.

These are aspects impossible to decipher from the Revit logs alone. These applications do not provide as abundant records of the process as Revit does. Therefore, we sought to grasp the intensity of software usage by counting the software running time per user.

For the above users, the session times of the representative BIM-related applications were aggregated for joint analysis with the BIM log. The software included BIM modeling software (Revit, ARCHICAD, Rebro), management software (Solibri and Navisworks), and drafting software (AutoCAD). The running duration of those applications is automatically recorded separately from the log collection process. The time summations per user from 1 January 2019 to 28 February 2019 are adopted for the analysis. The accumulated time may indicate longer than the actual working time because the different versions of each software may have been run simultaneously by an identical user.

### 2.3. Algorithm

The classification aimed to separate the log files based on the types and frequency of issued commands during a single software session. The clustering process requires the structured aggregation of the issued command types per user. The authors tailored Python scripts to parse the unstructured text in the collected log files. The results were accumulated in a Comma Separated Value (CSV) format for the subsequent analysis.

First, the Python code reads lines from log files to extract the unique IDs of issued commands triggered by the tag "Jrn.Command". The extracted IDs were added into a single CSV file as a command sequence per log. Second, the invalid command "ID_CANCEL_EDITOR" was omitted from the sequence since this is issued whenever the user presses the escape key. Third, the CSV files per log are further aggregated into one table. Each line represents a log with the summed counts per command within a software session. Logs are named to distinguish the users and organizations. The clustering algorithm runs on this dataset to partition the logs according to their commonalities in command counts. The algorithm up to this process is outlined in Table 3.

*Table 3. Algorithm for data collection and cleansing.*

```
Algorithm 1: The data cleansing and summation
Input: log files j embedding username Un and sequential number Jn
Output: output table in csv format t, df
```

```
 1: FOR all files in j
 2:   FOR all lines in j(Un, Jn) Do
 3:     IF "Jrn.Command" is in line THEN
 4:       extract Command_ID, Command_Timestamp from line
 5:       write CommandID, Command_Timestamp into t(Un, Jn)
 6:     ENDIF
 7:   ENDFOR
 8: ENDFOR
 9: FOR all files in t
10:   FOR all lines in t(Un, Jn)
11:     Count commands per type
12:     Write username, counts(command_type) into df
13:   ENDFOR
14: ENDFOR
```

Model-based clustering is a popular technique relying on finite mixture models that proved efficient in modeling heterogeneity in data (Celebi, 2015). The algorithms are deployed to the tailored Python script by Scikit-learn, a Python module that integrates many state-of-the-art machine learning algorithms. Among the coverage of the package, K-means (MacQueen, 1967) and Variational Bayesian Gaussian Mixture Model (VBGMM) (Corduneanu and Bishop, 2001) are the two plausible estimators for the purpose.

Upon comparing the silhouette coefficient values for partitioning the dataset into 20 clusters, VBGMM was selected as the best estimator among these algorithms. This choice is consistent with the fact that K-means implicitly assumes hyperspherical clusters in shape and numbers of objects in clusters are equal; thus, it is challenging for the K-means algorithm to extract structures from the data that violate this assumption.

The number of clusters to partition is non-deterministic. While the silhouette coefficient is applied to experiment with the different cluster numbers, no prominent tendency was observed, considering that the randomness of the initial values affects the clustering results. Too few clusters seemed not to yield the desirably discriminating results for a broad range of data sources. Therefore, we hypothetically employed 20 clusters and attempted to interpret the results through visual analytics by incorporating the user survey. The clustering algorithm jointly executed with the earlier procedure is expressed in Table 4.

*Table 4. Algorithm for data collection and cleansing.*

```
Algorithm 2: Clustering by VBGMM
Input: CSV file df, Cluster number k empirically determined by Silhouette Coefficient
Output: Expected label in VBGM El
```

```
1: Convert df into matrix
2: Initialize parameters k
3: Perform BayesianGaussiannMixture on df
4: Get El for each prediction in matrix
5: Append El for each line in aggregation.csv
```

## 2.4. Clustering Result

When we developed a clustering method in a previous paper (Ishizawa and Ikeda, 2021), we classified the clusters into four superordinate groups based on their similarity among each other: Void, Dominant, Major, and Minor. Though we follow that result in this paper, we also examined the number of included commands per log file belonging to each cluster for more accurate classification by project contribution. As shown at the bottom of Fig. 3, there existed a significant difference in the average number of commands between Dominant and Major Clusters. It implies that the gap in contribution to the project is not negligible. From this aspect, #3, previously classified as a Major cluster, was recategorized as a dominant cluster. The updated overview of four cluster groups plotted on the size and nominal contribution to the model is illustrated in Fig. 4. Hereafter, we denote the log files classified into each cluster as Void, Dominant, Major, and Minor logs.
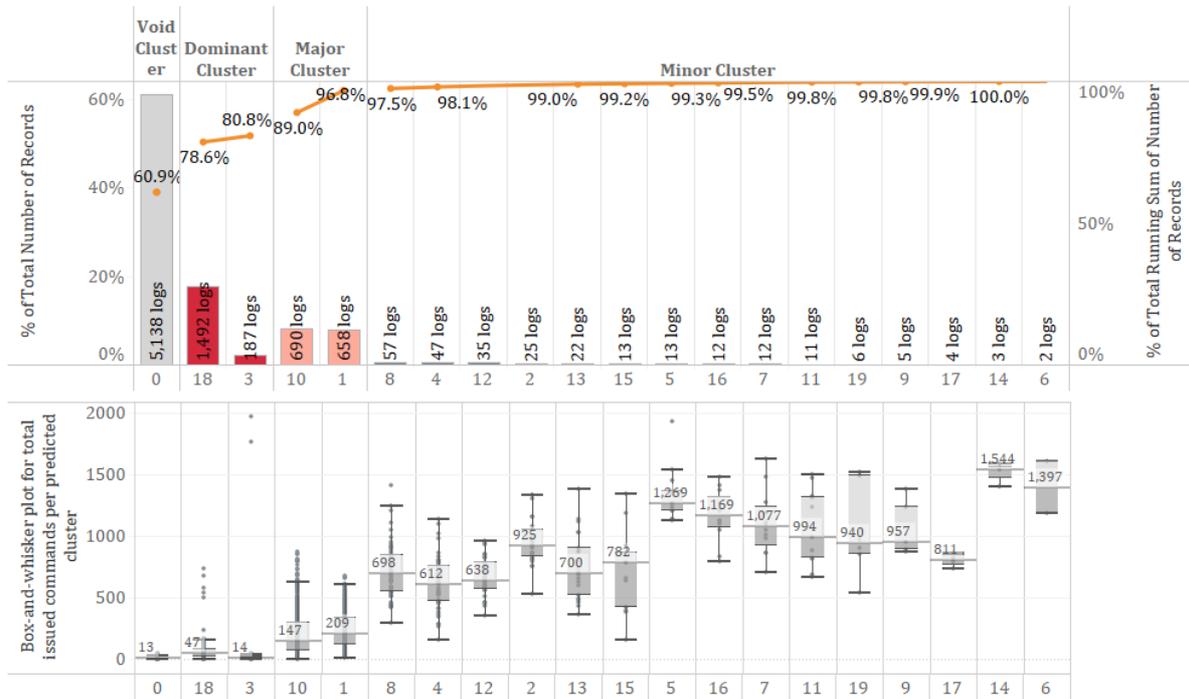


FIG. 3: Overview of predicted clusters and the numbers of total issued commands per log under respective clusters, with four super-cluster types
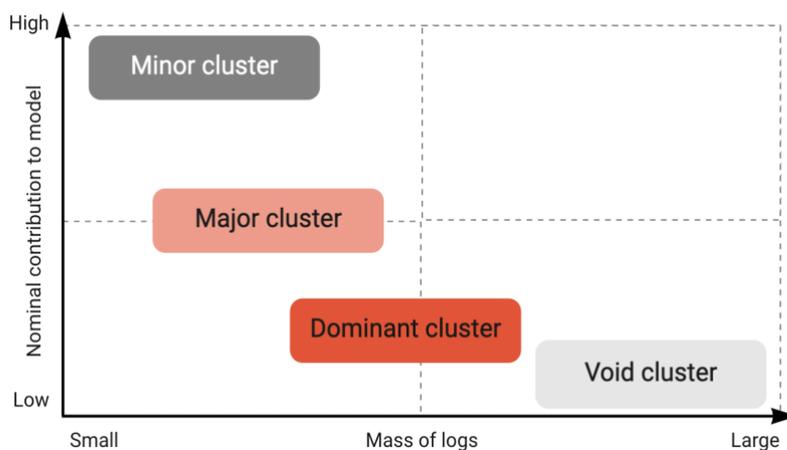


FIG. 4: The implication of four cluster groups plotted on the size and nominal contribution to the model

# 3. ANALYSIS

## 3.1.BIM log visual analytics

Visual analytics enables us to discover patterns by projecting data on various subspaces. It is more likely to yield significant findings than methods like principal component analysis when the data comprises overwhelming dimensions compared to the data volume, as in the case of BIM logs (Andrienko et al., 2020).

Fig. 5 shows the number of log files collected and the number of commands executed in them by affiliation. While the number of data providers for both Permanent staff and External staff does not differ substantially, the total amount of logs collected from the External staff was 1.80 times larger than Permanent staff, and the overall count of commands issued was 3.04 times larger from External staff. The BIM workload per person is considerably higher for External personnel.

Within each affiliation, however, the BIM contributions of users are not evenly distributed. One user accounted for more than half of the number of commands executed among 84 employee users. Similarly, specific External employees undertook significantly more work. The top seven of all data contributors executed more than half of the commands issued.

As seen above, the degree of the project BIM contribution is not merely determined by affiliation. In Japan, cooperative relationships are generally established without being restricted by the definition of job responsibilities, and thus the reality of work in projects is considered case-by-case (Buntrock, 2002; Ogasawara and Yashiro, 2017).
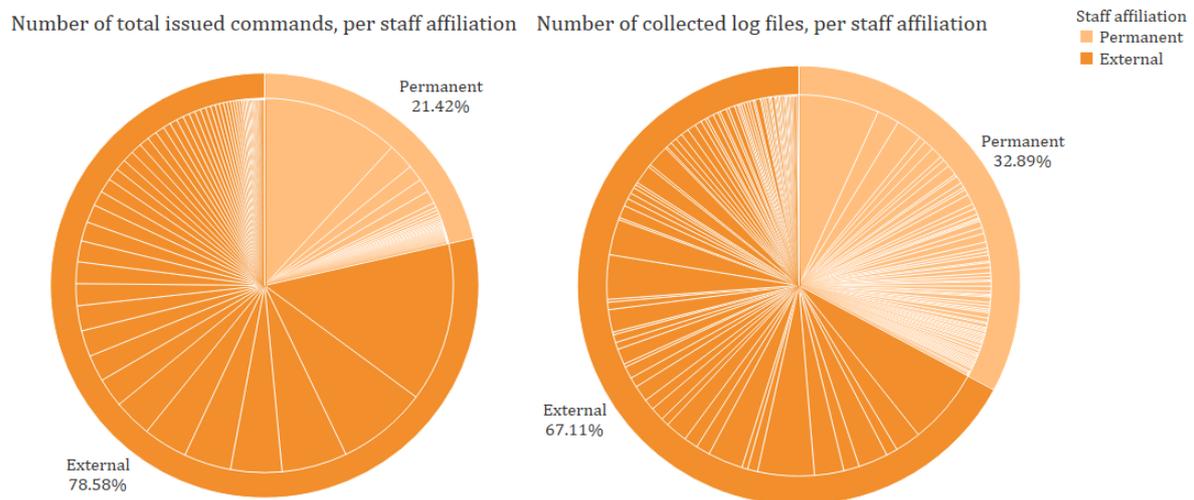


*FIG. 5: Number of log files and issued commands per staff affiliation.*

Clustering results of the logs per individual user are presented in Table 5 to comprehend the users' BIM activities during the data collection period. Logs under Major and Minor Clusters contain a higher average number of commands and thus have a higher nominal contribution. In contrast, the logs in Dominant and Void Clusters have very few commands in a session; particularly, the Void logs generally exhibit an insufficient amount of execution for significant modeling work.

In light of the factors above, we experimentally classify users into four cluster groups. Proficient user group (G1) refers to the users who nominally contribute the most to the task with the Minor log, Collaborative group (G2) means users who do not hold the Minor log but bind to the Major log, Contributing (G3) is for the users who hold the Dominant log, and General (G4) represents users who only carry the Void log. All users except User 086 own the Void log. All Proficient users have Void, Dominant, and Minor logs, but few are without Major logs (036, 039). A small number of users do not have dominant clusters, whereas all Collaborative users have Major logs.

In the following, we will focus on these four user groups and their affiliations to examine the meaning of this classification.

Fig. 6 shows an aggregated ratio of the command types for the respective user groups. categorized into seven types: Modeling commands and Drawing commands for direct contribution to the modeling progress, Editing commands and Save commands for indirect contribution, Import/export commands and Workset configuration for functionality required for collaboration, and Miscellaneous for the rest.

*TABLE. 5: Supergroups of users correspond to whose cluster types*

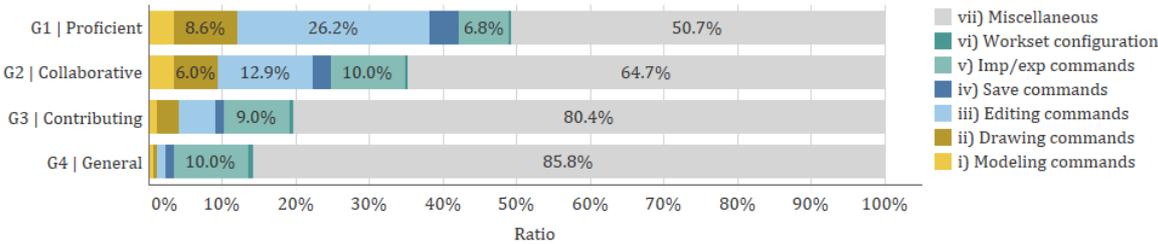|  | G1 \| Proficient | G2 \| Collaborative | G3 \| Contributing | G4 \| General |
|---|:---:|:---:|:---:|:---:|
| *Void Log* | ✓ | ✓ | ✓ | ✓ |
| *Dominant Log* | ✓ | ✓ | ✓ |  |
| *Major Log* | ✓ | ✓ |  |  |
| *Minor Log* | ✓ |  |  |  |



*FIG. 6: The ratio of command types per user supergroups*

While the Collaborative users show more Miscellaneous commands that do not indicate meaningful contribution to project progress, the proportion of commands for modeling and drafting is comparable to the top users. The difference lies in its composition ratio's lesser number of editing commands. Accordingly, there seems to be little significant difference between the G1 and G2 users in model creation. The difference between these users appears to be largely due to editing and modification.

On the other hand, the percentages of modeling and drafting commands stayed low in G3 and G4, and the non-modeling commands (total from iii to vii) exceeded 90%. Only 0.5% of modeling commands were performed by G4 users, indicating that they hardly engage in substantial modeling work. However, the Import/export command displays relatively high values, suggesting that some users execute various settings necessary for collaboration rather than modeling itself.

It is intriguing to notice that there is no prominent difference in the overall average number of commands per session between External and Permanent, even though External personnel supply most of the activity on BIM. This observation indicates that the number ratio within these supergroups causes the overall skew in the workload ratio toward External personnel.

We will now analyze the data grouped by projects to comprehend the responsibilities of the users in the actual projects as categorized above. Eight Revit projects had more than three members involved during the data collection period. Other projects were not captured here because they mainly used other BIM software such as ARCHICAD, were at such an early design phase that collaboration was not active yet, or was relatively small in scale. The analysis illustrates the work split among the 50 users assigned to these projects.

Fig. 7 depicts the total count of issued commands during the data collection period based on the affiliation of the project team members, divided into Permanent and External staff. The External staff executed more commands in all projects except Project F, prominently in Projects B, C, G, and H. It is noteworthy that the portions for External staff were all carried out by one to three staff members.
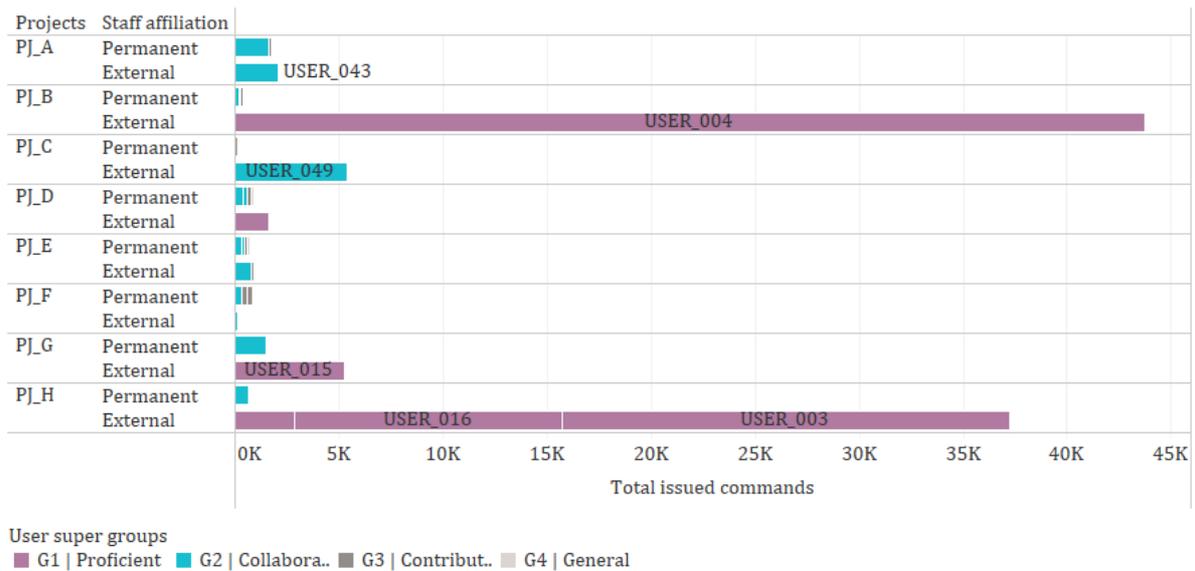
*FIG. 7: Average total issued commands per staff affiliation and projects*

## 3.2. Cross analysis and identifying keystones

As a next step, the authors further investigate the affiliation and supergroup of the staff assigned to each project. A cross-analysis is conducted to estimate their project contribution outside of Revit. Fig. 8 reports the total time dedicated to each of the six software sessions per project and the team members' affiliation, department, and supergroup.

Most projects have G1 or G2 BIM staff, who are the driving force for the BIM implementation. Other staff are designated as G3 or G4; nevertheless, their BIM contribution is not necessarily low, as these users may also utilize ARCHICAD or Rebro for considerably more time.

Here is where the G2 staff comes into focus. As seen in the previous section, while the G2s contribute more to modeling, their per-session average count of commands remains fairly lower than G1s. Their Revit utilization time also follows a similar tendency. However, these users recorded greater usage time than average in other software. In addition to AutoCAD for drafting environments and other modeling software, several users primarily rely on integrated environments such as Navisworks and Solibri.
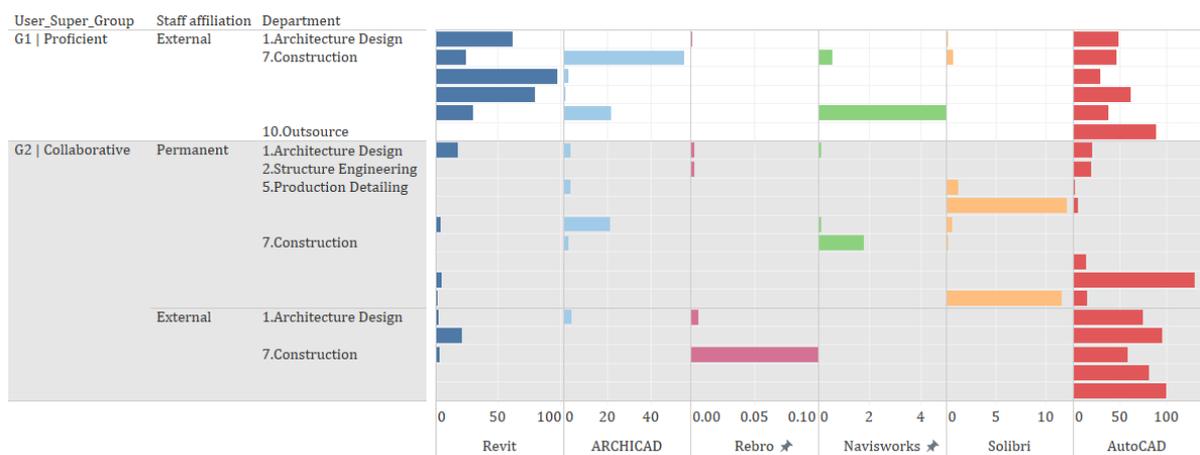


*FIG. 8: Project staff, supergroups and aggregated software session times*

The number of members within the four supergroups is tabulated by their affiliation and department in Table 6.

There is only four G1 personnel (4.7%) among Permanent staff but 30 External personnel (30.6%), which supports the previous projection that the main modeling workforce is from External personnel. Nevertheless, the Permanent staff classified as G2 accounts for 31% of the total, comparable to the 32.6% of the External staff. Few G3 and G4 users can be acknowledged because the External personnel is mostly BIM operators. Nevertheless, as mentioned in the previous section, it cannot be immediately concluded that this personnel is necessarily low in BIM use since they often have a high degree of non-Revit use.

The lower overall BIM utilization by Permanent staff is not because the utilization is sluggish overall but because there is relatively few G1 personnel. Therefore, the G2 personnel will likely bridge the workflow inside and outside the project.

Fig. 9 as a plot of the relationship between the accumulated Revit session time and the total executed commands shows a strong correlation between the two; the overall distribution of G1 – G4 is generally in line with the intensity of software use, while the distribution of G2 and G3 intersects with each other. This pattern indicates that Collaborative users such as G2 are difficult to stratify from one-dimensional information alone (simple metrics). It is appropriate to use machine learning approaches such as clustering to identify them.

TABLE 6: Number of data donors per department and affiliation

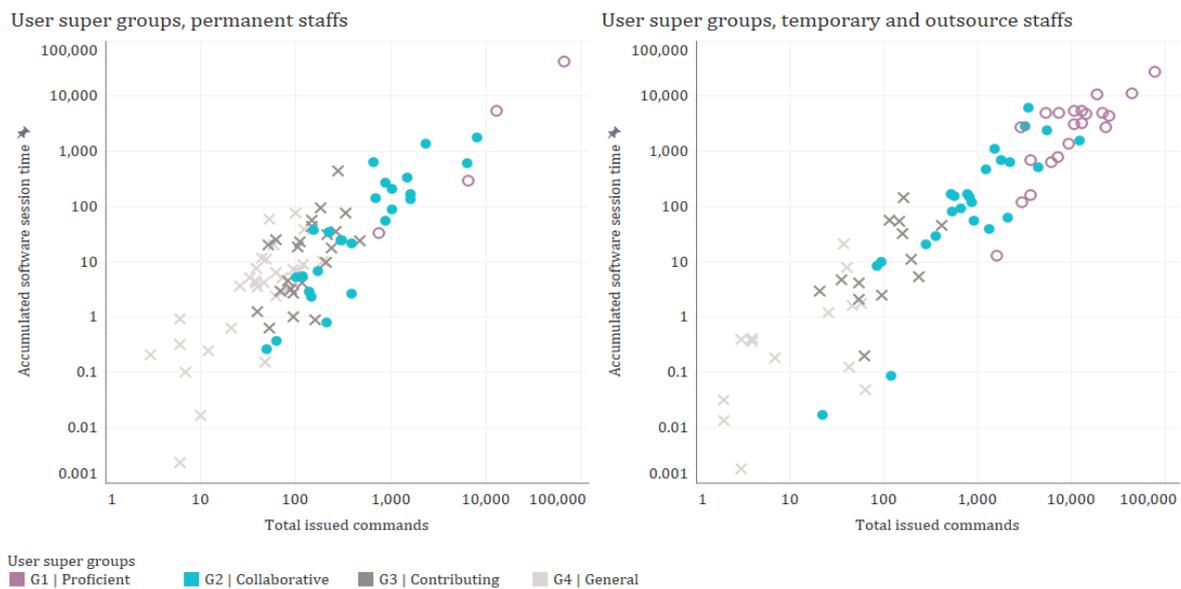| Section | Department | Permanent | | | | External | | | | Grand Total |
|---|---|---|---|---|---|---|---|---|---|---|
| | | G1 | G2 | G3 | G4 | G1 | G2 | G3 | G4 | |
| Design | Architecture Design | 1 | 2 | 5 | 6 | 6 | 6 | 4 | 1 | 31 |
| | Structure Engineering | | 7 | 7 | 1 | 4 | | 1 | 2 | 22 |
| | MEP Engineering | | 1 | 3 | 4 | | 1 | 1 | | 10 |
| | Computational Design | 1 | 4 | 1 | 3 | | 1 | 1 | | 11 |
| Construction | Production Detailing | 2 | 5 | 3 | 1 | 2 | 5 | 1 | 1 | 20 |
| | Quantity Survey | | | | 1 | | | | | 1 |
| | Construction | | 6 | 2 | 6 | 8 | 10 | 3 | 6 | 41 |
| | MEP Construction | | 1 | 1 | 6 | | | | | 8 |
| Administration | | | | 1 | 3 | | 1 | 1 | 2 | 8 |
| Outsource | | | | | | 10 | 8 | 3 | 9 | 30 |
| Grand Total | Total | 4 | 26 | 23 | 31 | 30 | 32 | 15 | 21 | 82 |



FIG. 9: User supergroups plotted on the total issued commands and accumulated software session time (Revit)

## 3.3. Verification by interviews

Given the prerequisite for collaboration with BIM Operators, expanding the G2 hierarchy of Permanent staff is crucial t on decision-makingIt is helpful to verify the mining results by investigating the cooperative relationship of Permanent staff classified as G2.

Semi-structured interviews were conducted with 6 of the 26 G2 Permanent staff members who consented to participation. The demographic information of interviewees is tabulated in Table 7, and the interview guideline is as detailed in Appendix A. Interviews with each user lasted one hour and were conducted via online calls using Microsoft Teams. The purpose and methodology of the study were announced in advance. However, the analysis results were not communicated to interviewees to avoid bias; instead, they were reported to interviewees who requested them after the process. The results were organized using an analysis template with a pattern coding technique and tabulated using the most important themes.

A common thread among the interviewees was that they always collaborated with BIM operators who split the work under their division's domain. The operators' skills were not always superior - employees were occasionally teaching them. A collaborative relationship existed in which the operators handled the principal modeling tasks, and G2 personnel consistently ensured that the External operators stayed focused on the BIM tasks (Table 8).

*TABLE 7: Interviewees' demographic information*

| # | | Interview date | Affiliation | Position | Experience in AEC |
|---|---|---|---|---|---|
| 1 | G2 | July 10, 2020 | Architectural design, Permanent | Project architect of two residential design-build projects | > 3 years |
| 2 | G2 | July 2, 2020 | Architectural design, Permanent | Project architect of a complex high-rise project, design development | > 4 years |
| 3 | G2 | June 29, 2020 | Construction Planning, Permanent | BIM technical support for ten (10) construction projects | > 12 years |
| 4 | G2 | June 30, 2020 | Construction Planning, Permanent | BIM technical support for nine (9) construction projects | > 10 years |
| 5 | G2 | July 2, 2020 | Structural design, Permanent | Project structural engineer of a commercial design-build project | > 11 years |
| 6 | G2 | July 8, 2020 | Architectural design, Permanent | Project architect of an office design-build project | > 14 years |

*TABLE 8: Availability and role of the BIM operator in projects*

| Question | Drafted results |
|---|---|
| Model co-creator | • (#1) A BIM operator (External). |
| | • (#2) Two BIM operators (External). |
| | • (#3) Seven BIM operators (all External). Some operators are skillful in visual scriptings such as Dynamo or Grasshopper or real-time rendering by Lumion or TwinMotion. |
| | • (#4) Five to ten BIM operators (all External) |
| | • (#5) A BIM operator (External) and another Permanent structural engineer |
| | • (#6) A BIM modeling firm (External) |
| Project model contributor | • (#1) The BIM operator undertook most modeling work, who has experienced decades in the industry and is especially skillful for detail design and construction documentation. |
| | • (#2) The BIM operators exclusively undertook modeling and drawing tasks. |
| | • (#3) The BIM operators. Upon receiving the project request, the interviewee assigns seven BIM operators. |
| | • (#4) The BIM operators, unless the task is highly urgent. Some operators are to be trained before being deployed to the construction projects. |
| | • (#5) The modeling tasks are mostly done by the BIM operator. As the Permanent staff, the engineer reviewed and verified the model against the design and calculation results. |
| | • (#6) Though the modeling process was mainly self-service until the schematic design, some repetitive tasks like designing staircases are shared with the BIM operators. In the detail design phase, the drawing-related tasks are largely outsourced, and the Permanent staff concentrated on project communication to eliminate the informational inconsistency. |

Although each G2 employee responded that the External BIM operator performs most tasks, they still complete certain tasks independently. Particularly for architects, this included the production of perspective images to communicate design intent and submission drawings that demanded specific knowledge; for constructors, this included urgent tasks that imposed a deep understanding of project contexts.

All respondents viewed the presence of BIM operators positively and stated that their collaboration could benefit the project by allowing employees to specialize in tasks that cannot be outsourced. Through their own project experience, the respondents pointed out that the employees themselves should not undertake the same level of work as the operators but should acquire the same level of knowledge. The rationale for this point was primarily the need to ensure proper communication and cooperation and the importance of being respectful to experienced BIM operators (Table 9).

It is intriguing to note that all interviewees initially stated that they were uncertain about why they were selected for BIM-related interviews because they regarded themselves as making no significant contribution to BIM. They were aware that their BIM workload was not as large as that of the operators on the project, and hence they did not consider themselves worthwhile BIM talents for the company.

*TABLE 9: BIM collaboration scheme*

| Topic | Drafted results |
|---|---|
| Role of project architect as Permanent staff | ● (#1, #2, #6) Rendered images to visualize architect impression<br>● (#1) Submission documents to the authority clearance<br>● (#3, #4) Communicational support. Project requests often require to be decomposed for the operators to execute. Permanent staff can swiftly understand the background context and interpret the required deliverable precisely. |
| Need for BIM operator | ● (#1) BIM operator is ideal to have in any project. Permanent architects could carry out quick changes, but gaining BIM skills during the project is challenging.<br>● (#2) BIM operator is vital for BIM project regardless of project scale because Permanent staff should spend most of the project time for decision making<br>● (#3) The need for a BIM operator is parallel to CAD. Permanent staff must be responsible for some tasks, including the client's decision-making or the authority submission. BIM operators need supervision to connect with the project need appropriately.<br>● (#4) Although the central technical support for BIM will be diminished over the years, BIM operators will be in need in the future. Reducing communicational stress is necessary for a better relationship with each other.<br>● (#5) Regardless of whether they are BIM technicians or not, supporting staff are vital for timely project progress. The works of modeling, coordination, engineering, and documentation must be balanced in total. The more skillful BIM staff is, the more Permanent staff enhances engineering expertise.<br>● (#6) Most projects in general contractors require collaboration, including BIM, in terms of scale and complexity. BIM collaboration is inevitable to meet high expectations for on-time delivery, especially during the detail design phase. |
| The cooperative relationship with BIM operators | ● Permanent staff should equip intermediate BIM skills to enable smooth communication with BIM operators.<br>● Insufficient BIM skills will lead to a lack of respect for skillful BIM operators.<br>● Permanent staff should concentrate on the tasks where they are fully responsible. Some tasks can be shared with the External staff for timely project delivery.<br>● While technical challenges in BIM will be transitionally resolved, collaboration with the BIM operators will have been expected to pursue additional value to projects. |

# 4. DISCUSSION

## 4.1. Keystone species discovered by log mining

All participants were stratified into four clusters from G1 to G4, corresponding to the four superordinate clusters, based on the results of VBGMM clustering. While some users classified as G4 had only void logs showing limited contribution, G1 users with high contribution had logs corresponding to all the superclusters. The G1 group was the one who intensively performed the modeling work, with the distribution being heavily concentrated on

External personnel. The differences in modeling duties between Permanent and External personnel were evident from the collected logs. Thus, there is no question that the External staff undertakes most of the modeling work in this ecosystem.

G2 employees utilized BIM less intensely than the G1; hence their contribution to the project BIM was relatively low. Unlike the G3 and G4 groups, however, the percentage of modeling-related commands is comparable to that of the G1 group, indicating that a substantial contribution to modeling progress is observed. This composition implies that the G2 personnel has a reasonable degree of skills and literacy in BIM. Furthermore, the G2 human resources were widely spread among Permanent and External employees.

The abundance of G2 users in both affiliations is worth reiterating. At least one G2 personnel was identified in all but a few departments. Larger organizations tend to organize a promotion department to consolidate and assign BIM personnel. It was considered reasonable to enhance the efficiency of human resource utilization. This strategy may be well-suited for task-specific personnel such as G1 talents. On the other hand, G2 personnel, who would be pivotal in collaboration, are likely to arise in each department from the respective necessities. When highly specialized personnel are intensively placed, it would be preferable to broadly appoint collaborative users like G2.

The project-level analysis also revealed that the G1 workforce is the primary promoter. However, simultaneously, the G2 workforce occupies the secondary BIM driving force position, including a bridge to other software use.

Since it was assumed that the G2 Permanent employees were the keystone species of the project, the verification interview was conducted. As a result, it was discovered that they positively viewed the collaboration with External personnel and acquired BIM knowledge to handle very important tasks independently and respectfully interact with them. It is noteworthy that they did not perceive themselves as occupying an important position in its BIM promotion. Fig. 10 summarizes the concept of keystone BIM players in the ecosystem.
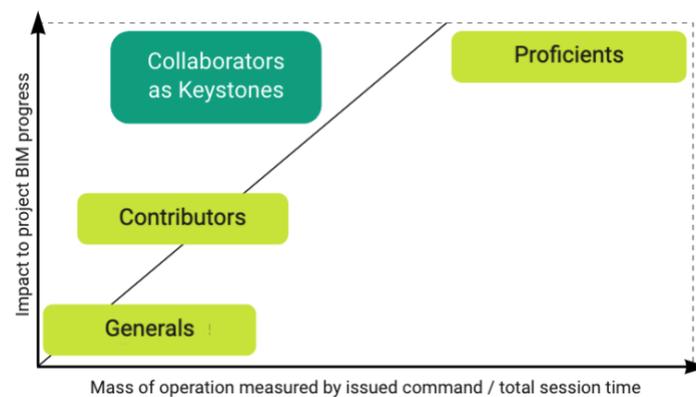


FIG. 10: Collaborators as the keystone species in project BIM ecosystem

## 4.2. Implication of the keystone species in Japanese AEC ecosystem

Large general contractors have led the spread of BIM in Japan. The majority of the activities are undertaken by External personnel hired on a short-term basis. Permanent employees are almost exclusively assigned to tasks of higher importance to the project. Nevertheless, the Permanent employees must possess adequate BIM knowledge and experience to ensure appropriate direction and approval. It is extremely valuable for general contractors to strengthen the G2 staff since expanding the G1 workforce is hard to achieve.

BIM and CAD operators often have more experience and domain knowledge than the average Permanent project member. In order to establish cooperative collaboration regardless of job responsibilities, it is worthwhile for them to be motivated to acquire knowledge and strive to attain the equivalent level of knowledge themselves, rather than just to stand in a leadership position. In a high-context environment, detailed instructions are often bypassed in quick communication. However, the exact interpretation of the substance is challenging for outside personnel. Therefore, employees who can conscientiously understand the instructions are encouraged to act as interpreters who also comprehend the work performed in BIM.

BIM skills have been focused on modeling; thus, training has emphasized modeling skills. However, in such a collaborative environment, it is more appropriate to focus on the knowledge required for collaboration rather than on specific operations. Due to the steep learning curve of BIM and the high learning cost of learning to model, it is easy to fall behind when trying to learn it in parallel with practical work. It is important for the company and for the industry to improve the G3 and G4 personnel levels by increasing overall efficiency by focusing on more practically effective points that can be easily applied to the current business.

Even more important is the penetration of BIM to professionals outside of BIM; as BIM is a communication channel, parallel conversations in non-BIM environments are not desirable. However, BIM frequently comes to a mere aggregation or transcription of consolidated information in reality; therefore, the information available from BIM is regarded as unreliable or incomplete. While practitioners should have the right to choose their tools, information sources alienated from project conversations become untrustworthy or even worth, useless.

The Tokyo office of the data-providing contractor houses more than 2,000 employees, and the members under the design and production departments alone easily exceed 1,000. These facts imply that G3s and G4s are not inferior to the whole but are even early adopters when covering the whole professionals. It would be highly realistic to recognize the contribution of G3s and G4s and promote them to participate in a broader range of BIM applications. Arising and consolidating their importance in project communication should positively influence rather than encourage them to become G1s.

The general contractor runs approximately 200 projects concurrently. To deploy at least one collaborative BIM staff for each design and construction, 400 G2 users are expected, equivalent to nearly 30% of the technical workforce. G2 users currently account for about 30% of the surveyed personnel; however, this is the percentage of personnel limited to BIM-enabled professionals. Even more non-BIM personnel currently exist. In line with Rogers's diffusion of innovation, G2 is expected to disperse from early adopters to most BIM users (Fig. 11). This layer will expand when the technology is sufficiently penetrated; it is theoretically possible for G2 to occupy the entire majority. BIM log mining provides the indispensable dashboard to identify and nurture the individual layers of organization BIM users.



FIG. 11. Distribution of user supergroups jointly interpreted with the adopter categorization based on innovativeness, Source:(Everett, 1983), copyright (2008) by the title of publisher.

In cases like the Japanese construction industry, where a cooperative process is a prerequisite, the labor split does not completely follow the functional separation but is defined broadly and vaguely according to the duties and authorities. As with the conventional practice in construction, although subcontractors may be responsible for specific tasks, the architect at the center is expected to be the final responsible one and be in a position to supervise the project comprehensively.

Considering the organizational design and individuals' skills are deeply relevant to the issue, the proposed methodology depicts the fundamental strategy to improve BIM use continuously. Viewing a BIM model from the outside does not provide any insight into the day-to-day workings of the model. Using tools that are understandable to such stakeholders is extremely important to generate more interest through exposure.

Large private companies like general contractors are highly influential in the BIM deployment in Japan. Their project cases are influential enough to improve construction productivity of the entire industry, especially with BIM, when their organizational design and learning aim to build a cooperative ecosystem rather than focus on modeling skills.

## 5.  LIMITATION OF STUDY AND FUTURE PROSPECT

This study concentrated on a single general contractor. Thus it is not obvious if the results apply to other organizations in the sector. Despite the plausibility of the method, obtaining employee attributes besides software logs is not always easy. If broader data is available to benefit from this technique, it will be possible to draw clearer trends in the entire ecosystem of Japan's construction industry in a data-driven manner.

Another limitation lies in the log mining of other software. Despite several studies on IFC, the mining methods for most software without operation logs are still unexplored; thus, a procedure to extract the equivalent knowledge from Revit has not been successfully implemented. In this study, we attempted to consider this area by analyzing the time spent; however, as seen earlier in the study, the software session time is not necessarily proportional to the BIM contribution. Alternatively, users who show regular collaboration with other users may be considered keystone users regardless of their usage time. It should be possible to monitor the activity intensity from the communication platform or cloud environment processes.

Finally, drafting professions specializing in preparing application drawings for the governmental authority, called agency workers, have existed in major urban areas since the 1930s, prior to the institutionalization of licensed architects with the enactment of the Architects Act in 1950 (Hayami, 2006). It must be highlighted that the research on CAD/BIM operators in Japan is still underway; there are few statistics or studies on their employment status, roles, contributions, and treatment. The architecture can only be established through collaboration; therefore, it is imperative to recognize the existence of such supporters in the project in order to foster discussions on better collaboration.

## 6.  CONCLUSION

This article conducted a cross-analysis of Revit log files and the aggregated session time of six different BIM applications in a major Japanese general contractor. External staff performed the greater part of the BIM work; most of the Permanent staff's BIM efforts were accomplished by very few employees. The clustering process in BIM log mining yielded four main types of logs. User stratifies these four types of logs; users containing all types of logs are named Proficient users (G1), who have the widest range of log types and contribute the most to the modeling process. The second group, Collaborative users (G2), showed the percentage of modeling contribution comparable to the Proficient users, although the contribution intensity was relatively low. Visual analytics revealed that while Proficient users comprise only a few employees, Collaborative users are widely spread across the organization and projects, comprising about 30% of Permanent and External staff.

Collaborative users tend to use Revit sparingly but spend much time with other software. Among such Collaborative users, it indicated that Permanent staff, in particular, are likely the key to promoting BIM projects.

Semi-structured interviews further explored the common characteristics of Collaborative users. The findings demonstrated that Collaborative users are positive about collaborating with Proficient BIM operators; while they can perform certain operations themselves, they partition the labor to concentrate on their specialty. Notably, their learning motivation was to acquire a comparable level of knowledge for showing respect to the operators. As such, the Collaborative members among Permanent staff are considered the keystone species in the ecosystem of large-scale projects, serving as the pillars of knowledge accumulation and the collaboration hub in organizations.

The contributions of the thesis are immense as the findings contribute to the existing body of knowledge. It also contributes to furthering good industry practice. First, the proposed method extends the existing BIM log mining technique. The vast amount of datasets explored and verified the utility of the proposed method, which combined the visual analytics and partitional clustering algorithm as a novel approach. It should be reiterated that the proposed method applies to multidisciplinary logs, which is essential functionality for analyzing the design-and-build projects. Second, the article introduces the keystone species concept that became possible by analogizing the BIM environment as an ecology. As the keystone species are the ones who have a tremendous influence on the

surroundings once they are removed, the identified keystone species should be nurtured and even augmented to upscale the BIM collaboration. Last, the research shed light on the cooperative relationship in the Japanese AEC industry. Although it was partially recognized that collaborative attitudes are widely seen in the construction field, it was hardly mentioned that similar events happen throughout the project, including the design phase. The research on BIM operators or other external supporters in Japan has been scarcely made. Their employment status, roles, contributions, and treatment should be uncovered for further constructive discussions. The architecture can only be established through collaboration; therefore, it is imperative to recognize the existence of such supporters in the project to foster discussions on better collaboration.

# REFERENCES

Abdel-Wahab, M. and Vogl, B. (2011) 'Trends of productivity growth in the construction industry across Europe, US and Japan.' *Construction Management and Economics*, 29(6) pp. 635–644.

Andrienko, N., Andrienko, G., Fuchs, G., Slingsby, A., Turkay, C. and Wrobel, S. (2020) *Visual Analytics for Data Scientists*. Cham: Springer International Publishing.

Autodesk (2017) *About Journal Files*. Autodesk Knowledge Network. [Online] [Accessed on 24th March 2019] https://knowledge.autodesk.com/support/revit-products/getting-started/caas/CloudHelp/cloudhelp/2017/ENU/Revit-GetStarted/files/GUID-477C6854-2724-4B5D-8B95-9657B636C48D-htm.html.

Barison, M. B. and Santos, E. T. (2011) 'The Competencies of BIM Specialists: A Comparative Analysis of the Literature Review and Job Ad Descriptions.' *In Computing in Civil Engineering (2011)*. Miami, Florida, United States: American Society of Civil Engineers, pp. 594–602.

Bennett, J., Flanagan, R., Norman, G. and PLC, C. & C. (1987) *Capital & Counties Report: Japanese Construction Industry*. London: Centre for Strategic Studies in Construction, University of Reading (Capital and countries report).

Buntrock, D. (2002) *Japanese Architecture as a Collaborative Process: Opportunities in a Flexible Construction Culture*. London: Spon Press.

Celebi, M. E. (ed.) (2015) *Partitional Clustering Algorithms*. Cham: Springer International Publishing.

Chien, K.-F., Wu, Z.-H. and Huang, S.-C. (2014) 'Identifying and assessing critical risk factors for BIM projects: Empirical study.' *Automation in Construction*, 45, September, pp. 1–15.

Corduneanu, A. and Bishop, C. M. (2001) 'Variational bayesian model selection for mixture distributions.' *In Proceedings Eighth International Conference on Artificial Intelligence and Statistics*. Morgan Kaufmann, pp. 27–34.

Ejima, S., Oka, M. and Ikegami, T. (2019) 'Concept of Keystone Species in Web Systems: Identifying Small Yet Influential Online Bulletin Board Threads.' *In Proceedings of the 10th ACM Conference on Web Science - WebSci '19*. Boston, Massachusetts, USA: ACM Press, pp. 81–85.

Everett, R. (1983) *Diffusion of Innovations, Third Edition*. New York: The Free Press.

Forcael, E., Martínez-Rocamora, A., Sepúlveda-Morales, J., García-Alvarado, R., Nope-Bernal, A. and Leighton, F. (2020) 'Behavior and Performance of BIM Users in a Collaborative Work Environment.' *Applied Sciences*, 10(6) p. 2199.

Gao, W., Wu, C., Huang, W., Lin, B. and Su, X. (2021) 'A data structure for studying 3D modeling design behavior based on event logs.' *Automation in Construction*, 132, December, p. 103967.

Hayami K. (2006) 'Formation and Development Process of System of Architectural Agent for Procedures Required by Law:A study on enactment process of the Kenchikushi Law for architects & building engineers Part 2.' *Journal of Architecture and Planning (Transactions of AIJ)*, 71(601) pp. 199–204.

Hileman, J., Kallstenius, I., Häyhä, T., Palm, C. and Cornell, S. (2020) 'Keystone actors do not act alone: A business ecosystem perspective on sustainability in the global clothing industry.' Angeli, F. (ed.) *PLOS ONE*, 15(10) p. e0241453.

Ishida, K., Shide, K. and Iki, T. (2016) 'Study on Corporate Organization in Construction Companies Through Technology Introduction of BIM.' *Journal of Architecture and Planning (Transactions of AIJ)*, 81(726) pp. 1743–1753.

Ishizawa, T. and Ikeda, Y. (2021) 'Visual log analysis method for designing individual and organizational BIM skill.' *Journal of Architectural Informatics Society*, 1(1) pp. a1–a21.

Kaneta, T., Furusaka, S. and Deng, N. (2017) 'Overview and problems of BIM implementation in Japan.' *Frontiers of Engineering Management*, 4(2) p. 146.

Kouhestani, S. and Nik-Bakht, M. (2020) 'IFC-based process mining for design authoring.' *Automation in Construction*, 112, April, p. 103069.

Ku, K., Pollalis, S. N., Fischer, M. A. and Shelden, D. R. (2008) '3D model-based collaboration in design development and construction of complex shaped buildings.' *Journal of Information Technology in Construction (ITcon)*, 13(19) pp. 258–285.

MacQueen, J. (1967) 'Some methods for classification and analysis of multivariate observations.' *In*. The Regents of the University of California.

Mathews, M. (2015) 'Defining Job Titles and Career Paths in BIM.' *In CITA BIM Gathering 2015*. Dublin: Dublin School of Architecture, p. 7.

Ministry of Land, Infrastructure, Transport and Tourism (2017) 'The current status of construction industry and construction workers.' Ministry of Land, Infrastructure, Transport and Tourism.

Ministry of Land, Infrastructure, Transport and Tourism (2021) *Fact-finding investigation into the adoption and utilization of Building Information Modeling in the building and construction sector, translated by the author*. Ministry of Land, Infrastructure, Transport and Tourism.

Ministry of Land, Infrastructure, Transport and Tourism and Ministry of Health, Labor and Welfare (2020) *Press release: to secure and develop human resources in the construction industry, translated by the author*. [Online] [Accessed on 21st November 2021] https://www.mlit.go.jp/report/press/content/001380300.pdf.

Munk, M., Kapusta, J. and Švec, P. (2010) 'Data preprocessing evaluation for web log mining: reconstruction of activities of a web visitor.' *Procedia Computer Science*. (ICCS 2010), 1(1) pp. 2273–2280.

Nakamura T. and Ito T. (2018) '3D-CAD School for Working Adults and Education Effectiveness on Campus - Education Promotion Program Responding to Working Adults' Needs for Re-learning -.' *In*. Japanese Society for Engineering Education.

Ogasawara, M. and Yashiro, T. (2017) 'The Task Distribution Among Architects: Matrix of Responsibilities in US projects.' *Journal of Architecture and Planning (Transactions of AIJ)*, 82(738) pp. 2029–2039.

Ogasawara, M. and Yashiro, T. (2018) 'A Comparative Study of the Design Process in General Construction Companies and Design Firms in Japan.' *Journal of Asian Architecture and Building Engineering*, 17(1) pp. 31–38.

Pan, Y. and Zhang, L. (2020a) 'BIM log mining: Exploring design productivity characteristics.' *Automation in Construction*, 109, January, p. 102997.

Pan, Y. and Zhang, L. (2020b) 'BIM log mining: Learning and predicting design commands.' *Automation in Construction*, 112, April, p. 103107.

Pan, Y. and Zhang, L. (2021) 'Automated process discovery from event logs in BIM construction projects.' *Automation in Construction*, 127, July, p. 103713.

Pan, Y., Zhang, L. and Li, Z. (2020) 'Mining event logs for knowledge discovery based on adaptive efficient fuzzy Kohonen clustering network.' *Knowledge-Based Systems*, 209, December, p. 106482.

Park, H.-J. and Lee, J.-H. (2010) 'Exploring Integrated Design Strategies for the Optimal Use of BIM.' *Architectural research*. Architectural Institute of Korea, 12(2) pp. 9–14.

Seow, J. (2014) 'Shortage of skilled staff in some sectors with 63,900 vacancies in June.' *THE STRAITS TIMES*. [Online] 16th September. [Accessed on 19th December 2019] https://www.straitstimes.com/singapore/shortage-of-skilled-staff-in-some-sectors-with-63900-vacancies-in-june.

Sharag-Eldin, A. and Nawari, N. O. (2012) 'BIM in AEC Education.' American Society of Civil Engineers, April, pp. 1676–1688.

Shide K. (2015) 'Current Status of BIM in the Building Industry - Comparison of the United States, Singapore, South Korea and Japan -.' *Development Engineering*, 35(2) pp. 109–112.

Suzuki, K. and Sui Pheng, L. (2019) *Japanese Contractors in Overseas Markets: Bridging Cultural and Communication Gaps*. Singapore: Springer Singapore (Management in the Built Environment).

Teicholz, P. (2013) *Labor-Productivity Declines in the Construction Industry: Causes and Remedies (Another Look)*. AECbytes Archived Article. [Online] [Accessed on 17th November 2018] http://www.aecbytes.com/viewpoint/2013/issue_67.html.

Uhm, M., Lee, G. and Jeon, B. (2017) 'An analysis of BIM jobs and competencies based on the use of terms in the industry.' *Automation in Construction*, 81, September, pp. 67–98.

Webster, A. C. (1993) 'Japanese Building Design and Construction Technologies.' *Journal of Professional Issues in Engineering Education and Practice*, 119(4) pp. 358–377.

Yamada H. and Isoyama H. (2018) *A study on future prospects and human resources strategy in the field in conjunction with Ministry of Land, Infrastructure, Transport and Tourism*. Tokyo, Japan: Policy Research Institute for Land, Infrastructure, Transport and Tourism (PRILIT) (PRILIT Research Report), p. 38.

Yarmohammadi, S., Pourabolghasem, R. and Castro-Lacouture, D. (2017) 'Mining implicit 3D modeling patterns from unstructured temporal BIM log text data.' *Automation in Construction*, 81, September, pp. 17–24.

Yarmohammadi, S., Pourabolghasem, R., Shirazi, A. and Ashuri, B. (2016) 'A sequential pattern mining approach to extract information from BIM design log files.' *In 33th International Symposium on Automation and Robotics in Construction*. Auburn, AL, USA: International Association for Automation and Robotics in Construction, pp. 174–181.

Yoshikai, K., Tsutsui, T. and Takaguchi, H. (2018) 'Evaluation of Design Efficiency Improvement by Utilizing BIM by Proceedings Analysis.' *Journal of Architecture and Planning (Transactions of AIJ)*, 83(749) pp. 1297–1303.

Zhang, L. and Ashuri, B. (2018) 'BIM log mining: Discovering social networks.' *Automation in Construction*, 91, July, pp. 31–43.

# APPENDIX. SEMI-STRUCTURED INTERVIEW GUIDE

**Start of interview**

- Introduce the interviewer to the interviewee.
- Provide a background of the interview:
  - The aim is to verify the discoveries from BIM log mining conducted on the software records across the firm, including the interviewee.
  - The reason for selecting you as an interviewee is because the interviewee's BIM activity may represent certain tendencies in terms of organizational BIM implementation.
  - The interviewer requests the interviewee's project and situation. Please answer your personal experience and opinion; it needs not consider the broader perspective.
  - The mining result will not be disclosed since the analysis is in progress; however, the interviewee will share the published article afterward.
- There will be an opportunity for questions at the end.
- The interview begins when the interviewee understands and agrees with the above condition.


1. Interviewee personal details
   1.1. Name:
   1.2. Position: (management / project architect / BIM promotion team / BIM operator)
   1.3. Qualification: (licensed architect / licensed engineer / none)
   1.4. Years of experience:
2. Project details the interviewees were in charge during the log collection period
   2.1. Project name:
   2.2. Project location:
   2.3. Project owner:
   2.4. Type of project: (residential / commercial / industrial / educational / ...)
   2.5. The number of floors:
   2.6. Estimated project duration in month: (levels of aboveground / levels of the underground)
3. Use of BIM software (Revit)
   3.1. What was the primary reason for implementing BIM in the project?
   3.2. Who was the decision-maker to adopt BIM for the above?
   3.3. Did the project have the BIM execution plan (BEP)?
   3.4. To what extent was the project required to utilize Revit or other BIM applications?
4. Role of interviewee
   4.1. Have you been navigating Revit on your own?
      4.1.1. If yes, have you been modeling by yourself?
      4.1.2. If no, how have you been proceeding with the project by Revit?
   4.2. Have you had anybody else who (co-)created the project model?
      4.2.1. If yes, were they the dedicated BIM operators?
      4.2.2. If yes, were they the External staff?
   4.3. Was the managerial personnel involved in the BIM process?
      4.3.1. If yes, were they supportive of implementing and leveraging BIM in the project?
      4.3.2. If no, what did you observe was the chief reason?
5. Contribution to modeling
   5.1. Who most contributed to the progress of the modeling?
   5.2. How do the rest project members collaborate with the model contributor identified in Q5.1?
6. Project self-evaluation
   6.1. What was the benefit of Revit identified throughout the project?
   6.2. Did the BIM implementation progress cooperatively?
   6.3. In your opinion, do you think it was ideal to proceed with BIM in such collaboration scheme mentioned in Q4?(Hayami, 2006)
   6.4. What was the primary factor that contributed to enhancing project productivity?
7. Shortcomings and proposals for future projects

  7.1.  What were the shortcomings identified in implementing BIM to the project?
  7.2.  Do you think you should undertake more modeling roles than current ones in future projects?
    7.2.1.  What is the reason for the above opinion?

**End of interview**

- Ask the interviewee if they have any further information to share.
- Ask the interviewee for further questions
- Thank the interviewee for their time and help