Journal of Computational Design and Engineering, 2020, 0(0), 1–11

doi: 10.1093/jcde/qwaa092
Journal homepage: www.jcde.org

RESEARCH ARTICLE

A voice-based annotation system for collaborative computer-aided design

Raquel Plumed[1], Carmen González-Lluch[1], David Pérez-López[2], Manuel Contero[2] and Jorge D. Camba[3,*]

[1]Department of Mechanical Engineering and Construction, Universitat Jaume I, 12071 Castellón, Spain; [2]I3B, Universitat Politècnica de València, 46022 Valencia, Spain and [3]Department of Computer Graphics Technology, Purdue University, West Lafayette, IN, 4907 USA

*Corresponding author. E-mail: jdorribo@purdue.edu

Abstract

In this paper, we describe a voice-based interaction mechanism to annotate 3D models directly from a computer-aided design (CAD) modeling environment. The audio signal is captured and automatically transcribed to a textual 3D note, which is attached to the geometry and made available to other product information and business processes across the enterprise via a product data management system. Our approach provides a more natural and intuitive method to capture design and engineering knowledge that is particularly effective when large amounts of information need to be communicated. We discuss the rationale of the software architecture and the value of this modality for capturing knowledge in a collaborative engineering context. Finally, we examine the results of an experiment to validate our proposal. Our results show that 3D annotations are an effective mechanism to communicate design knowledge, which suggests the need for further developments in the areas of multimodal interaction methods and interfaces for CAD and collaborative tools.

Keywords: voice annotations; computer-aided design; design documentation; collaborative design

1. Introduction

Engineering design and product development processes involve the creation, consumption, and exchange of large, diverse amounts of information. Managing this information is critical to ensure data integration and connectivity, improve collaboration and process efficiency, increase productivity and scalability, and reduce risks and potential sources of error. However, much of the information generated and consumed during a product's lifecycle is unstructured in form. For example, many design decisions are communicated verbally through informal conversations, discussions, and sharing of ideas and opinions. Even during formal meetings and collaborative sessions, the information that is recorded and documented does not always capture the full extent of the interactions that occur during these sessions, and thus part of the tacit knowledge and context may be lost. Capturing and processing this type of information (especially the relationships among different items of information) are challenging and time consuming, and thus, not cost effective (Anerousis & Panagos, 2002).

Mechanisms to manage explicit and tacit knowledge during business processes typically follow a three-stage approach: knowledge creation, knowledge retention, and knowledge reuse (Levallet & Chan, 2016). In product engineering, design rationale capture, retrieval, and feedback are an example of this approach. Design rationale is an essential aspect of product knowledge that aims to represent information about the reasoning, motivation, and justification for design decisions as well as describe the relationships to other decisions (Szykman, Sriram, & Regli,

Received: 7 October 2020; **Revised:** 15 November 2020; **Accepted:** 26 November 2020

2001). As a particular aspect of design rationale, design intent is generally used in the context of the virtual representation of the product to refer to the modeling decisions that were made to build the model, which determine how the model will react to changes and alterations (Otey, Company, Contero, & Camba, 2018). From a research standpoint, design intent has received comparatively less attention than design rationale, as it is generally assumed that design intent in communicated implicitly by the parametric structure of the computer-aided design (CAD) model.

Although several methods have been developed to capture and record design intent and rationale (Otey *et al.*, 2018), practical implementations and success stories in industrial settings are rare (Bracewell, Ahmed, & Wallace, 2004; Camba, Contero, & Johnson, 2014). Circumstantial evidence suggests that lack of incentives and motivation may play a role, particularly in systems that require significant user intervention, as many designers and engineers do not see a clear value and are reluctant to spend extra time documenting designs with rationale (Szykman *et al.*, 2001; Chandrasegaran *et al.*, 2013). Nevertheless, in certain fields such as computer science and software development, documenting source code is critical for maintenance and reuse (Steidl, Hummel, & Juergens, 2013) and as such, has achieved high levels of success and sophistication.

In this paper, we present a voice-based annotation system for collaborative CAD where speech recognition technology is used to capture design rationale and knowledge in the form of 3D annotations that are created in real time within the CAD environment during a modeling session. We describe the software architecture and the implementation details of a functional prototype using a commercial CAD system. Our software builds on a product data management (PDM)-based 3D annotation platform previously developed by the authors (Camba, Contero, Company, & Pérez, 2017).

In the next section, we justify the need for our research by reviewing relevant work in the area of knowledge capture and representation and discuss a series of application scenarios for voice-based annotation tools. Next, we describe the high-level architecture and implementation details of the proposed software infrastructure, emphasizing how the voice functionality builds on and extends the capabilities of the 3D annotation platform previously developed. The proposed prototype was validated through a series of experiments with a group of CAD users in a collaborative setting. The statistical analyses and results of these studies are discussed in Sections 4 and 5. Finally, conclusions and future work are presented in Section 6.

2. Related Work

Over the years, a number of strategies have been proposed to capture and document design rationale. Most notably, Kunz and Rittel's (1970) Issue-Based Information System and the numerous variations and alternative approaches, such as the Procedural Hierarchy of Issues model (McCall, 1991), the Question, Option, Criteria (MacLean, Young, Bellotti, & Moran, 1991), and the Issue Solution Artifact Layers (Liu, Liang, Kwong, & Lee, 2010), have served as the basis for many knowledge representation solutions (Conklin, Selvin, Buckingham Shum, & Sierhuis, 2001; Bracewell *et al.*, 2004).

Recent approaches that leverage storytelling and video sharing functionality to capture and communicate knowledge during the engineering process have shown great potential (Camba, Contero, & Salvador-Herranz, 2014; Zammit, Gao, Evans, & Maropoulos, 2018), although the practical value and impact in industrial settings have not been fully assessed. Other mechanisms based on word mappings, multiple-domain matrices, and especially annotations are considered more natural and intuitive, and thus more useful for providing a general view of the information and the relationships between items of information (Rasoulifar, Eckert, & Prudhomme, 2014). In this context, advances in the areas of model-based definition (MBD) and product manufacturing information (PMI) have enabled the development of new markup tools and annotation-based mechanisms for CAD where information is linked directly to part geometry (Sandberg & Na¨ssstro¨m, 2007; Camba *et al.*, 2014; Lundi, Lejon, Dagman, Näsström, & Jeppsson, 2017; Ma, Song, Xie, & Zhou, 2017). Authors Jones *et al.* (2020) recently experimented with a model-based approach to information access to improve engineering information retrieval, which even allows for non-text documents to be indexed. Nevertheless, textual representations of design and engineering knowledge are suitable for computer consumption (Yong, Deyu, Hong, & Gang, 2019) as they enable connectivity and integration with other product information and business processes across the enterprise.

Annotations, specifically 3D annotations, enable the exchange of design intent and rationale with other users directly through the 3D model and in a manner that is similar to how software engineers and programmers document source code (Camba *et al.*, 2014). However, the unstructured nature of this type of information makes it cumbersome and time consuming to represent in textual form, particularly when compared to verbal communication. Furthermore, engineers and designers often used vague expressions in their verbalizations of a problem or a design approach, particularly during the early stages of the design process, where the design problem is ill-defined and designers iterate possible solutions (Khan & Tunçer, 2019). These practices make it difficult to formalize context-specific information and establish semantics in CAD models. Despite some recent efforts to automate the creation of annotations, the process depends heavily on user input as the knowledge is not linked to the geometry of the 3D product representation (Soria & van der Hoek, 2019) and the results are not comparable to those provided directly by engineers (Cheng, He, Lv, & Cai, 2019). If the user does not actively maintain the annotations or neglects to retrieve the information to annotate, some important information may be lost. Many efforts have been devoted to the study of how annotations can directly support design methodologies, by capturing the design process, retrieving the rationale, and supporting and integrating with other engineering tools. However, results are still insufficient (Ding & Liu, 2010).

From a user's point of view, voice annotations are more expressive than textual annotations (Chalfonte, Fish, & Kraut, 1991). There are elements of verbal communications such as tone, emphasis, or pitch that are difficult or impossible to replicate with text. Verbal communication is also more convenient and faster to create and consume than text, since we tend to speak faster than we can type. Working with lengthy textual content is notoriously difficult, especially when using portable devices with small screen sizes. In this regard, the use of multimodal interfaces that implement gestures and speech is considered suitable for CAD modeling, as they offer an improved and more natural user experience than conventional input devices (Oviatt, 1999). However, commercial CAD systems still rely heavily on traditional methods of interaction based on mouse and keyboard.

Some application spaces that demonstrate the capability of voice-based annotations in the context of knowledge capture

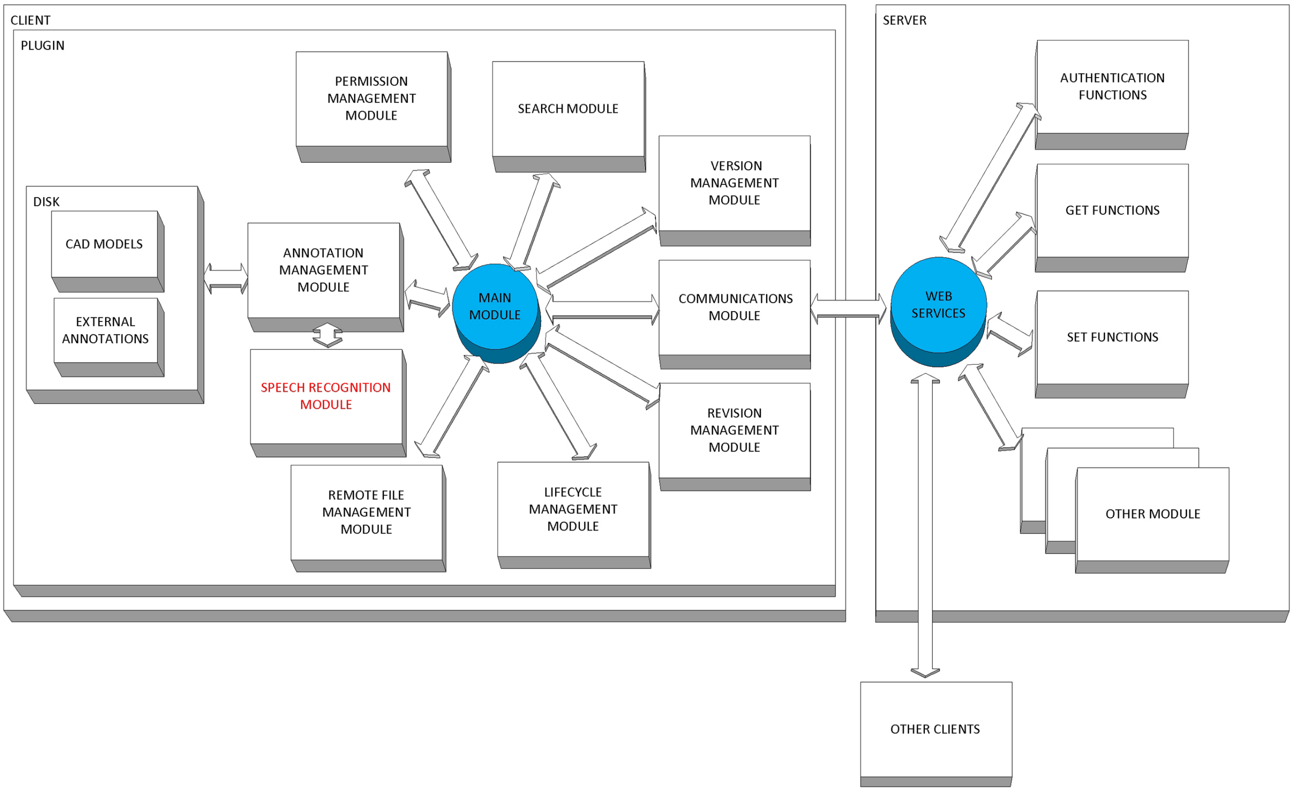


Figure 1: System architecture with speech recognition module.

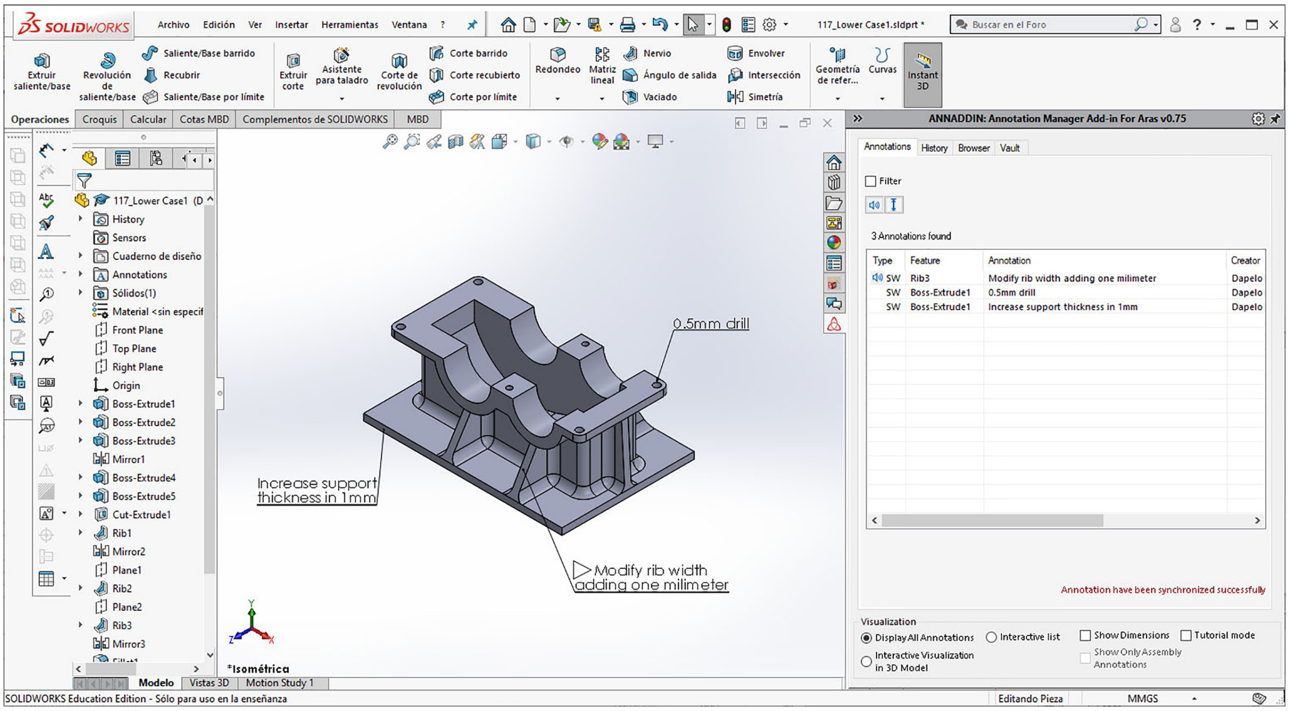
and exchange were identified by Camba, Naya, Perez-Lopez, and Contero (2020):

- Collaborative design review sessions: Voice-based annotations can streamline collaborative design workflows by providing an interaction mechanism to capture the large amounts of data generated through discussions, comments, and feedback, make it immediately available in the PDM system, and automatically connect it to explicit product attributes and properties. These data often refer to specific aspects of the CAD model, which is shared in real time among all stakeholders. Solutions in this application space, particularly in the context of virtual environments, have been shown to strengthen communication and knowledge management (Lenne, Thouvenin, & Aubry, 2009).
- Customer involvement and co-design: When a 3D model is used as a mediator in the communication process, voice annotations provide a natural interface that enables an active dialog with customers and suppliers and allow designers to capture their feedback and knowledge, instantly making the information available to other users through the PDM platform.
- Document generation: Voice-based annotations and the corresponding transcriptions can facilitate the creation of technical documentation by connecting operational descriptions, conditions, and flows to convey work instructions, ensuring quality and format consistency, up-to-date dynamic content (the information is linked to the 3D model), and traceability within the PDM system.
- CAD training: In student-centered learning environments, particularly in the context of online and distance education, voice-based annotations can enable and facilitate the delivery of personalized feedback. In a 3D modeling class, feedback can be extensive, time consuming, and challenging to verbalize without referencing the model or certain aspects of the design requirements. For example, assessing the quality of a parametric 3D model requires a thorough examination of the model structure, dimensions, and constraints, which often involves checking dozens of elements and determining the level of correctness based on a set of multidimensional criteria. Using voice-based annotations, instructors can simplify the creation of feedback while providing richer and more expressive explanations, creating a more personal connection with the student.

In this paper, we describe and evaluate a system for capturing design and engineering knowledge in a CAD model directly through voice interaction. The information is automatically transcribed to a 3D annotation, linked to specific parts of the model, and fully searchable and integrated with existing design and engineering workflows via the PDM system.

3. Annotation System and Speech Recognition Module

The system described in this paper implements a voice interpreter on the software architecture previously developed by the authors to manage extended annotations in CAD models (Camba *et al*., 2017). The architecture provides a framework in the context of a PDM system for integrating product and manufacturing information (PMI) within a 3D model directly from the CAD environment. The architecture is client–server; the server side of the system provides services to client applications such as PDM server authentication, file download/upload, lifecycle management, and others, and the client provides specialized functionality, as illustrated by the various modules shown in Fig. 1.



ANNADDIN: Annotation Manager Add-in For Aras v0.75

Annotations | History | Browser | Vault

☐ Filter

3 Annotations found

Type	Feature	Annotation	Creator
🔈 SW	Rib3	Modify rib width adding one milimeter	Dapelo
SW	Boss-Extrude1	0.5mm drill	Dapelo
SW	Boss-Extrude1	Increase support thickness in 1mm	Dapelo

Annotation have been synchronized successfully

Visualization

● Display All Annotations ○ Interactive list ☐ Show Dimensions ☐ Tutorial mode

○ Interactive Visualization in 3D Model ☐ Show Only Assembly Annotations

Figure 2: Module interface (right) with annotations tab active.

The client is implemented as an add-on for a CAD system (i.e. SolidWorks) and is fully integrated within the 3D modeling environment. The interface is organized in four tabs: Annotations, History, Browser, and Vault. The "Annotations" tab displays all the annotations in the 3D model in a tabular form where each annotation is shown on a separate row, as illustrated in Fig. 2. The "History" tab provides a mechanism for tracking the changes performed to the annotations associated with the model. Basic PDM functionality (e.g. Check-In, Check-Out, change state, etc.) is available via the "Vault" tab.

The speech recognition module is responsible for recording the user's voice, automatically transcribing the signal, creating a textual annotation from the audio signal, and storing the information in the PDM server. The module also provides additional functionality such as filtering, searching, synchronization, and information visualization, and is comprised of three functional blocks: (1) audio recording and playing, (2) speech recognition, and (3) query management, as shown in Fig. 3.

The first functional block (audio recording and playing) was developed using the NAudio library (Heath, 2016); the second is based on Google's Speech-to-Text application programming interface (API; Google, 2017), and the third was developed to manually build JSON queries against a REST service from Google using the Newtonsoft.json library (Newton-King, 2012). The system performs a Google authentication via an API key, which allows tracking data from different users.

From a CAD user's point of view, voice annotations are created by clicking on the speaker icon available in the "annotation" tab in the interface and using the PMI module of the CAD system to attach the annotation to the geometry. When the first annotation is created, a language selection box is displayed. The selected language is saved for subsequent annotations. Audio recording starts automatically after the user speaks to the microphone and can be stopped manually by clicking the "Stop recording" button, or automatically after 30 seconds (to avoid generating large files). When the recording stops, a local data file containing the audio is sent to a Google Cloud service, which returns the transcribed speech in a textual form. The text and the local audio file are then automatically linked to the annotation both in the CAD file and the PDM system. Voice annotations are represented in the 3D model by a triangular shape that mimics the traditional "play" button followed by the transcribed text, and with a speaker icon in the annotations table, as shown in Fig. 4. Voice annotations can be played directly from the CAD environment.

For example, to create the annotation shown in Fig. 4, the user must click the speaker icon in our module, which activates a text box in the 3D model next to the mouse pointer. This is the same text box that is created when a text annotation is added via the SolidWorks PMI module. The next step is to click on the model feature that needs to be annotated (the rib, in our example) and then the area where the text will be created. Next, the user can directly speak to the microphone to create the note (in our example, "Modify rib width adding one millimeter") and stop the recording process by clicking the stop button in the dialog box. The audio recording is then sent to the Google Cloud service to be transcribed. Once complete, the transcribed text is automatically added to the 3D text note next to the selected feature in the model, and a new row is added to the annotator manager window ("annotations" tab), as shown in Fig. 4, right.

An additional key point in our system is the ability to filter and search voice annotations. These functions allow users to reduce the amount of on-screen information based on specific criteria, such as Text, Date, Creator, and Feature. The functions are implemented via the "Filter" checkbox, available on the "annotations" tab in the Annotation manager interface. When activated, the user can enter a text string to define the search criteria, as shown in Fig. 5. Multiple criteria can be applied simultaneously by adding additional search fields to the query. For example, the

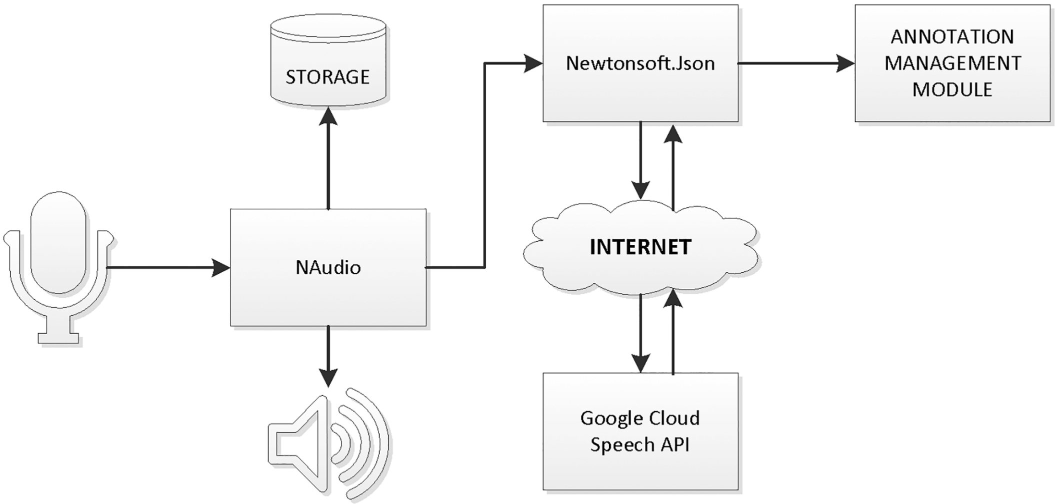
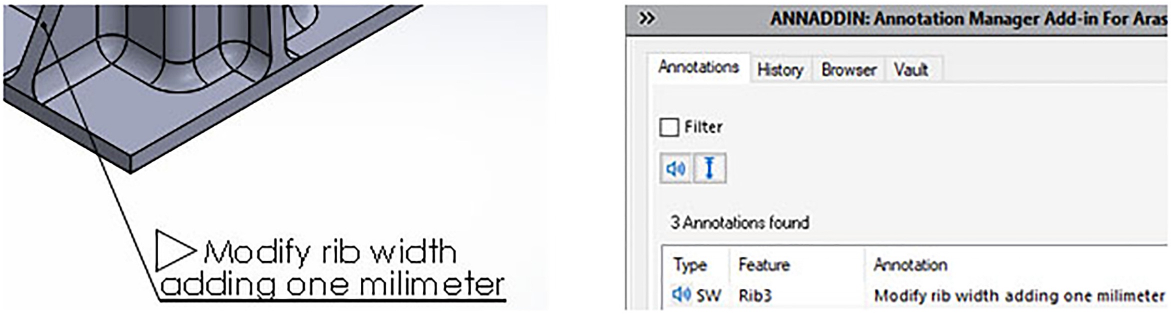


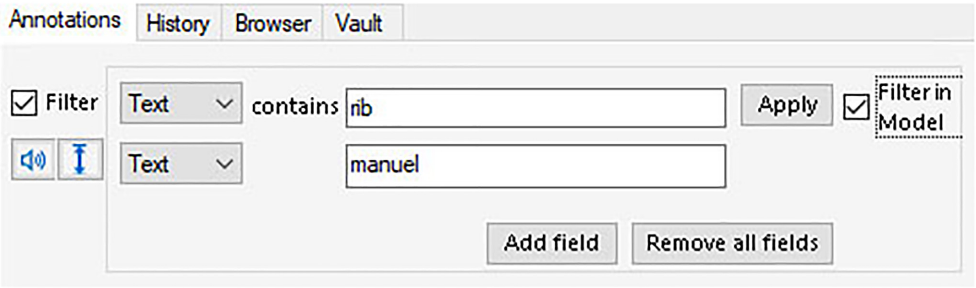
Figure 3: Speech recognition flow diagram.



3 Annotations found

Type	Feature	Annotation
SW	Rib3	Modify rib width adding one milimeter

Figure 4: Representation of the annotations in the user interface.



☑ Filter Text contains rib ☑ Filter in Model

Text manuel

Figure 5: Annotation search and filtering capabilities.

user can search for annotations created by a particular user and containing a certain keyword. Search results are shown in the annotations table in the module interface and in the 3D model (as long as the “Filter in Model” checkbox is selected). The search function works locally, but it is also implemented in the “vault” tab, so annotations can be searched at the PDM level.

The system has two main limitations that are due to the inherent constraints of the transcription service used (i.e. Google Cloud): the need for an internet connection and the inability to autodetect the language being spoken. These limitations are not critical for the research objectives of this work, which was centered on the idea of using voice notes to annotate 3D CAD geometry and making the information available for further processing.

4. Materials and Methods

In order to validate the proposed audio annotation system, we conducted a series of experiments with a group of 67 junior engineering students enrolled in a CAD course. All subjects were informed about the objective of the study, the mechanisms to ensure data confidentiality, and provided consent to participate.

The course consists of two weekly sessions (1-hour lecture and 3-hour lab). During the lab sessions, students use a 3D CAD package (i.e. SolidWorks®) to create 3D parts, assemblies, and technical drawings. Students work in pairs on a semester-long design project in which they create 3D models of parts and assemblies (and the corresponding 2D drawings) of a real product. Students select their products from various options proposed by

Figure 6: Examples of assemblies for the project proposal and elements to redesign (marked with circles).

the instructor. Some examples are illustrated in Fig. 6. Students are also asked to redesign one of the parts of the original product (e.g. the handle in the case of the blender, and the arm rest in the desk chair), as highlighted in Fig. 6. The redesigned part must be original, visually appealing, ergonomic, functional, and easy to manufacture.

Each group of students is required to submit the assembly model eight weeks after the project is assigned. The work is reviewed by the instructor who provides feedback with comments and recommendations to improve the product. After redesigning the specific part of their product, students submit their CAD files. As part of the final submission, students are asked to identify three parts in the assembly that were modified based on the instructor's feedback after the first submission. The goal is for students to explain the modifications and improvements that were introduced to the product using different annotation methods, i.e. using a text file (i.e. Microsoft Word), text annotations in the 3D model, or voice annotations.

For our study, we conducted a comparative analysis of the three annotation methods based on various criteria, and a qualitative study in which participants' data were collected to assess the user experience and effectiveness of each of these methods. All participants in the study received training on how to create quality 3D models. We used the definition of CAD quality and the quality dimensions proposed by Company, Contero, Otey, and Plumed (2015) and Otey, Company, Contero, and Camba (2019) and the corresponding self-assessment rubric. Part modeling concepts were explained in a regular class setting. Instruction on assembly modeling and technical drawings was delivered online due to the COVID-19 pandemic. Students also received instruction on the use of annotations as a tool to convey design intent and enrich 3D models. The quality dimensions as described in the assessment rubric are shown in Table 1. The instructor's feedback for the first submission refers specifically to items: 3.2, 3.3, 4.1, 4.2, 4.3, 5.1, 5.2, 6.1, 6.2, and 6.3 in Table 1.

Detailed information about the annotation software was provided to participants, which included step-by-step instructions on how to install the SolidWorks plugin that contains the speech recognition module, the creation of both text and voice annotations in the CAD model, and the submission procedure for the annotated files. Students were asked to select three of the parts they originally modeled and explain the changes introduced in the final submission using a different annotation mechanism for each part:

1) Creating a text file (*TFi*) using a text editor (e.g. Microsoft Word) and submitting the file to the Course Management System.
2) Creating text annotations (*TAn*) directly on the 3D model using standard PMI tools and submitting the CAD files to the Course Management System.
3) Creating voice annotations (*VAn*) directly on the 3D model using our custom software and submitting both the CAD files and the XML files generated by the plugin to the Course Management System.

To evaluate the usability of the text file (*TFi*), text annotations (*TAn*), and the voice annotation mechanism (*VAn*), participants completed an online questionnaire that addressed their previous experience using annotation tools in texting applications, games, or other types of online communication, and the user experience of the annotation method with respect to: (1) adequacy and clarity of the explanations; (2) ease of use; (3) speed; (4) adaptability to the user's work preferences; (5) readability and comprehension; (6) ease of modification; (7) efficiency of information exchange; and (8) user predisposition for regular use. The questionnaire was designed using five-point Likert scales: (1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; and (5) Strongly agree.

4.1 Text readability parameters

Readability is a metric that measures the ease with which a text can be understood. Every language has a characteristic writing style; therefore, the same readability metrics are not always applicable to all languages. Since all our participants were native Spanish speakers, we considered the most appropriate metrics to assess texts in this language. In these metrics, the higher the score, the more readable the text.

To extract text legibility parameters from the annotated models and text files, we developed a software tool using Microsoft Visual Studio 2017, Visual Basic .Net, the .Net Microsoft Office Word 2018 API, and the Solidworks 2018 API. The software parses the textual content from the files, applies various mathematical formulas to the text to determine the readability metrics, and writes the results to a file. The process is illustrated in Fig. 7.

Our tool can process annotated 3D models created in SolidWorks as well as Word documents. When processing annotated models, the system retrieves all SolidWorks models in a folder

Table 1: CAD quality criteria used to assess 3D modeling skills (Company *et al.*, 2015).

1. The model is valid
1.1 The model file can be located and opened without any operations in progress.
1.2. The model file can be reopened, has the correct content, and can be used by other users.
2. The model is complete
2.1. The model replicates the shape of the part/definition problem.
2.2. The model replicates the size of the part/definition problem.
3. The model is consistent
3.1. Profiles are free from duplicated and segmented lines.
3.2. Profiles are fully constrained (without excessive fixed constraints).
3.3. The models are aligned and oriented with respect to the global reference system and suitable datums.
4. The model is concise
4.1. The model was created with adequate and nonrepetitive operations.
4.2. Pattern operations (translate-and-repeat, rotate-and-repeat, and symmetry) were used when possible.
4.3. The model tree contains adequate datums and is free from unnecessary dependencies.
5. The model is clear
5.1. Modeling operations in the model tree are labeled to emphasize their function, instead of the operation name.
5.2. Related modeling operations are grouped in the model tree to emphasize parent–child relationships.
6. The model conveys design intent
6.1. The model tree is like a “script” that describes the elements that constitute the part and their functionality.
6.2. The model is created to prevent the loss or transfer of critical design dimensions when the model is altered, without losing symmetries or patterns in the part.
6.3. The model enables redesign without errors or undesirable changes (i.e. the model is flexible and robust).

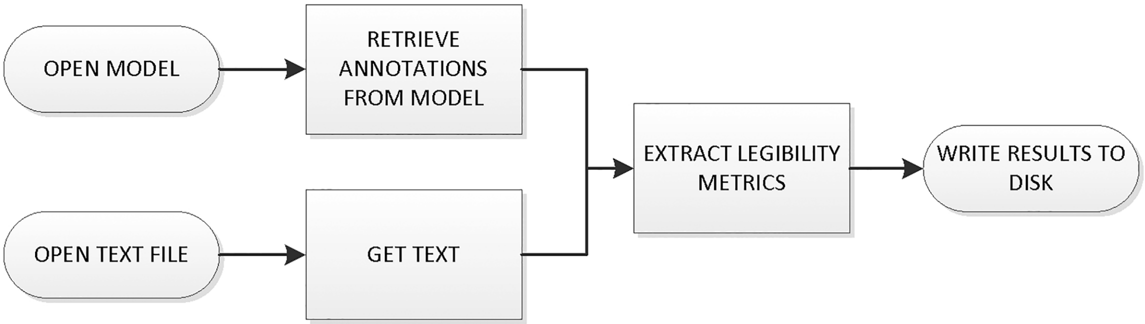


Figure 7: Text readability flow diagram.

and opens each file in a hidden SolidWorks window. All annotations are then extracted using the appropriate SolidWorks API functions. The system calculates the following metrics: number of words, estimated reading time, Fernandez Huerta readability (Fernandez Huerta, 1959), Gutierrez understanding (Gutiérrez de Polini, 1972), Szigriszt-Pazos perspicuity (Barrio-Cantalejo *et al.*, 2008), INFLESZ scale (Barrio-Cantalejo *et al.*, 2008), μ legibility (Muñoz, 2006), and Crawford understanding (Crawford, 1989). The information is then saved to a csv file. The same steps are performed when processing Word files but using the Word API instead of the SolidWorks API.

5. Analysis and Results

All files submitted by the participants were processed after submission to verify whether annotations (voice and text) were correctly added. Word documents were also filtered, so irrelevant content that was not directly related to the annotations (e.g. titles, student names, embedded images, etc.) was deleted. A total of 62 files were analysed. The mean rank of these parameters for each annotation method is shown in Table 2. An alpha level of 0.05 was used for all statistical tests.

A Kruskal–Wallis H test revealed a statistically significant difference among the distribution of the number of words used in each method of annotation and the estimated reading time. Indeed, when students explained their modifications in a text file (*TFi*), they tend to use more words (mean rank = 116.73) than when they created annotations in the 3D model via PMI tools, *TAn* (mean rank = 75.85), or voice, *VAn* (mean rank = 87.93). Naturally, when descriptions and explanations are not supported by images or illustrations of the 3D model, users must include more details to make the explanations clear, which naturally take longer to read.

The Kruskal–Wallis H test did not reveal any significant differences between the annotation methods in any of the readability metrics. The type of annotation mechanism does not seem to influence the clarity and conciseness of the annotations.

5.1 Questionnaire analysis

A total of 63 responses to the online questionnaire were collected. The questionnaire was designed to assess nine aspects related to the use of the various annotation methods through a five-point Likert scale: (1) Strongly disagree; (2) Disagree; (3) Neutral; (4) Agree; and (5) Strongly agree.

Our results revealed a preference for annotations embedded in the 3D model (*TAn*, *VAn*) over annotations in a text file (*TFi*), as shown by the frequencies of the responses “Agree” and “Strongly

Table 2: Kruskal–Wallis H test results (metrics).

	χ^2 (2, N = 186)	P-value	TFi (N = 62) Mean rank	TAn (N = 62) Mean rank	VAn (N = 62) Mean rank
Words	18.66	$P < 0.001$	116.73	75.85	87.93
Estimated reading time	18.88	$P < 0.001$	116.73	75.85	87.93
INFLESZ	.017	$P = 0.992$	93.05	93.24	94.21
LegMu	3.17	$P = 0.205$	83.83	100.31	96.36
Fernandez Huerta	.027	$P = 0.982$	92.69	93.53	94.27
Gutierrez	.031	$P = 0.985$	92.53	93.84	94.13
Szigriszt-Pazos	.017	$P = 0.992$	93.05	93.24	94.21
Crawford	1.48	$P = 0.477$	100.29	90.18	90.03

Table 3: Frequency of responses ("Agree" and "Strongly agree").

	TFi			TAn			VAn		
Aspects studied	%Agr	%Str.Agr	Total	%Agr	%Str.Agr	Total	%Agr	%Str.Agr	Total
Adequacy	28.6%	9.5%	**38.1%**	47.6%	47.6%	**95.2%**	58.7%	22.2%	**81.0%**
Ease of use	33.3%	34.9%	**68.3%**	38.1%	58.7%	**96.8%**	49.2%	22.2%	**71.4%**
Speed	20.6%	14.3%	**34.9%**	44.4%	44.4%	**88.9%**	34.9%	44.4%	**79.4%**
Adaptability	28.6%	15.9%	**44.4%**	41.3%	52.4%	**93.7%**	49.2%	25.4%	**74.6%**
Readability comprehen.	39.7%	30.2%	**69.8%**	49.2%	49.2%	**98.4%**	41.3%	47.6%	**88.9%**
Ease of modification	39.7%	47.6%	**87.3%**	42.9%	50.8%	**93.7%**	23.8%	12.7%	**36.5%**
Information exchange	39.7%	30.2%	**69.8%**	44.4%	52.4%	**96.8%**	44.4%	47.6%	**92.1%**
Regular use	20.6%	7.9%	**28.6%**	25.4%	66.7%	**92.1%**	30.2%	19.0%	**49.2%**

Agree" in Table 3. The relative frequencies of the participants' responses to each aspect of the annotation mechanism are shown in Fig. 8.

The Kruskal–Wallis H test was used to analyse the participants responses. Our results are shown in Table 4 and reveal significant differences among participants in all the aspects analysed.

A post hoc analysis (in pairs) shows that *TAn* performed significantly better in terms of ease of use than *TFi* and *VAn*. *TFi* was the least valued method in terms of speed, readability and comprehension, and information exchange; *VAn* was the least preferred method in terms of ease of modification; and *TAn* was the preferred method for regular use of annotations.

The participants' previous experience, preferences, and use of voice messages in other applications are illustrated in Table 5. In general, all participants were familiar with the use of voice messaging in various forms of remote communication and most did not express a clear preference in terms of text or voice messaging. All participants claimed to be active users of instant messaging applications, almost half of whom (49.2%) usually did not use voice messaging, and 47.6% said they used both text and voice messaging. Similarly, 60.3% of participants admitted playing online games, half of whom (30.2%) preferred verbal communication over text messaging. Finally, 93.7% of the participants claimed to use other messaging applications regularly, 44.5% of whom claimed to use both text and voice messaging.

To determine whether the participants' previous experiences influenced their preferences for a specific annotation method, we compared the answers to "predisposition for a regular use" with their preferences on an instant messaging app (i.e. WhatsApp, since all participants were active users). To this end, for each participant, we determined the annotation method that was scored the highest and calculated the relative frequencies. Our results show that *TFi* was the preferred option 3.2% of the time, *TAn* was selected as the best option by 57.1% of the participants, and *VAn* by 12.7%. For the remaining 27%, different annotation methods were scored equally, so they were not included in the comparative analysis.

The Kruskal–Wallis H test revealed that the relationship between the participants' previous experience with instant messaging tools and the preferred annotation method is significant χ^2 (2, N = 46) = 1.55, $P = 0.46$, which suggests that users who regularly use voice messaging in other applications are more likely to use voice annotations, or both voice and text annotations, in collaborative CAD.

6. Conclusions and Future Work

In this paper, we described a novel voice-based interaction mechanism built on a collaborative software architecture that integrates 3D model annotations with other design and engineering processes via a PDM system. The voice-based functionality facilitates the capture of design knowledge directly from audio recordings that can be created in the CAD environment, which are transcribed automatically, converted to 3D annotations, and finally linked to specific geometric elements of the design. Our approach enables connectivity by making design knowledge searchable and suitable for computer consumption. Our software implementation demonstrates the value of the proposed architecture as a mechanism to facilitate design documentation, which paves a path to new multimodal interaction methods and interfaces for more effective CAD and collaborative engineering tools, particularly for communicating and sharing design intent and rationale. In this regard, voice-based knowledge capture and management in collaborative design domains can expand product definition and documentation beyond annotated geometry and convey richer information.

Our validation studies confirm the value and usefulness of our approach for conveying design information embedded in the

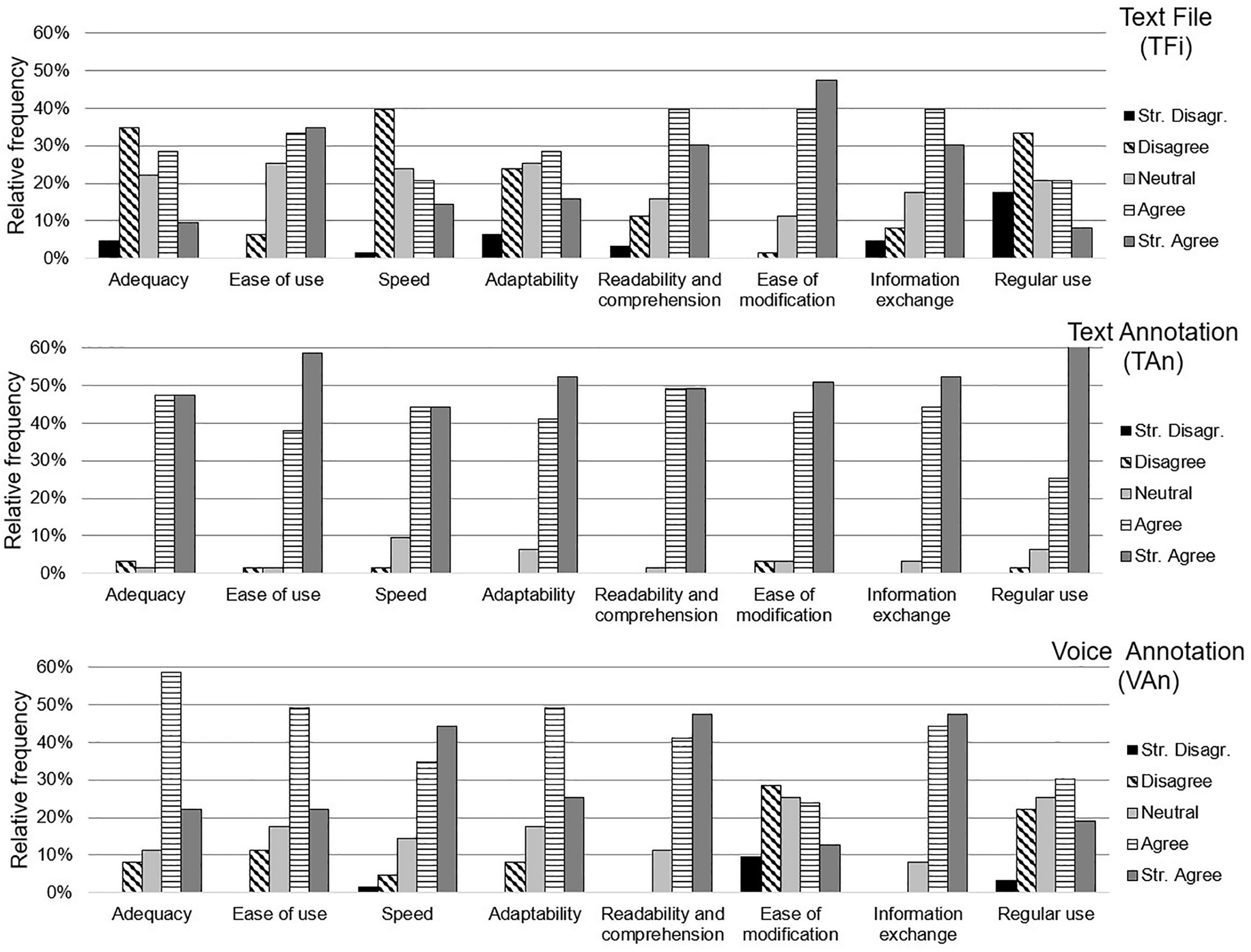


Figure 8: Bar chart with the score frequencies for each aspect studied.

Table 4: Kruskal–Wallis H test results (parameters).

	χ^2 (2, N = 189) ($P < 0.001$)	Mean rank adequacy		
		TFi	TAn	VAn
Adequacy	53.17	59.43	125.79	99.79
Ease of use	24.03	86.74	120.29	77.98
Speed	44.03	59.46	115.81	109.73
Adaptability	40.64	65.74	124.73	94.53
Readability and comprehension	14.12	75.92	107.56	101.52
Ease of modification	55.10	112.41	117.14	55.44
Information exchange	15.55	75.00	108.12	101.88
Predisposition for regular use	71.14	60.02	138.4	86.58

Table 5: Relative frequencies regarding the use of messaging.

Response	Instant messaging	Online games	Other applications
I am not a user	0%	39.7%	6.3%
I prefer text messaging	49.2%	12.7%	30.2%
I prefer voice messaging	3.2%	30.2%	19%
I use both text and voice messaging	47.6%	17.4%	44.5%

3D CAD model. Our results show that most users prefer 3D annotations in the CAD model (both text and voice-based) over annotations made in text editors. In this regard, the availability of the geometry when creating the annotations and the ability to document in-context and link content directly to the 3D model provide a powerful means for documentation, particularly in scenarios where considerable amounts of information need to be communicated, such as design review sessions and the delivery of feedback to students. Regarding the differences between annotations created manually and voice-based annotations, we speculate that the participants' lack of experience and training with a relatively complex system such as the proposed tool coupled with the timing of the experiment at the beginning of the Covid-19 pandemic may have influenced our results. In this regard, as future work, we are interested in conducting further long-term studies with larger groups and in industrial environments to determine the effectiveness of our approach when compared to traditional methods of design documentation and communication. From a technical standpoint, we are studying and developing methods to further improve the collaborative aspects of the system. For example, by automatically recognizing users in a design review session, voice annotations can be automatically linked to the creators as they engage in conversations, which can enable data analytics and the automatic generation of reports and similar documents.

Finally, our system is built on the PMI functionality provided by the CAD platform (in our case, the resources provided by the SolidWorks MBD module), which means that all existing capabilities such as Geometric Dimensioning and Tolerancing (GD&T) are still available to expert users. Therefore, our system could enhance these particular types of annotations by automatically recognizing and transcribing GD&T data from voice descriptions. In our view, text and voice-based annotations can provide an additional layer of information to communicate design decisions and improve traceability while being able to adapt to specific domains.

Acknowledgement

This work was supported by the Spanish grant DPI2017-84526-R (MINECO/AEI/FEDER, UE), project "CAL-MBE, Implementation and validation of a theoretical CAD quality model in a Model-Based Enterprise (MBE) context," and the grant UJI-A2017-15 (Universitat Jaume I), project "New metrics to improve the quality of 3D CAD models."

Conflict of interest statement

None declared.

References

Anerousis, N., & Panagos, E. (2002). Making voice knowledge pervasive. *IEEE Pervasive Computing*, 1(2), 42–48.

Barrio-Cantalejo, I. M., Simón-Lorda, P., Melguizo, M., Escalona, I., Marijuán, M. I., & Hernando, P. (2008). Validación de la Escala INFLESZ para evaluar la legibilidad de los textos dirigidos a pacientes. *Anales del Sistema Sanitario de Navarra*, *31*(2), 135–152.

Bracewell, R. H., Ahmed, S., & Wallace, K. M. (2004). DRed and design folders: A way of capturing, storing and passing on knowledge generated during design projects. In *ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. 235–246).

Camba, J., Contero, M., & Johnson, M. (2014). Management of visual clutter in annotated 3D CAD models: A comparative study. In A. Marcus (ed.), *Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments. DUXU 2014. Lecture Notes in Computer Science*(Vol. 8518). Cham: Springer. https://doi.org/10.1007/978-3-319-07626-3_37.

Camba, J., Contero, M., & Salvador-Herranz, G. (2014). Speak with the annotator: promoting interaction in a knowledge-based CAD environment built on the extended annotation concept. In *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*(pp. 196–201).

Camba, J., Contero, M., Johnson, M., & Company, P. (2014). Extended 3D annotations as a new mechanism to explicitly communicate geometric design intent and increase CAD model reusability. *Computer-Aided Design*, 1(57), 61–73.

Camba, J. D., Contero, M., Company, P., & Pérez, D. (2017). On the integration of model-based feature information in product lifecycle management systems. *International Journal of Information Management*, *37*(6), 611–621.

Camba, J. D., Naya, F., Perez-Lopez, D., & Contero, M. (2020). From voice to knowledge: A proposal for a voice annotation system to support collaborative engineering design processes. In *53rd Hawaii International Conference on System Sciences (HICSS-53)*, HICCS. Maui, HI. January 7–10.

Chalfonte, B. L., Fish, R. S., & Kraut, R. E. (1991). Expressive richness: a comparison of speech and text as media for revision. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*(pp. 21–26).

Chandrasegaran, S. K., Ramani, K., Sriram, R. D., Horváth, I., Bernard, A., Harik, R. F., & Gao, W. (2013). The evolution, challenges, and future of knowledge representation in product design systems. *Computer-Aided Design*, *45*(2), 204–228.

Cheng, Y., He, F., Lv, X., & Cai, W. (2019). On the role of generating textual description for design intent communication in feature-based 3D collaborative design. *Advanced Engineering Informatics*, *39*, 331–346.

Company, P., Contero, M., Otey, J., & Plumed, R. (2015). Approach for developing coordinated rubrics to convey quality criteria in MCAD training. *Computer-Aided Design*, *63*, 101–117.

Conklin, J., Selvin, A., Buckingham Shum, S., & Sierhuis, M. (2001). Facilitated hypertext for collective sensemaking: 15 years on from gIBIS. In *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia*. Århus, Denmark, Association for Computing Machinery.

Crawford, A. (1989). Fórmula y gráfico para determinar la comprensibilidad de textos del nivel primario en castellano. *Lectura Y Vida*, 4, 18–24.

Ding, L., & Liu, S. (2010). Markup in engineering design: A discourse. *Future Internet*, 2(1), 74–95.

Fernández Huerta, J. (1959). Medidas sencillas de lecturabilidad. *Consigna (Revista PedagóGica de la Sección Femenina de Falange ET y de las JONS)*, *214*, 29–32.

Google. (2017). Speech to text: Automatic speech recognition. Retrieved from https://cloud.google.com/speech-to-text?hl = es, Accessed 15 July 2020.

Gutiérrez de Polini, L. E. (1972). Investigación sobre lectura en Venezuela. *Primeras Jornadas de Educación Primaria*. Ministerio de Educación, Caracas.

Heath, M. (2016). Audio and MIDI library for .NET. Retrieved 13 February 2019, from https://github.com/naudio/NAudio.

Jones, D., Snider, C., Matthews, J., Yon, J., Barrie, J., Robinson, K., & Hicks, B. (2020). Model-based information navigation for engineering documents. *Computers in Industry*, *121*, 103254.

Khan, S., & Tunçer, B. (2019). Speech analysis for conceptual CAD modeling using multi-modal interfaces: An investigation into architects' and engineers' speech preferences. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AI EDAM*, *33*(3), 275–288.

Kunz, W., & Rittel, H. W. (1970). *Issues as elements of information systems*. Institute of Urban and Regional Development, University of California, 1970.

Lenne, D., Thouvenin, I., & Aubry, S. (2009). Supporting design with 3D-annotations in a collaborative virtual environment. *Research in Engineering Design*, *20*, 149–155.

Levallet, N., & Chan, Y. E. (2016), Knowledge loss and retention: The paradoxical role of IT. In *Successes and failures of knowledge management*(pp. 97–111). Morgan Kaufmann.

Liu, Y., Liang, Y., Kwong, C. K., & Lee, W. B. (2010). A new design rationale representation model for rationale mining. *Journal of Computing and Information Science in Engineering*, *10*(3), 031009.

Lundi, M., Lejon, E., Dagman, A., Näsström, M., & Jeppsson, P. (2017). Efficient design module capture and representation for product family reuse. *Journal of Computing and Information Science in Engineering*, *17*(3), 031002–1-031002-16.

Ma, Q. Y., Song, L. H., Xie, D. P., & Zhou, M. J. (2017). Development of CAD model annotation system based on design intent. *Applied Mechanics and Materials*, *863*, 368–372.

McCall, R. J. (1991). PHI: a conceptual foundation for design hypermedia. *Design Studies*, *12*(1), 30–41.

MacLean, A., Young, R. M., Bellotti, V. M., & Moran, T. P. (1991). Questions, options, and criteria: Elements of design space analysis. *Human-Computer Interaction*, *6*(3), 201–250.

Munoz, M. (2006). Legibilidad y variabilidad de los textos. *Bol. De Investigacio´n Educacional*, *21*(2), 13–25.

Newton-King, J. (2012). Newtonsoft. Retrieved 13 February 2019, from https://www.newtonsoft.com/json.

Otey, J., Company, P., Contero, M., & Camba, J. D. (2018). Revisiting the design intent concept in the context of mechanical CAD education. *Computer-Aided Design and Applications*, *15*(1), 47–60.

Otey, J., Company, P., Contero, M., & Camba, J. D. (2019). Assessment of parametric assembly models based on CAD quality dimensions. *Computer-Aided Design and Applications*, *16*(4), 628–653

Oviatt, S. (1999) Ten myths of multimodal interaction. *Communications of the ACM*, *42*, 74–81.

Rasoulifar, G., Eckert, C., & Prudhomme, G. (2014). Supporting communication between product designers and engineering designers in the design process of branded products: a comparison of three approaches. *CoDesign*, *10*(2), 135–152.

Sandberg, S., & Na¨sstro¨m, M. (2007). A proposed method to preserve knowledge and information by use of knowledge enabled engineering. In *ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*(pp. 207–212). American Society of Mechanical Engineers.

Soria, A. M., & van der Hoek, A. (2019). Collecting design knowledge through voice notes. In *Proceedings of the 12th International Workshop on Cooperative and Human Aspects of Software Engineering*(pp. 33–36). IEEE.

Steidl, D., Hummel, B., & Juergens, E. (2013). Quality analysis of source code comments. In *21st IEEE International Conference on Program Comprehension (ICPC)*(pp. 83–92).

Szykman, S., Sriram, R. D., & Regli, W. C. (2001). The role of knowledge in next-generation product development systems. *Journal of Computing and Information Science in Engineering*, *1*(1), 3–11.

Yong, Y., Deyu, H., Hong, W., & Gang, Z. (2019). Data mining and recommendation of engineering note items in MBD dataset. In *Proceedings of the 2019 11th International Conference on Computer and Automation Engineering*(pp. 1–6). ACM.

Zammit, J., Gao, J., Evans, R., & Maropoulos, P. (2018). A knowledge capturing and sharing framework for improving the testing processes in global product development using storytelling and video sharing. In *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*(Vol. *232* (13), pp. 2286–2296).