



Creative Construction Conference 2019, CCC 2019, 29 June - 2 July 2019, Budapest, Hungary

## Toward a qualitative RFIs content analysis approach to improve collaboration between design and construction phases.

Mathieu FOKWA SOH<sup>a</sup>, Daniel BARBEAU<sup>b</sup>, Sylvie DORE<sup>a</sup>, Daniel FORGUES<sup>a,\*</sup>

<sup>a</sup>*École de technologie supérieure, Montréal, Canada*

<sup>b</sup>*Groupe CANAM, Boucherville, Canada*

---

### Abstract

Requests for Information (RFIs) are formal processes, used in the industry of Architecture, Engineering and Construction (AEC), to obtain information not contained or inferable in the contract documents between the design and construction phases. RFIs produce rich, precise and structured sources of information. Analysis of RFIs content can help identify recurrent problems. The goal of this article is to present a method to identify problem areas during the construction phase of AEC projects through the analysis of RFI documents. Recent advances in the qualitative analysis of document content make this quest possible and fast. This article proposes to the scientific communities and AEC industry professionals, a systematic method based on the qualitative analysis of RFIs in order to propose some types of information to consider for a design more adapted to the construction phase. An example of the application in a steel construction project demonstrates the feasibility of this method and proposes some points to consider to improve the design of steel structures.

Keywords:

© 2019 The Authors. Published by Budapest University of Technology and Economics & Diamond Congress Ltd.

Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2019.

*Keywords: construction phase; design phase; qualitative content analysis; request for information,*

---

### 1. Introduction

The success of a project in the AEC industry depends on, among other things, the quality of collaboration between designers and contractors [1,2]. Designers provide information that contractors use to create buildings that meet the customer's needs [3]. However, eighty-eight percent of decisions related to the duration and the cost of projects are made by designers [4]. They often have no experience or knowledge of the construction operations that follow and do not always provide all the information needed during the construction phase. Contractors, who are later involved in the project supply chain [5,6], will formulate RFIs to request additional information and propose modifications [7,8]. The RFIs they emit are numerous (average 796 RFIs per project [8]), and they arrive when the design phase is quite advanced. At this moment, design changes have a significant impact on the cost and duration of the project [9,10]. These impacts represent on average of 79% of the total modification costs and 9.5% of total projects costs [11]. If the information requested in the RFIs is known in advance by the designers and communicated properly in the project documents, the cost and duration of the project could be reduced. This article proposes to use qualitative content

analysis of RFIs, to highlight recurrent information needs requested by contractors from the designers, in order to allow the designers to take this information into account when preparing construction documentation. To do this, the technique of *the qualitative analysis of the contents* is presented. Then, a methodology is proposed and a case study proposes the application of this method in the context of the steel construction industry.

## 2. Literature review

An RFI is by definition produced to request information or clarification from designers [12]. The quantity of RFIs depends on the quality of the design [7]. RFIs are very expensive [13]. It takes on average nine days to reply to one RFI [8]. The RFI response cost for projects of 1 - 10 million USD can range from 598 to 2 078 USD per RFI [13], which corresponds to 13 532 hours-person on average per project.

RFIs are a reliable source of information. For [12] RFIs are the standard communication approach between designers and builders. They respond to a standard writing protocol that aims to make information needs very efficient. Each RFI relates to a single technical problem. Questions and proposals for changes in RFIs are usually clear, complete and signed by an official who assumes responsibility for the request. These features of RFIs made by [12] are part of the standard steel practice code established by the American Institute of Steel Construction (AISC), and give the RFIs great credibility in terms of source of information. This quality can be exploited to improve the quality of construction projects, especially the quality of the design phase.

The use of RFIs to improve the quality of construction projects has already been the subject of some studies. [26] used RFIs to provide quantitative information on the relationship between selected RFI variables and the performance of the shop drawing production process [26]. They also determined significant associations between shop drawing production performance and production performance in terms of cost and duration. Regression techniques were used to identify individual RFI variables with significant influence on the models to predict the production performance of shop drawings.

In an analysis of the causes, effects and indicator of design flaws, [3] also proposes to quantify the causes of RFIs as a criterion for evaluating design flaws in construction projects. And in this case, it is through workshops with manufacturers that the causes of RFIs have been identified.

These studies demonstrate that RFIs can play a role in improving the quality of projects. However, there is a need for a method to analyse RFIs and retrieve the information needed to be included in design document, in order to improve the quality of the design phase.

Qualitative content analysis (QCA) is a research method that applies to textual data intended for human understanding [14]. The aim is to propose and classify knowledge and understanding of phenomena present in texts through a subjective interpretation [15]. This analysis can be done manually or automatically with software.

In manufacturing engineering, a few studies have been conducted using QCA. For example, Dong et al (2004) used analysis of the contents of design documents to study the performance of design teams [16]. Mabogunje (1997) proposed to measure the creativity of a design team based on the documents generated during the design phase [17]. Zhang et al. (2017) have developed a systematic method for delving into design journals and discovering the social relationships and performance of design teams [18]. To our knowledge, no initiatives propose to use QCA and RFIs to improve the quality of design in the AEC industry.

## 3. Methodology

There are three approaches in qualitative content analysis [19]: the conventional approach that describes the limited phenomena in the literature from the contents [20] ; the directed content analysis approach, used to validate or extend, a framework or theory conceptually [21] ; the summative content analysis approach used to identify and quantify certain words or themes in order to understand the contextual use of words or content [22]. This article wanted to find

common themes in a series of textual documents. So, the summative content analysis approach was adopted to identify the type of information requested in RFIs.

The case study for this article is a project carried out by a steel structure design and manufacturing firm between August 2014 and December 2017. The number of RFIs for the project was 11 426. The sum of RFIs response time was 96 277 days, with an average of 8.4 working days for each RFI.

To analyze the content of RFIs, the following steps were followed (see figure 1):

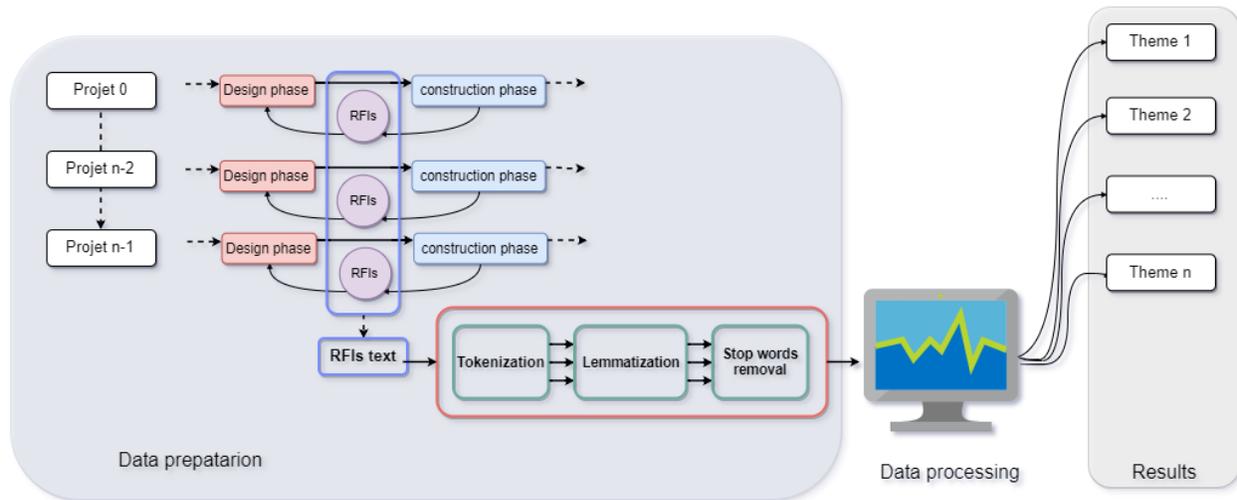


Figure 1: Methodology

### 3.1. Data preparation

In this step, the RFIs of previous projects are firstly collected in digital text format (docs, txt, csv). Secondly, they are treated to make them suitable for the machine [23]. This includes *tokenization*, *lemmatization* and *stop word removal*. The process of lemmatization or stemming, is a process of representing words in such a way as to retain only their meaning; for example, nouns are converted to singular masculine nouns, and verbs are converted to the infinitive. Stop word removal is used to simplify the text by removing words that have no major importance in communication, and, tokenization is used to divide the flow of textual content into words, terms, symbols or other significant elements called tokens [24]. After this step, the texts are ready to be processed by software.

### 3.2. Data processing

In this step, the most important words and themes are collected and grouped together as a table [25]. Themes represent the main problems faced by the contractors.

### 3.3. Data post processing.

To validate the importance of the themes found, the number of RFIs related to these themes are counted. Furthermore, the waiting time associated to each related RFI is extracted from the original database and summed for all the RFIs related to a given theme.

## 4. Results.

The results are as follows. The list of themes most present in RFIs and their impact on the number and response time of RFIs are presented in Table 1. Some RFIs contain several themes. For this reason, the total of the percentage is greater than 100%.

Table 1. Theme frequency for the case study

themes	Example of tokens	RFIs containing theme	% of RFIs containing theme	elapsed time for RFIs	% of the elapsed time for RFIs
total Items		11 426		96 277	
areas	area (area, areas)	5 298	46.37%	58 689	60.96%
connection	Member connection (connect, connected, connecting, connections)	3 041	26.61%	35 841	37.23%
trusses	Additional truss (trusses)	2 458	21.51%	28 499	29.60%
missing	missing (miss, missed) hole	1 556	13.62%	17 180	17.84%
detail	Complete detail (detail, detailed, detailing)	1 238	10.83%	12 002	12.47%
plate	plate (plates) size	1 165	10.20%	7 432	7.72%
weld	Dimension weld (welded, welding, welds)	1 137	9.95%	5 106	5.30%
confirm	Confirm level (confirmation, confirmed, confirmer, confirming)	900	7.88%	7 828	8.13%
Area + connection		2 137	18.70%	25 075	26.02%
Area + connection+ trusses		909	7%	12 533	13.01%

The theme “area” is the most present theme in the RFIs. It is contained in 5 298 out of 11 426 RFIs which represents 46.37% of RFIs. The theme “area” also corresponds to the highest response time of the RFIs, which means 58 689 out of 96 277 or 60.96% of the elapsed days for the RFIs. In the same order, the themes "connection" and "trusses" are the second and third themes most present in the RFIs. They are contained respectively in 26.61% and 21.51% of the RFIs and represent respectively 37.23% and 29.60% of the elapsed time for the RFIs.

Certain themes are present together in certain RFIs. The last two lines of Table 1 contains results for groups of themes. For example, 2 137 RFIs (18.70% of RFIs) contain both "area" and "connection" themes. they account for 25 075 (26.02%) of the elapsed days for responses to the RFIs. Nine hundred and nine RFIs (7% of RFIs) contain "area" "connection" and "trusses" themes and represents 12 533 (13.01%) of RFI response elapsed days. This shows that there are links between themes. Figure 2 shows how the themes are linked in pairs. The top number associated with each link represents the number of RFIs containing both themes and the bottom number the total elapsed time days for those RFIs. Thus, the missing and area theme are contained in 1 066 RFIs and represent 13 367 days elapsed.

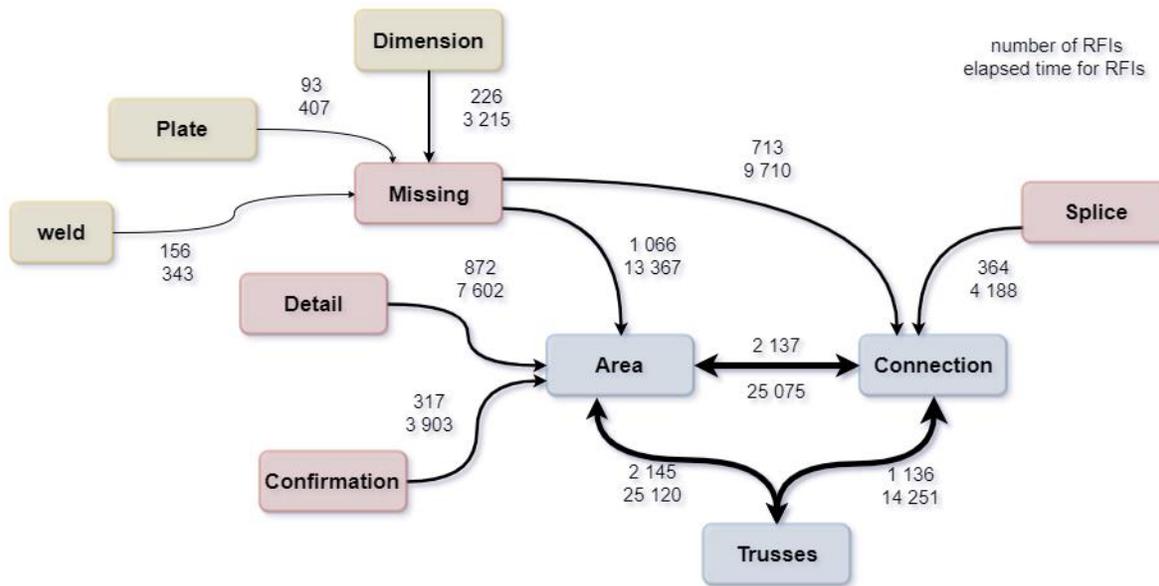


Figure 2: Link between pair of themes

## 5. Observations and recommendations

Observation 1: the theme “area” is present in 43% of RFIs and represents 60.96% of RFI's response time. Area does not refer to a technical problem. It would be used in RFIs to define the geographical location of structural elements. thus, the qualitative analysis of the RFIs can provide information on the geographical location which is object of most RFIs.

Recommendation 1: if the designers identify the characteristics of the most sensitive geographical areas of the RFIs, this observation may allow them to better direct the design efforts according to the zones of similar characteristics within future projects.

Observation 2: “area” is connected to five themes (Figure 2) of which “missing”, “detail”, and “confirmation” refer to a type of information need as well as to “trusses” and “connection” themes, which refer to parts of the steel structure. The theme “area” is then the location used to request information about connections and trusses.

Recommendation 2: the qualitative analysis of RFIs in this case study, can provide information on the types of building elements whose data are not complete at the end of the design phase.

Observation 3: on the basis of Figure 2, the set of RFIs would then express:

- The need for details or confirmation of connections,
- The need for details or confirmation regarding trusses,
- The absence of dimensions on plates or welds belonging to a connection,

Recommendation 3-1: the qualitative analysis of the RFIs can identify, in a precise way, the information that the contractors need. This information can in the long run be the basis of the design rules. These rules can be for this example: to use for connections, symmetrical parts, standard, or even modular parts. This will significantly reduce the detail to be provided or confirmed.

Recommendation 3-2: Recommendation 3-3: better documenting the connections of trusses could reduce the number of RFIs by almost 10% and the global response time of RFIs by almost 15%.

## Conclusion:

This article proposes a systematic method that uses qualitative analysis of the RFI contents to identify the type of information to be considered during the design phase in order to better take into account the information needs of contractors. The analysis of the contents makes it possible to highlight themes present in the RFIs. In the case study in this article, the following topics were identified: “connection” “area” “trusses” “detail” “missing” “confirmation” and “splice”. In the context of steel construction, these themes represent: geographic locations that are most subject to RFIs. These themes also represent the structural elements that are the subject of the RFIs, and these themes also represent the type of information sought (missing, confirmation, detail). Consideration of these themes in the design and documentation of future projects could significantly reduce the quantity of RFIs, the cost, and duration of the project

However, this analysis does not consider synonyms and word contexts in documents. An alternative to this is to continually feed the register of synonyms and contexts, in order to improve the understanding of the contents of the documents studied.

To further this study, it would be possible to analyze several similar projects developed in the same design office to uncover contractor information need currently not provided by designers and eventually establish design rules.

## References

- [1]. Latham Sir M. Constructing the team: Final report of the government/industry review of procurement and contractual arrangements in the UK construction industry. London: HMSO; 1994.
- [2]. Egan J. The Egan report-rethinking construction. Rep Constr Ind task force to deputy prime Minist London. 1998;
- [3]. Tilley PA. Causes, effects and indicators of design and documentation deficiency. Queensland University of Technology, Brisbane; 1998.
- [4]. Evers HGA, Maatje IRF. Cost based engineering and production of steel constructions. In: Steel Design Codes—Fourth International Workshop on Connections in Steel Structures. 2000. p. 14–22.
- [5]. Forgues D, Koskela L. The influence of a collaborative procurement approach using integrated design in construction on project team performance. *Int J Manag Proj Bus* [Internet]. 2009;2[3]:370–85. Available from: <https://www.scopus.com/inward/record.url?eid=2-s2.0-77956273236&partnerID=40&md5=efb21cfc0ad706031202f4cc40d1270%5Cnhttp://www.emeraldinsight.com/doi/pdfplus/10.1108/17538370910971036%5Cnhttp://dx.doi.org/10.1108/17538370910971036>
- [6]. Forgues D, Jordanova I. An IDP-BIM framework for reshaping professional design practices. In: Construction Research Congress 2010: Innovation for Reshaping Construction Practice. 2010. p. 172–82. [https://doi.org/10.1061/41109\(373\)18](https://doi.org/10.1061/41109(373)18)
- [7]. Mohamed S, Tilley PA, Tucker SN. Quantifying the time and cost associated with the request for information [RFI] process in construction. *Int J Constr Inf Technol*. 1999;7(1):35–50.
- [8]. Hughes N, Wells M, Nutter C, Zack J. Impact & control of RFIs on construction projects. *Navig Constr Forum*TM. 2013;
- [9]. Fokwa Soh M, Barbeau D, Dore S, Forgues D. Design rules to improve efficiency in the steel construction industry. 2018;
- [10]. Jeong W, Chang S, Son J, Yi J-S. BIM-integrated construction operation simulation for just-in-time production management. *Sustainability*. 2016;8[11]:1106.
- [11]. Burati Jr JL, Farrington JJ, Ledbetter WB. Causes of quality deviations in design and construction. *J Constr Eng Manag*. 1992;118(1):34–49.
- [12]. Andrews W. RFI recommendations. *Mod Steel Constr*. 2005;46(10):37–41.
- [13]. Sparksman PJ. Quantifying the time and cost associated with the Request for Information (RFI) OR Technical Query (TQ) process-a designers perspective. 2015;
- [14]. Downe-Wamboldt B. Content analysis: method, applications, and issues. *Health Care Women Int*. 1992;13(3):313–21.
- [15]. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today*. 2004;24(2):105–12.
- [16]. Dong A, Hill AW, Agogino AM. A document analysis method for characterizing design team performance. *J Mech Des*. 2004;126(3):378–85.
- [17]. Mabogunje A. Noun phrases as surrogates for measuring early phases of the mechanical design process. In: the Proceedings of the 9th International Conference on Design Theory and Methodology, 1997. 1997.
- [18]. Zhang L, Wen M, Ashuri B. BIM log mining: measuring design productivity. *J Comput Civ Eng*. 2017;32(1):4017071.
- [19]. Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15(9):1277–88.
- [20]. Kondracki NL, Wellman NS, Amundson DR. Content analysis: review of methods and their applications in nutrition education. *J Nutr Educ Behav*. 2002;34(4):224–30.
- [21]. Mayring P, Fenzl T. Qualitative inhaltsanalyse. In: *Handbuch Methoden der empirischen Sozialforschung*. Springer; 2014. p. 543–56.
- [22]. Potter WJ, Levine-Donnerstein D. Rethinking validity and reliability in content analysis. 1999;
- [23]. Han J, Pei J, Kamber M. Data mining: concepts and techniques. Elsevier; 2011.
- [24]. Vijayarani S, Janani R. Text mining: open source tokenization tools—an analysis. *Adv Comput Intell*. 2016;3(1):37–47.
- [25]. Mei Q, Zhai C. Discovering evolutionary theme patterns from text: an exploration of temporal text mining. In: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining. ACM; 2005. p. 198–207.
- [26]. Burns TM. The effect of selected request for information variables on steel fabrication outcomes. Indiana State University; 2007.