Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10294

Revealing axiomatic design relations in patent documents with natural language processing (NLP)

Marco Losanno [1,2,✉], Vito Giordano [1,2], Marco Consoloni [1,2], Filippo Chiarello [1,2] and Gualtiero Fantoni [1,2]

[1] University of Pisa, Italy, [2] Business Engineering for Data Science Lab (B4DS) research group, Italy

✉ marco.losanno@phd.unipi.it

ABSTRACT: Natural Language Processing (NLP) has been widely applied in design, particularly for analyzing technical documents like patents and scientific papers to extract engineering design knowledge. This work aims to enhance this process by integrating the Axiomatic Design methodology with NLP techniques applied to patent texts. The objectives are to (1) extract Functional requirements (FRs) and Design parameters (DPs), and (2) identify how FRs and DPs are related in text (Axiomatic relations). The second objective is particularly challenging due to limited focus on understanding semantic relations in literature, and previous studies often extract Axiomatic relations in an unstructured way. The approach achieves 60% precision for the first objective and 30-50% for the second. Moreover, a case study shows the practical application of this methodology to assist the work of designers.

KEYWORDS: design theory, machine learning, semantic data processing, axiomatic relation, natural language processing (NLP)

1. Introduction

Designers today have access to a vast repository of digitally stored design and manufacturing data from the past. These data could be used to improve the engineering design process. Even with the abundance of information on both successful and failed design data, transforming documentation into actionable insights to inform future design practices remains a formidable challenge across various industries.
Natural Language Processing (NLP) techniques have enabled computer-based systems to process large volumes of unstructured technical information written in natural language. These techniques have been extensively used to extract Engineering Design (ED) knowledge from technical documents such as patents, maintenance reports, assembly instructions and process manuals. For example, Cascini et al. (2007) extract components and identify their abstraction levels through an automatic functional analysis of the patent text. Similarly, Fantoni et al. (2013) extract information about functions, physical behaviours and states of devices directly from the patent documents. Puccetti et al. (2023) explore the use the NLP techniques with the purpose to identify the technologies mentioned in patents' text and automatically perform a technology landscape analysis. Sarica et al. (2023) introduced a methodology that leverages large-scale multidisciplinary design knowledge bases to automatically generate semantic networks of entities and relationships based on textual design descriptions.
Despite these advancements, extracting relationships between ED concepts remains a complex and underexplored task. Some studies have attempted to extract ED relationships, but often without defining the types of relationships between entities (Sarica et al., 2023). Other studies focus on grammatical or lexical relationships (Guo et al., 2016), instead of relationships more specific to engineering design. The challenge of defining and extracting meaningful ED relationships remains largely unaddressed.
Giordano et al. (2024) addressed the challenges in extracting ED relationships and provided a set of guidelines for this task. These include: (1) defining the concepts and relationships to be extracted, (2) utilizing Named Entity Recognition (NER) to identify concepts, (3) evaluating the performance of

concept extraction, (4) extracting relationships between concepts, and (5) assessing the performance of relationship extraction. Following these guidelines, this paper develops and implements an approach for automating the extraction of Axiomatic relationships from patent texts.

This work aims to use NLP techniques to identify in the text of patents the main concepts related to Axiomatic Design Theory. These concepts are: (1) Functional requirements, what a system or product must do, (2) Design parameters, parameters which characterize the design of a system or product, (3) Axiomatic relations, relations between Functional requirements and Design Parameters. The approach utilizes NER with rule-based techniques to extract Functional requirements and Design parameters systematically and to map the Axiomatic relations between them. A case study in the field of protective face masks was defined, demonstrating the application of this approach to assist designers in their work. Specifically, the approach helps the designers to find the fundamental or alternative functionalities (FRs) of a system or product and understand how these can be satisfied by corresponding DPs.

2. Axiomatic design

2.1. Axiomatic design theory

In this section, we outline the definition of the main concepts related to Axiomatic design (AXD), which includes Functional Requirements (FRs), Design Parameters (DPs), and the Axiomatic relation.

AXD is a system design methodology developed by MIT professor Nam P. Suh in the 1980s that provides a systematic and scientific basis for making design decisions. This technique uses matrix methods to analyse the transformation of customer needs into FRs and DPs for defining the design of a system or product.

The primary objective of AXD is to reduce the trial-and-error nature of the design process by offering a systematic, scientific framework for translating FRs into DPs. For those familiar with AXD, this translation is mathematically expressed as FRs = [A] × DPs, where [A] represents the design matrix that captures the relations between FRs and DPs. The approach ensures that, once customer needs are articulated as FRs, designers can identify the optimal combination of DPs to fulfil all specified FRs.

Two fundamental axioms guide the AXD process for developing effective designs. Firstly, the Independence Axiom states that a good design maintains FR independence, meaning each FR is satisfied by a unique DP. This results in an uncoupled design. Secondly, the Information Axiom states that among multiple uncoupled designs, the optimal one minimizes information content, ensuring a higher probability that DPs successfully fulfil the FRs. A notable study by Ullah (2005) explored a numerical approach to evaluating these axioms using linguistic information. By following these principles, AXD improves the design process, leading to more reliable and optimized solutions.

To fully understand AXD and how we can use NLP to extract the information related to it in text, it is important to have a clear definition of FRs and DPs. In literature, there are different definitions of FR:

- **IEEE (1990):** *"a function that a system (. . .) must be able to perform."*
- **Suh (1998):** "*are a minimum set of independent requirements that completely characterizes the functional needs of the product (or software, organizations, systems, etc.) in the functional domain."*
- **Shankar et al. (2020):** *"is what a product must do, be able to perform, or should do. In each of these, functional requirements describe action and doing. Others have slightly refined this view by positioning that the FRs are the desired and resulting behaviors of the system."*

Building on these definitions, we have defined a FR as an action required to be performed by a system, product or product component (e.g. *"a helmet must* **protect a user***"*).

For what concern DP, the following definitions were individuated:

- **Otto and Antonsson (1993):** *"design parameters are those for which the designer selects values as a part of the design process."*
- **Suh (1998):** *"are the key physical variables (or other equivalent terms in the case of software design, etc.) in the physical domain that characterize the design that satisfies the specified FRs."*
- **Saglar Onay et al. (2020):** *"A variable that is evaluated during the design process determining the characteristics of the final object or system."*

A DP is essentially a physical quantity that characterizes the design of a system, product or product component (e.g. *"the* **helmet thickness***"*).
The relation between FR and DP, that is expressed in the AXD matrix, is called Axiomatic relations. We defined the Axiomatic relation as the relation that exists when a DP influences a FR (e.g. *"the* **helmet thickness** *must be enough to* **protect the user***"*).

2.2. NLP for axiomatic design

There are different NLP approaches used in literature for extracting AXD concepts. In the early stages, different NER techniques, which are a subset of NLP, were used. These techniques relied on three main methods: gazetteer-based, which uses predefined lists of entities to match them in text; rule-based, which applies specific rules and patterns to identify entities; and machine learning-based, which uses machine learning algorithms to detect entities based on context.
Focusing on NLP systems to extract AXD information, Chen et al. (2008) developed a NER rule-based retrieval system that extracted FRs from product specifications using a "Verb-Noun" linguistic structure. A similar methodology was employed by Liu and Tate (2010) which applied grammatical rules based on the Subject-Verb-Object (SVO) structure to identify FRs within patents. These approaches extract design information in a good, structured way, but the effectiveness is heavily dependent on the authors' writing style and grammar. However, they are simple, computational inexpensive and quite precise.
With the advent of Large Language Models (LLMs), more advanced techniques have been explored. These AI-powered LLMs uses deep learning architecture and enormous amount of textual data to perform different tasks in NLP. Akay and Kim (2021) introduce a recursive decomposition process of the text to extract a hierarchical structure of FRs and DPs using a question-answering system of BERT. Then, Akay et al. (2023) integrated an LLM with prompting to develop a chatbot designed for extracting FRs and DPs from design specifications, called Virtual Design Assistant. Similarly, Kwon et al. (2024) employed a BERT-based question-answering model to extract design requirements from technical specifications. These approaches use the high generative capabilities of LLMs to extract FRs and DPs from text, but they are unstructured and lack linguistic modeling of AXD concepts and quantitative evaluation of how LLM-generated outputs aligned with AXD theory. Additionally, these methods are computationally expensive and require large, high-quality datasets for systematic training and testing. For this reason, in this study we have used a simple, structured method with NER techniques to support the work of the designers.

3. Methodology

The methodology proposed in this work is composed of the following three phases: *1) Data Preprocessing*, where patent texts were segmented into sentences in order to facilitate the extraction of AXD concepts; *2) Identification of Functional requirement and Design parameter*, in which FRs and DPs were recognized within the patent sentences using a NER system; and *3) Identification of Axiomatic relations*, where sentences containing FRs and DPs were analysed to determine the Axiomatic relations.

3.1. Data preprocessing

The first step followed in this study was the preprocessing of the patent text used. Initially, the title, abstract, claims, and description of the patents, were selected as the textual data. Textual data was then segmented into sentences to simplify the identification of Axiomatic relations between FRs and DPs. In fact, extracting the relations between concepts (i.e., FRs and DPs) across different sentences is challenging, because the context often changes from one sentence to the next. Consequently, the Part-of-Speech (POS) tagging was applied for extracting the sentence structures necessary to create the rule patterns for the FRs and DPs extraction. For this task the Python package SpaCy was used.

3.2. Identification of functional requirements and design parameters

Obtained the dataset with the sentences tagged with the POS, the identification of FRs and DPs was performed using a NER rule-based technique. The extraction approach involved defining rule patterns with curated lists of key elements to identify FRs and DPs. For the extraction of FRs, which represent actions performed by systems or products, functional verbs were used. We have defined functional verbs as the verbs that represent functions carried out by systems or products (e.g. "store", "move"). These types of verbs were selected for their likelihood to capture actions related to systems or products. For DPs, which refer to physical variables, a list of general physical quantities (e.g. “diameter”, “length”) was employed to locate specific DPs. Hence, for the first step we have created two list of functional verbs (to extract FRs) and physical quantities (to extract DPs):

- **List of functional verbs:** initially, 203 functional verbs were extracted from the work of Hirtz et al. (2002). These verbs were then expanded to include third-person conjugations in the present simple, past simple, and present continuous tenses, resulting in a total of 812 verbs.
- **List of parameters:** contains physical quantities obtained initially from Wikipedia[1]. We have used web scraping techniques, resulting in a list of 105 parameters. This list was cleaned, refined and then expanded. Finally, the parameters were pluralized, obtaining 340 parameters.

Created the two lists we have defined rules pattern for extracting FRs and DPs. To develop FR patterns, initially, a basic format of *“Verb + Noun”* was defined, as suggested by Chen et al. (2008), with a functional verb as *“Verb”*. Then, this pattern was refined: below we can see the final rules.

1. Verb (*Functional verb* at infinitive or present 3rd person) + Article (0 or 1) + Adverb or Adjective (0 or more) + Noun (1 or more)
2. Verb (*Functional verb* at continuous or past simple) + Article + Adverb or Adjective (0 or more) + Noun (1 or more)

Two rules were defined based on the type of functional verb conjugation used. This because the POS-tagging misidentifies adjectives as verb conjugations in past simple and continuous forms (e.g. “dispersing substances”, “integrated circuit”). Hence, the inclusion of an article between such verbs and nouns helps to handle the error. Additionally, we did not define rules for extracting FRs in the passive form. This decision was made because entities extracted in the passive form typically represent states rather than actions (e.g., “agitator is connected”), and filtering them would be an overly complex task. For developing DP patterns, a simpler pattern of just *“Noun”* was identified, where the *“Noun”* denoted a parameter. Subsequently, this pattern was analysed and refined. The final rule is described below.

1. Noun (0 or more) + Noun (*Parameter* at singular or plural) + Not Noun

The rule defined extracts DPs only when there is no subsequent noun. This adjustment was defined because parameters often are used as adjectives describing components or products (e.g. “flow sensor”, “pressure sensor”). To mitigate this type of error, *“Not Noun”* has been added to the pattern. Additionally, the nouns before the parameters were also added to also extract the context of the physical quantities (e.g. “wheel diameter”, “nozzle width”).
The rules defined for extracting FRs and DPs, were analysed and refined with an iterative process of performance evaluation. A sample of 600 sentences, annotated with the entities identified by the defined rules, was manually checked to calculate precision and recall metrics (Tsai et al., 2006). In literature, the results of several NLP tasks have traditionally been evaluated using human subjects or using previously annotated data (Lee et al., 2020). In the AXD domain, there is a lack of annotated documents. For these reasons, we rely on human evaluation in a manner similar to Giordano et al. (2021). The performance obtained using FR rules defined is 65.10% of precision and 34.73% of recall. Instead, the performance obtained with DP rules is 62.72% of precision and 85.77% of recall. We can see a low recall for the extraction of FRs and this is derived from three principal causes: 1) The use of an article in the first FR extraction rule to limit POS-tagging errors, which reduces the number of FRs extracted; 2) The use of a limited list of verbs restricts the ability to capture a diverse range of FR expressions, as only those containing functional verbs listed in the gazetteer can be extracted; 3) FRs can be expressed in multiple ways, not exclusively through verbs. For instance, the sentence *“the controller allow the movement of the robot arms”* can be rephrased as *“the controller moves the robot arms”*, where the function performed by the controller is *“move”* instead of *“movement”*. As we can see, FRs can be written as noun phrase *“movement of the robot arms”* and not only, but our method cannot extract FRs which are not written as verb. The low recall rate does not significantly impact our research, because this work aims to find in a simple way the fundamental correct FRs of a system, not all of them.

[1]https://en.wikipedia.org/w/index.php?title=List_of_physical_quantities&oldid=1215749861

3.3. Identification of axiomatic relations

Extracted the FRs and DPs with the rules defined, we have analysed how they can be Axiomatic related in patent sentences. For identifying Axiomatic relations, we have created a list of textual forms, referred to as Axiomatic Pointers (AXP), which explicitly indicate how a DP affect a specific FR. An AXP is a word or group of words that show the influence of a DP on an FR in the text and is situated between the two entities in the sentence. For instance, in the sentence *"the motor control the valve based on the pressure level"*, the phrase *"control the valve"* define a FR that is affected by *"pressure level"* and it is explicated by *"based on"* which is the AXP.

The list of AXPs was defined evaluating the sentences contained FRs and DPs with three types of analysis: (A) Distance analysis, to understand the typical distance or proximity between the FRs and DPs, (B) Inspection analysis, to manually check recurring patterns or syntactic structures indicative of AXP, and (C) a Word frequency analysis, to find linguistic forms (unigram, bigram, trigram) and conventions prevalent for expressing an AXP which may not have been identified during the inspection analysis. Performing the Distance analysis we have seen that the median of the number of words between FR and DP in sentences was 18, and 25 if DP precedes FR. These values were used to for the extraction of Axiomatic relations with the AXPs, because a lot of distance from the FRs and DPs reduced a lot the probability of having an Axiomatic relation. With the other two analysis we have created the list of AXP. The list of AXPs contain 176 textual forms which can find potential Axiomatic relations. These defined pointers were classified based on two factors:

1. **Type of relation:** specify how the DP influence the FR. This factor was defined to aid in understanding how the AXP relates the FR and DP. We have defined three types of AXP (1) *Need*: when the DP is necessary for the FR. For instance, *"a larger tube diameter is* **needed** *to flow the liquid"*; (2) *Allow*: when the DP allow or facilitate the FR. For instance, *"the rotation of the wheels* **permit** *to move the object"*; (3) *Cause*: when the DP cause or affect the FR. For instance, *"the machine control the speed* **due to** *the control signal received"*.
2. **Type of form:** specify the position of FR and DP in the sentence. This factor classifies in which type of phrases the AXP can be used to identify an Axiomatic relation. We have two categories of AXP: (1) *DP before*: when the DP is before the FR in the sentence. For instance, *"the force must be* **sufficient** *to open the valve"; (2) FR before*: when the FR is before the DP in the sentence. For instance, *"to open the valve* **according** *to the pressure level"*.

Created the list we have evaluated the precision of each AXP in finding Axiomatic relations. The precision was estimated by evaluating up to 100 sentences for each extracted. The recall was not assessed, as the focus was on determining the ability of each AXP in identifying correctly Axiomatic relations. The results of this evaluation are presented in Table 1.

Table 1. Type of AXPs with precision estimate

AXP category	Precision category	No. AXP	Precision AXPs
Need (DP before)	53.27%	30	selected (54%), required (68%), sufficient (66%) ...
Allow (DP before)	53.08%	34	used (55%), order (60%), applied (45%) ...
Cause (FR before)	46.54%	24	based on (52%), according to (55%), corresponding to (33%) ...
Cause (DP before)	36.12%	43	controlled (52%), cause (49%), causes (32%) ...
Need (FR before)	28.38%	22	need (16%), require (48%), requires (47%) ...
Allow (FR before)	11.36%	23	supported by (9%), maintained by (0%), secured by (0%) ...

Table 1 presents detailed information about the different AXPs and we defined above. The first column lists the AXP categories, while the second provides the precision estimates for capturing the Axiomatic relations within each category. The third column indicates the number of AXPs in each category, and the final column contains the precision estimates for individual AXPs within the respective categories.

The data indicates that the overall precision in identifying Axiomatic relations is not particularly high. In general, a low precision in extracting Axiomatic relations derived from two factors. The first is the

distance between the FRs, DPs, and the AXP. The greater the distance, the lower the probability of identifying an Axiomatic relation because the sense of the relation may change. The second factor is punctuation between the entities. The punctuation marks between FRs and DPs reduces the probability of identifying an Axiomatic relation because the context may change. However, if we analyse the precision of the specific AXP some of them have a high precision, such as *"required"* or *"sufficient"*. For instance, in different cases *"sufficient"* is used in this way: *"[. . .] a rigidity to the first transparent shield 26 sufficient to protect the face of the wearer from flying debris and fluids." (US8291512B2)* where the FR is *"protect the face"*, which is Axiomatic related with the DP *"rigidity"*.
Using the defined list of AXPs, we extracted sentences with FRs, DPs, and the AXPs connecting them.

4. Results

4.1. Case study: protective face mask

The data used for this study is a set of patents related to protective face mask. It was collected a set of 34,658 patents classified under IPC code A41D13/11 (Protective Face Masks). This data was obtained thanks to the research-driven company Errequadro[2] The decision of using patents on protective face masks was defined because it is a simple system where the design is it not too complex, making the explanation of our NLP system easier. From these patents were applied the data preprocessing explained in Section 3.1 and were obtained 2,185,616 sentences.

4.2. Functional requirements and design parameters identified

Obtained the sentences from the protective face mask patents, the extraction of FRs and DPs was performed following the methodology explained in Section 3.2. The extraction of FRs and DPs resulted in 964,113 (44.2%) of sentences containing at least one type of entities. As we can see, more than 50% of the sentences not contain AXD concepts. This observation suggests that many patent sentences focus on describing the invention's structure rather than its functions. However, it may also reflect limitations associated with using the gazetteer. Additional results from the extraction process are showed below.

Table 2. Type of FRs and DPs extracted

AXD Concept	Entity type	No. Entities	Examples of entities
FR	FR at infinitive	384,100	protect the face, fix the mask, maintain the shape
	FR at third person of present	188,305	absorbs the droplets, reduces discomfort, supports a user
	FR at continuous	158,713	filtering the air, protecting the respiratory tract, providing a barrier
	FR at past	26,442	reduced the load, adjusted the position, regulated the flow
DP	DP at singular form	1,266,950	filtering power, thickness, sterilization power
	DP at plural form	103,494	carbon dioxide levels, head sizes, air flow rates

Table 2 shows the types of FRs and DPs extracted. The first column lists the different categories of entities, with FRs categorized by verb conjugations and DPs by their singular or plural forms. The second column indicates the number of entities extracted, while the third provides examples.
The data indicate that 757,560 FRs and 1,370,444 DPs were extracted in total. This observation confirms that patent documents primarily emphasize structural descriptions over functional details. Additionally, Table 2 shows that the infinitive form is the most frequently used verb conjugation for expressing FRs. This is followed by the third person in the present simple and continuous tenses, while the past tense is rarely employed. Instead, for DPs, the singular form is the mostly used. This finding is consistent with the

[2]https://www.errequadro.ai/

Table 3. The most frequent functional verbs and parameters used for FRs and DPs

Functional verb	No. in Patents	Parameter	No. in Patents
provide	108,581	position	116,670
form	63,124	area	115,335
prevent	58,002	direction	94,523
reduce	31,850	time	89,930
adjust	26,821	length	69,506
protect	21,672	space	54,319
increase	21,008	size	46,224
cut	20,314	temperature	43,036
limit	19,003	pressure	37,212

nature of design specifications, which often focus on individual components or attributes of a system, rather than collective groups.

Table 3 shows the most frequently used functional verbs and parameters identified in the extraction of FRs and DPs within this domain of patents. The first two columns list the functional verbs and their corresponding frequencies, while the last two columns present the parameters and their counts.

As we can see, the most functional verb used is "provide". However, it is not always used to describe as FR but depend on the context. For instance, in the sentence "*use of multiple detectors may provide redundancy in the event a detector 670 is damaged*." *(US10420386B1)*, the verb *"provide"* does not denote a specific function of a system, product, or component. Similarly, contextual errors can also occur with parameters. For example, in the sentence *"[. . .] clean the hands from cross contamination and return to the on-going work area" (US10960238B1)*, the parameter *"area"* does not refer to a measurable physical quantity of a system or product.

These contextual errors are heavily influenced by the nature of the text. In domains like patents, which discuss systems or products, such errors are less frequent. In fact, as we see in Section 3.2 the performance of the extraction method shows that functional verbs and parameters are generally quite precise in capture FRs and DPs. Moreover, the verbs and parameters listed in Table 3 are particularly relevant to the context of protective face masks. Many of the verbs describe functions related to protection (e.g. "protect", "reduce", "prevent") or aspects of wearing (e.g. "adjust", "fix). For instance, in the sentence, *"head-and-face covers may also be used to block heat radiation to protect the skin of a wearer" (US20050160514A1)*, the verb *"protect"* directly identifies the FR *"protect the skin"*. Similarly, the frequent parameters often describe the physical dimensions or measurable properties of the product (e.g. "position", "size", "length"). For instance, in the sentence, *"further, the tie strap position may slip or change once the mask is installed, again compromising an effective seal" (US5699791A)*, the parameter *"position"* specifies the DP *"strap position"*.

4.3. Axiomatic relations identified

Identified the FRs and DPs from the protective face mask patents, we have extracted the Axiomatic relations following the methodology explained in Section 3.3. The total number of Axiomatic relations extracted is 15,818 on 235,156 (10.8%) sentences which contain FRs and DPs. Additionally, we can see in the Table 4 the more frequent type of Axiomatic relations extracted, classified with the AXPs.

Table 4. Axiomatic relations extracted

Relation Category	No. Relations	No. Relations for AXPs	Examples
Cause (FR before)	6,490	according to (3,968), due to (831), based on (585), caused by (375), corresponding to (218), depending on (163)	This disposable medical surgical mask convenient to **adjust elasticity** adjusts *according to* **face width** and ear distance of oneself, alleviates the pulling force of ear, is favorable to the user to use.

(*Continued*)

Table 4. Continued.

Relation Category	No. Relations	No. Relations for AXPs	Examples
Allow (DP before)	4,949	used (2,579), order (1,091), applied (329), make (178), makes (134)	The **diameter** is 5 mm to 10 mm, preferably approx. 7 mm, in *order* to **enable relaxed breathing**.
Need (DP before)	2,513	configured (944), required (307), sufficient (246), enough (165), necessary (161)	Additionally, the seal 80 may reduce the **force** *required to* **hold the respirator** 10 to the skin of the user.
Cause (DP before)	1,306	adjusted (400), control (400), controlled (107), cause (69), adjust (64)	A space can be formed, and the **air flow** in the mask can be appropriately *controlled* to **prevent air leakage** from the mask.

Table 4 consists of four columns, each detailing specific aspects of the extracted Axiomatic relations. The first column categorizes the type of Axiomatic relations based on the type of AXPs used. The second column provides the total number of Axiomatic relations identified. The third column breaks down the number of relations extracted for each type of most used AXPs, while the fourth column offers examples of the extracted relations, with **FRs** and **DPs** highlighted in bold and *AXPs* italicized.
The Table 4 shows that the most frequent Axiomatic relations extracted typically use AXPs of Cause category, where the FR precedes DP in the sentence. Within this category, *"according to"* emerges as the most utilized AXP, accounting for a significant portion of the identified relations. We can see from the example of Table 4 how typically *"according to"* is used to express an Axiomatic relation, where the elasticity of medical surgical mask is adjusted in relation to the face width and ear distance.
Another frequent category of AXP is Allow, where the DP precedes FR in the sentence. This category often employs terms like *"used"* and *"order"*, indicating DPs that enable or facilitate specific FRs. Similarly, the Need category highlights the necessary DPs for fulfilling the FRs, with common AXPs such as *"configured"* and *"required"*. A smaller but significant subset of Axiomatic relations appears in the Cause category, where DPs precede FRs. Overall, the findings underscore the reliance on causal and facilitative relations in describing design features, especially in the Cause and Allow categories.

4.4. Product design identification

Extracted the Axiomatic relations, we have analysed them to find design information of the protective mask for supporting the designer work. We can see below a simple approach to assist the designer work. Table 5 shows the FRs arranged by frequency, along with all Axiomatic related DPs for each FR. Additionally, each DPs related contains the number of Axiomatic relations with the FR. This Table can show different important information to designer. First, we can see that some of the most frequent FRs

Table 5. FRs Axiomatic related with DPs

FR	DP
adjust the length	size (95), length (64), comfort level (8), direction (8), strap length (8), face size (7), head size (6)
changes color	temperature (88), body temperature (50), heat (6), pressure (5), area (4), replacement time (4), color change temperature range (3)
adjust the position	size (21), distance (5), assembly line direction (4), length (4), width (4), direction (3)
adjusting the tension	tension (48), power (14), amount (9), number (1), times (1)
adjust the size	position (7), face size (5), angles (3), clearance (2), length (2), positions (2), speed (2)
filtering the air	nose position (13), position (10), mouth position (7), oronasal position (5), area (3), size (3), density (1)
prevent infection	wind speed (7), number (4), position (4), time (4), diameter (3), contact degree (2), diameters (2)
maintain the shape	direction (4), length (3), width direction (3), flexibility (2), heat (2), degree (1), density (1)

may indicate important functions that the patent invention must fulfil ("adjust the length", "filtering the air"). While not all FRs may be correct, but some of them can be helpful in identifying the product or system design. Additionally, designers can analyse specific frequent FRs to uncover innovative or uncommon functions. For instance, searching the FR *"changes color"* led to the discovery of a particular function in a facial mask *"a facial mask having impregnated therein a first thermochromatic material, suitable to change colors at a predetermined temperature." (US20190125011A1)*. Such analyses can guide designers in discovering alternative or additional functionalities, enhancing and supporting the design process.

The designers can analyse all related DPs to identify those that influence the FRs. In this way, they could identify DPs that are unexpected and useful to characterize the design of the product. For instance, the FR *"filtering the air"* is Axiomatic related to the DP *"size"* and we found *"[. . .] suction fan 500 used in the mask 100 has a problem that the mask body 200 needs to be very large in size because it is used for filtering the air [. . .]" (KR200489651Y1)*. This sentence demonstrates how the size of the mask is important in the air-filtering function, as it must accommodate the suction fan. Such insights can help designers better understand the design characteristics.

However, some of the DPs in Table 5 are generic and lack contextual specificity, which can complicate the analysis of Axiomatic relations. For instance, DPs like *"size"* may seem ambiguous without additional context. Additionally, some of DPs are correct and some of them are not. This because some are parameter and not DPs. For instance, in the sentence *"the body temperature sensor 150 may be configured to change color according to the body temperature of the wearer." (KR102515012B1)* the FR is *"change color"* that is affected by *"body temperature"* which is the DP extracted and *"according to"* is the AXP. However, *"body temperature"* is not a DP because it is a physical quantity but is not a parameter alterable of the product design. Our extraction method cannot distinguish the DPs and parameter because it depends on the context of the sentence. For instance, *"size"* is correctly identified as a DP in the sentence *"[. . .] the filter 914 may have a variety of shapes and/or sizes, each designed to filter air breathed in through a nose or mouth of the wearer" (US20210345694A1)*, but it is a parameter in the sentence *"[. . .] to provide comfort to the user depending on the shape and size of their head" (US20220000209A1)*. Despite these challenges, identifying parameters alongside DPs remains valuable for understanding design constraints. Recognizing these constraints helps refine the design process, ensuring all relevant factors are considered, leading to more robust and effective designs.

5. Conclusions and future developments

This section discusses the main results obtained in this work and explains the future steps. The approach applied for extracting FRs and DPs have good performance on precision but have several limitations. One of this is the low recall of FRs extraction. This is caused by the limited rule coverage using the gazetteer. Another is caused by the POS-tagging that not tag always correctly, particularly with past simple/continuous verbs being mistaken for adjectives. Additionally, most of DPs extracted are correct but some of them are generic and not contextualised (e.g. "value", "number"). This is not a problem for the extraction, but it could be a problem for the analysis of Axiomatic relations. To address these issues, future improvements include developing additional rules for FR extraction, exploring alternative POS-tagging models, and refine the extraction rules to provide more context for DPs. Additionally, expanding and refining the gazetteer with different sources, such as the parameters established from the Theory of Inventive Problem Solving (TRIZ), could improve accuracy and coverage.

The approach applied to analyse and identify Axiomatic relations has found interesting finding. The textual relation between FRs and DPs can be identified with specific rules, as evidenced in the case study. However, the primary causes of uncorrected identification of an Axiomatic relation between FR and DP individuated are: (1) too much distance between the entities that can alter the sense of relation, (2) the punctuation that can change the context. The mitigation strategies that could be employed are the optimization of entity distance in the rules used, standardizing punctuation usage.

This work presents a rule-based NER system using a gazetteer to extract FRs, DPs, and Axiomatic relations from patent texts, assisting designers in understanding design characteristics and identifying innovative opportunities. Future work could focus on improving the method, developing a machine learning model to detect Axiomatic relations, and exploring negative Axiomatic relations that can identify design problems. Additionally, we plan to further investigate how our approach aligns with the fundamental axioms of Axiomatic Design. In conclusion, this work contributes to understanding design information in patent text and serves as a foundation for future research.

Acknowledgments

This research has been partly funded by PNRR - M4C2 - Investimento 1.3, Partenariato Esteso PE00000013 - "FAIR - Future Artificial Intelligence Research" - Spoke 1 "Human-centered AI", funded by the European Commission under the NextGeneration EU program and by the DETAILLs Project (DEsign Tools of Artificial Intelligence in Sustainability Living LabS) - European Union. Erasmus + KA2 - Cooperation partnership in higher education (Project Number: 2023-1-IT02-KA220-HED-000158755).

References

Akay, H., & Kim, S. G. (2021). Reading functional requirements using machine learning-based language processing. *CIRP Annals*, 70 (1), 139–142. https://doi.org/10.1016/j.cirp.2021.04.021

Akay, H., Lee, S. H., & Kim, S. G. (2023). Push-pull digital thread for digital transformation of manufacturing systems. *CIRP annals*, 72 (1), 401–404. https://doi.org/10.1016/j.cirp.2023.03.023

Cascini, G., Russo, D., & Zini, M. (2007, October). Computer-aided patent analysis: finding invention peculiarities. In *Trends in Computer Aided Innovation: Second IFIP Working Conference on Computer Aided Innovation*, October 8–9 2007, Michigan, USA (pp. 167–178). Boston, MA: Springer US.

Chen, Y. J., Chen, Y. M., Chu, H. C., & Kao, H. Y. (2008). On technology for functional requirement-based reference design retrieval in engineering knowledge management. *Decision Support Systems*, 44 (4), 798–816. https://doi.org/10.1016/j.dss.2007.10.003

Fantoni, G., Apreda, R., Dell'Orletta, F., & Monge, M. (2013). Automatic extraction of function–behaviour–state information from patents. *Advanced Engineering Informatics*, 27 (3), 317–334.

Giordano, V., Chiarello, F., Melluso, N., Fantoni, G., & Bonaccorsi, A. (2021). Text and dynamic network analysis for measuring technological convergence: A case study on defense patent data. *IEEE Transactions on Engineering Management*, 70 (4), 1490–1503.

Giordano, V., Consoloni, M., Chiarello, F., & Fantoni, G. (2024). Towards the extraction of semantic relations in design with natural language processing. *Proceedings of the Design Society*, 4, 2059–2068.

Guo, J., Wang, X., Li, Q., & Zhu, D. (2016). Subject–action–object-based morphology analysis for determining the direction of technological change. *Technological Forecasting and Social Change*, 105, 27–40.

Hirtz, J., Stone, R. B., McAdams, D. A., Szykman, S., & Wood, K. L. (2002). A functional basis for engineering design: Reconciling and evolving previous efforts. *Research in Engineering Design*, 13 (2), 65–82. https://doi.org/10.1007/s00163-001-0008-3

IEEE. (1990). IEEE Standard Glossary of Software Engineering Terminology. *IEEE Std 610.12-1990*, 1–84. https://doi.org/10.1109/IEEESTD.1990.101064

Kwon, B., Kim, J., Lee, H., Suh, H. W., & Mun, D. (2024). Construction of design requirements knowledgebase from unstructured design guidelines using natural language processing. *Computers in Industry*, 159, 104100. https://doi.org/10.1016/j.compind.2024.104100

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36 (4), 1234–1240.

Li, Z., & Tate, D. (2010, January). Automatic function interpretation: Using natural language processing on patents to understand design purposes. *Proceedings of the ASME 2010 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 44137, pp. 443–452). American Society of Mechanical Engineers. https://doi.org/10.1115/DETC2010-29097

Otto, K. N., and Antonsson, E. K. (March 1, 1993). "Tuning Parameters in Engineering Design." ASME. *J. Mech. Des. March 1993*; 115 (1): 14–19. https://doi.org/10.1115/1.2919311

Puccetti, G., Giordano, V., Spada, I., Chiarello, F., & Fantoni, G. (2023). Technology identification from patent texts: A novel named entity recognition method. *Technological Forecasting and Social Change*, 186, 122160.

Saglar Onay, N., Garip, S. B., & Garip, E. (Eds.). (2020). Re-Coding Homes: A Mass Customization Tool to Create Flexibility for Housing Units. In *Advances in Civil and Industrial Engineering* (pp. 1–19). IGI Global. https://doi.org/10.4018/978-1-5225-8958-7.ch001

Sarica, S., Han, J., & Luo, J. (2023). Design representation as semantic networks. *Computers in Industry*, 144, 103791.

Shankar, P., Morkos, B., Yadav, D., & Summers, J. D. (2020). Towards the formalization of non-functional requirements in conceptual design. *Research in Engineering Design*, 31 (4), 449–469. https://doi.org/10.1007/s00163-020-00345-6

Suh, N. P. (2001). Axiomatic design: advances and applications (Vol. 4). New York: Oxford university press. https://www.researchgate.net/publication/242487493_Axiomatic_Design_Advances_and_Applications

Tsai, R. T.-H., Wu, S.-H., Chou, W.-C., Lin, Y.-C., He, D., Hsiang, J., Sung, T.-Y., & Hsu, W.-L. (2006). Various criteria in the evaluation of biomedical named entity recognition. *BMC Bioinformatics*, 7, 92. https://doi.org/10.1186/1471-2105-7-92

Ullah, A. S. (2005). Handling design perceptions: an axiomatic design perspective. *Research in Engineering Design*, 16, 109–117 https://en.wikipedia.org/w/index.php?title=List_of_physical_quantities&oldid=1215749861