Proceedings of the ASME 2023
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC-CIE2023
August 20-23, 2023, Boston, Massachusetts

DETC2023-114927

CAD CHALLENGES APP: AN INFORMATICS FRAMEWORK FOR PARAMETRIC MODELING PRACTICE AND RESEARCH DATA COLLECTION IN COMPUTER-AIDED DESIGN

Yuanzhe Deng[1,2], Matthew Mueller[1], Matt Shields[1]

[1] PTC Education, PTC Inc., Boston, MA
[2] University of Toronto, Toronto, ON, Canada

ABSTRACT

Computer-aided design (CAD) is a key tool in modern engineering design and manufacturing, making design education and design research with CAD important fields of study. Effectively teaching modelling strategies in traditional classrooms is challenging and the design research community faces barriers in participant recruitment for research studies. In this paper, we propose a framework that connects the teaching and research community with design informatics in CAD. We productized the framework, named the "CAD Challenges" web application, and integrated it to Onshape, a commercially available cloud-CAD software. With free and easy access to this app, users gain access to a library of modelling challenges from within an Onshape document. The app automatically evaluates submissions and provides feedback, enabling asynchronous learning and the development of CAD expertise through practice. After challenge attempts, data on both the design process and the completed model are collected, enabling insight into the different strategies that can be used to create the same geometry. While providing a free and accessible training tool for learners, the big data generated through challenge attempts can provide valuable insight into how students learn CAD and the modeling strategies used by experts. Benefits and opportunities enabled by the framework are discussed in detail with preliminary research analysis.

Keywords: computer-aided design, engineering informatics, data mining, education, design theory and methodology

1. INTRODUCTION

Computer-aided design (CAD) is an increasingly important and ubiquitous tool for modern engineering design. CAD in mechanical engineering, often referred to as mechanical CAD (MCAD), is commonly used to design and model parts and assemblies, validate product functions and behaviors with computer simulations, and document specifications of physical products, all in a virtual environment [1]. While traditional CAD software was first invented as a virtual design studio to enable more efficient sharing and documentation of engineering design [2], people soon realized its capability in enabling the use for engineering design, analysis, and communication [3]. With continuous computational advancements, modern CAD software continues to be improved in areas such as enabling collaboration as a cloud service [4] and utilizing artificial intelligence (AI) technologies to aid the design process [5].

Given the importance of the technology, CAD education has been part of the engineering curriculum for decades. Traditional CAD education includes the training of spatial reasoning [6,7], use of parametric features [8], and high-level modelling strategies [9]. A CAD designer is likely able to grasp the declarative knowledge of using basic commands through classroom-based CAD education, but it will take years for them to develop their design expertise in CAD through deliberate practice and industry design experience [10]. Currently, there is no scalable solution to provide a widely accessible platform for such CAD practices, which should ideally enable asynchronous learning for the students.

For a complex enough model, there theoretically exists a near infinite number of approaches to model the same geometry, using different combinations of parametric features in CAD. Some approaches are likely more efficient than others [11], some approaches may result in more robust and flexible models for design iterations and model reuse [12], and some approaches may better reveal the design intent for review and collaboration [13–15]. While designing with an optimal and sustainable approach in CAD is crucial, cognitive design thinking during the CAD modelling process is of particular interest to the design research community [16]. However, empirical studies of design thinking largely rely on verbal protocols, case studies, and controlled experiments, all of which result in tedious data analysis and slow rate of discoveries [16]. As such, advancing the field using a big data approach will significantly improve the efficiency for future studies [17].

While the education community lacks a scalable platform to support asynchronous CAD modelling practice, the research community also lacks sufficient data for research analysis. This paper proposes a framework as a solution that addresses both gaps and connects the two fields using design informatics as the medium, which allows the two communities to learn from and benefit each other. As students (and CAD users in general) model in CAD, a large amount of design data (i.e., informatics) is generated, which can be captured for research in CAD-based design activities [18]. Further, research findings from this potentially rich dataset of design informatics also provide great insights for CAD education, fostering the concept of crowd learning [19].

Section 2 of this paper further introduces some background and related work in design education, research, and informatics, all in the context of CAD. Then we present the design of the proposed framework in section 3. Finally, we extrapolate the usefulness of the framework by discussing how we may enable existing research from both the literature and preliminary analysis, while supporting education through an accessible and scalable platform.

2. BACKGROUND

This section presents some related work in the fields of design education and design research in CAD. The background of design informatics, the medium of the framework, is also laid out in this section.

2.1 Design Education in CAD

When modelling and design in CAD are discussed in an educational setting, CAD knowledge is generally taught in two categories: declarative knowledge and procedural (also known as strategic or cognitive) knowledge. Declarative knowledge refers to the mastery of individual parametric features and commands available in a CAD software [20]. Procedural knowledge in CAD refers to the choice of modelling strategies, which involves making decisions and executing a series of commands to achieve a certain goal [9].

As declarative knowledge encompasses the fundamental skills required to start modelling geometries in CAD, it has traditionally been the primary focus of the dominant pedagogy in both education and industry [8]. In general, it is also the most commonly available type of resources written by CAD software providers through manuals and documentation. Hence, certification exams, offered by different CAD software providers, are mostly designed to test their users' level of declarative knowledge. While studying for such certification exams is valuable to a student's CAD learning process [21], the exams themselves also cannot fully reflect a user's industrial experience and proficiency in CAD [22].

Meanwhile, procedural knowledge in CAD determines the speed and efficiency in CAD [9], and it is often the main empirical difference between expert and novice designers [23]. Modelling strategies can be highly cognitive, causing the teaching of procedural knowledge with a traditional pedagogical approach far from ideal [24]. Instead, procedural knowledge in CAD should be taught through a cognitive apprenticeship approach, where students are able to observe and learn the high-order design thinking (i.e., the cognitive behaviors) from the teacher [25,26].

Like in most fields, deliberate practice has an important role in the acquisition of expert performance [10,27]. In recent years, a new concept of "speed modelling" has emerged and become popular in the CAD user community [28,29]. Competitions are often held in a tournament setting, where participants compete against each other by trying to model the same mechanical part in CAD, given the same manufacturing drawings of the part. With the goal of re-modelling the exact geometries of the reference part in CAD, the participant who completes the task in the shortest time wins. As one can imagine, to excel in such competitions, a high-level mastery of both declarative and procedural knowledge is indispensable.

While growing to be a design expert in CAD requires deliberate practice to develop the expertise, effective skill developments heavily draw upon a person's motivational resources, not mindless repetition [30]. Presumably, a platform that enables speed modelling challenges has a great potential in fostering deliberate practice in CAD by providing a library of models that require different skills, while tracking an individual learner's progress.

2.2 Design Research in CAD

As people continue exploring ways to advance methods, tools, and outcomes in design, researchers need to uncover and characterize the cognitive thinking behind the observable design behaviors [16]. To standardize and compare design research in the field, there are ontologies that try to capture the underlying regularities among all designs and design processes, representing them in a uniform way [31,32]. When researchers study design actions in CAD specifically, these ontologies are often applied [33,34].

Some design researchers study the novice-expert difference in behaviors and the characterization of expertise in design. It is commonly found that students and expert engineers differ the most in problem scoping and information gathering [10,35]. However, design fixation is a common difficulty encountered in engineering design, especially towards the end of the design process [36,37]. Early-stage use of CAD is particularly found to be linked with less novel designs, possibly the main cause of premature fixation [38]. While the use of CAD tools enhances visualization and communication, it can also lead to circumscribed thinking and bounded ideation [39].

As the complexity of design increases, modern industrial engineering designs are typically completed in teams, increasing the importance of effective collaboration and management in CAD and product development [40]. With the emergence of cloud-CAD platforms, synchronous collaboration is enabled in CAD, along with research questions on the affordance of collaborative CAD. Recent studies examine both the team efficiencies and effectiveness of synchronous collaboration in CAD, for both short- and long-term design projects [41–45].

A parallel branch of design research in CAD takes a geometric perspective. Many studies in the literature primarily focus on data extraction from public CAD models [46–49]. The preparation and the availability of these large CAD datasets enable researchers to use machine learning models to learn from existing CAD products. Then, tools are often built to assist human designers in CAD. For examples, some research explores the automatic generation of CAD sketches given images [50], the automatic mating of CAD assemblies [48,51], and the 3D object reconstruction from 2D images [52].

Data collection has been a widely encountered challenge for design research in CAD. It is expensive to run design experiments due to recruitment of participants (especially professional designers) and time requirements from researchers [10]. Furthermore, data analysis can be even more resource-intensive [16]. Further, collecting large geometric dataset on CAD models remains difficult and resource-consuming, especially with the extensive construction process of the models [46,47]. Our proposed framework is designed to collect both cognitive and geometric design data unobtrusively as users design in CAD, with minimal maintenance of the platform required. This enables design research through a big data approach.

2.3 Design Informatics in CAD

Informatics collection via CAD has long been a popular topic of research [53]. The use of data (i.e., informatics) in production research has a history of improving the processes, systems, products, and management of manufacturing production [54]. Meanwhile, data mining (or logging) in CAD for the study of design thinking has grown in popularity, given the challenges in research efficiency imposed by traditional research methods [16].

While CAD software is traditionally designed to store the final product design, research in engineering design continues exploring potential methods for capturing design activities unobtrusively, while users design in the CAD environment [33,55]. With a large amount of data being recorded in the back end, algorithms have been designed to automatically analyze the sequential patterns of design activities and extract meaningful design operations [56]. Others utilize advanced machine learning models (e.g., Markov chain [33,57], hidden Markov model [58], clustering [33,59], or deep learning [60–63]) to study large-scale datasets. Meanwhile, CAD logging is often combined with external devices (e.g., electroencephalography, mouse and eye tracking, screen video recording) to study design behaviors in greater detail [18]. Emerging analytical frameworks also explore analyses of multi-user CAD platforms, where multiple designers collaborate and work synchronously on the same design document [64].

Being able to mine design informatics at a large scale provides great resources for design researchers, but only a few types of CAD platforms support data collection to various extents. In relatively smaller scales, research teams can build and distribute their own CAD software, such that it is designed with very specific research objectives and measurable performance evaluation metrics with desired data collection capabilities [33]. Most commercially available CAD software also provides an application programming interface (API), which allows users (or researchers) to interact with the software and retrieve data from the software in an algorithmic approach (see work in [18,34,55,65] for example). However, as most traditional CAD software requires local installation, this approach is not ideal for efficient distribution of the research study and automatic data collection, where some manual operations are still unavoidable to retrieve and compile the collected data.

Recent advancements in graphics and computing have allowed CAD platforms to be natively developed in the cloud, running with a software-as-a-service (SaaS) model [4,66]. While a cloud-CAD platform is used in the same way as most traditional CAD software, users now access all CAD features and data through an internet browser, where no local installations are required. This potentially leads to increased scalability of research studies by enabling data mining of design activities through the web. The proposed framework in this paper aims to be one of the first attempts in exploring this potential benefit of cloud-CAD platforms.

3. FRAMEWORK DESIGN

While design education and design research are both important fields of study with potential for advancement, they both benefit and require support from each other. To connect the two fields with the use of design informatics, we constructed a framework which was productized as the *CAD Challenges* web application [67,68]. An overview of the framework is presented in Figure 1, and more details are discussed in this section.

The *CAD Challenges* app is built to integrate directly with Onshape [69], a cloud-based parametric CAD software which serves as the main platform of our framework. The application is

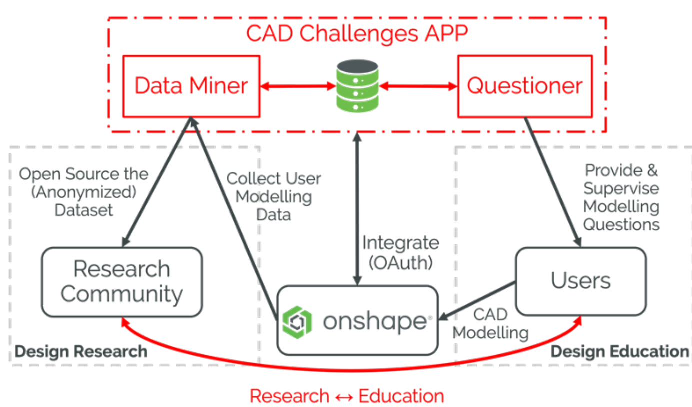


Figure 1. Overview of the framework.

composed of two parts, namely the *Questioner* and the *Data Miner*, which exchange and store information through a common database. While the *Questioner* provides and supervises all CAD modelling challenges for the users, the *Data Miner* collects and anonymizes all the user-generated modelling data for the research studies.

3.1 The CAD Platform

As the framework is built upon the goal of specifically supporting CAD education and research, a CAD platform is first required. Onshape is a parametric MCAD platform that runs in a web browser with a SaaS model. Its cloud-native architecture is necessary to the framework and provides many additional benefits.

Any user can subscribe to the *CAD Challenges* app from the Onshape App Store for free without installation requirements. This significantly lowers the barriers of entry to CAD modelling, and it enables scalable distribution of the service provided by both Onshape and our framework.

Onshape also provides a rich Representational State Transfer (REST) API that enables integration to the CAD platform through standard Hypertext Transfer Protocol Secure (HTTPS) web requests, which transfer information in JavaScript Object Notation (JSON) format [70]. With proper integration through Onshape's OAuth authorization framework, a third-party application can be granted access to users' CAD documents for a limited amount of time [70]. Once authenticated by the user, the application (1) synchronously initiates and evaluates modelling questions for the users with the *Questioner*, and (2) asynchronously collects design data from the users after completion of a question with the *Data Miner*.

3.2 The Questioner

The *Questioner* portion of the *CAD Challenges* app handles all front-end user interactions of the framework. An overview of the user experience is presented in Figure 2. Once a user subscribes to the *CAD Challenges* app in Onshape's App Store, they can directly open the app as a right-panel application in any Onshape documents that they have editing access to.

On the Index page, all questions available for attempt are listed in categories of difficulties. A list of potential question types that can be supported by the framework is presented in Table 1, which are all commonly tested in the certification exams offered by Onshape and other CAD software [71–73]. Once a question is chosen to start or retry (if the question was previously completed by the user), the user is redirected to the Modelling page. This is the page where details of the question (e.g., manufacturing drawings) are presented. The user also has the option to open the question in a new tab (the Reference document in Figure 2), where the drawings and a derived version of the source reference part(s) are shared in a public Onshape document. With the derived part(s), users can only view and measure the geometries of the reference, but the feature lists (i.e., the construction process of the original parts) are hidden and inaccessible to the public.

With a question presented, users attempt to model the given parts in their own Onshape document, with the goal of completing the task with the shortest time (i.e., a speed-modelling challenge). As a user submits the completed model for evaluation, API requests are automatically called by the *Questioner* to retrieve properties from the user's Onshape models, which are then compared with the properties of the reference model. For example, the mass, volume, surface area,

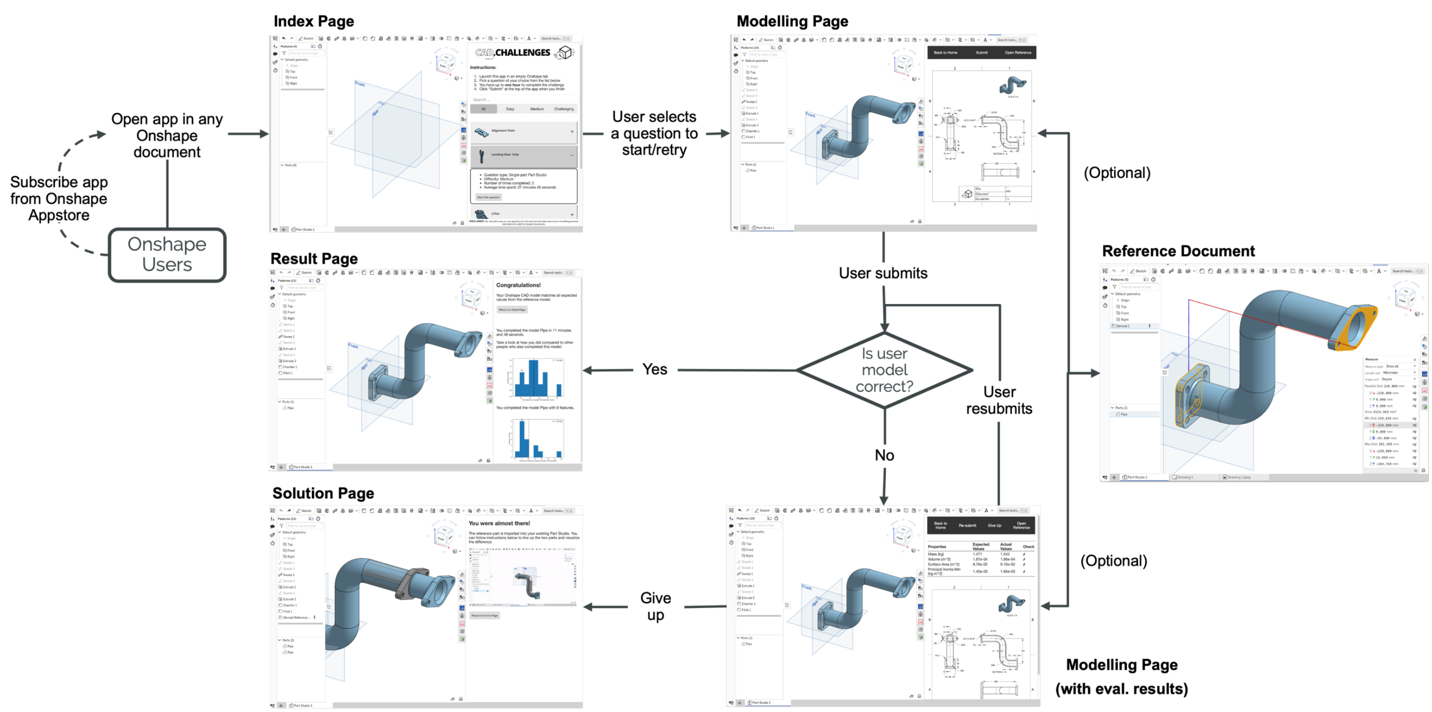


Figure 2. Overview of the user experience.

Table 1. Question types supported by the framework.

Question type	Modelling space in Onshape	Description	Abilities testing
Single-part modelling	Part Studio	The 2D manufacturing drawing(s) of a part is presented, and the user needs to re-create the 3D part in CAD	Mastery of the parametric modelling features available in the CAD software and efficient modelling strategies
Multi-part modelling	Part Studio	Drawings of each part in the model are provided with extra relational information between parts	The use of reference geometries from existing parts to expedite the creation of new parts
Mating	Assembly	The degrees and freedom of and relations between parts need to be fully defined using mating features in CAD	Mastery of mating features and modelling of the dynamic relationship and interaction between parts
Multi-step modelling	Part Studio & Assembly	After the completion a standard part modelling or mating question, modifications to various parameters are then tasked	Design of model flexibility and robustness for design iterations

and the principal inertia of the part are retrieved for part modelling questions; mating questions evaluate the principal inertia of the entire assembly. If any one of the assessed properties does not match the reference value, within a certain range of tolerance, the user will be prompted to revise their model with the evaluation results presented. Otherwise, if the model is correct, the Result page will present (1) the time spent by the user, (2) a distribution plot of the time spent by all other users who attempted this question, and (3) any other question type-specific statistics and distribution (e.g., the number of features used). For all distribution plots, users can visualize how others achieved in this question, as well as their relative position in the distribution.

For users who struggle with the question after numerous consecutive failed submission attempts, a "give-up" option is also available. By giving-up on a question, the reference parts will be automatically imported into the user's working document through the derive feature using API requests. Instructions on transforming the reference part are provided to align it with the working part, such that the difference between the two parts can be visualized. Again, the original feature list of the reference part is protected from access. Users are then encouraged to utilize this comparison and look for mistakes made during the modelling process. Additional attempts on the question are always allowed and encouraged to improve their modelling skills and strategies.

3.3 The Data Miner

The *Data Miner* portion of the *CAD Challenges* app handles all data collection of users' design product and process in the background. A list of data types that can be collected from Onshape is presented in Table 2. In the app, we collect such data at the point of the document's history when the user (1) successfully completes a question, (2) gives up on a question, and (3) submits the first failed attempt, if any. Data collection at these three occurrences enables rich informatics to (1) analyze different types of decision making in design, (2) compare properties between submission attempts for the modelling session, and (3) compare properties of the submitted product across different users. More potential research opportunities enabled by the retrievable data are further discussed in section 4.

For every change made in an Onshape document, an immutable *microversion* is automatically created and logged by Onshape in the document's history [74]. At every point of data collection occurrence, the microversion of that state of the

Table 2. Retrievable design research data from Onshape.

Type	Data	Descriptions	Information
Design product	Feature list definition	The list of CAD features that construct the model in a Part Studio and mating features that constraint the parts in an Assembly	The semantic naming and parameters specified for the features
	Geometric model reconstruction	In Part Studios only, the geometric mesh of the model can be exported in stereolithography (STL) format	Meshes are standardized format for 3D geometric analysis, suitable for applications such as machine learning
	Shaded view image capture	Shaded views of the entire modelling space can be captured from a specific view angle	A snapshot of the design product and its orientation
Design process	Microversion history	Every microversion entry records a brief description of the action and the timestamp of occurrence	A summary of all editing actions in the document

Table 3. Evaluation of our framework using the Large Data for Design Research framework [17].

Design principle	Definition as in [17]	Evaluation of our framework
Simulation scope	The design platform should simulate a wide range of recordable actions within a domain(s) where design can be undertaken.	Onshape is a commercially available mechanical CAD software; it provides a professional platform for the design of mechanical parts and assemblies.
Unobtrusive logging	The design platform should surreptitiously log designers' actions at a granular level throughout their design process.	All design data are retrieved using Onshape's REST API through HTTPS requests. Actions are not required from the users, except the initial authorization of access.
Design action types	The design platform should consist of actions for creating/ transforming design artifact(s) and analyzing design artifact(s) relevant attributes.	All user actions that result in a change of the design artifact(s) are recorded as the microversion history, and details of the design artifacts can be recorded in the form of parametric feature definition, screen capture, and geometric mesh.

document is logged and stored in the database. The data collection tasks are executed by a different processor as background jobs so the user can continue interacting with the web app to either continue working on a failed submission or start a new question. As all microversions of an Onshape document are read-only to the users, properties at historical states of the document are immutable, even if new changes are made to the working version of the same document. This unique feature allows us to non-intrusively query information stored in a specific microversion of the document.

4. RESULTS AND DISCUSSION

The goal of this framework is two-fold: facilitate design education and support design research. The *CAD Challenges* app enables users to deliberately practice their parametric modelling skills in CAD, while also asynchronously collects a large amount of data to be used for design research. An infinite number of questions can theoretically be made available, so any modeling skill can be practiced, or research question can be explored, with the framework.

We first evaluate our framework with an existing research framework in this section. Then, we outline some existing research studies in the literature that can be replicated utilizing this large-scale dataset. We then discuss the implication of these potential research results for improving design education. With preliminary analysis presented, limitations and future work directions are also discussed.

4.1 Framework Evaluation

Based on the consensual definition summarized by Mauro et al. [75], our data mining approach generates information assets that satisfy all three characteristics of big data:

- *High volume*: a large volume of design data can be collected from all user attempts through this integrated application in a commercially available CAD software;
- *High velocity*: data collection asynchronously runs as background tasks of the application, which is initiated as soon as a user completes a question attempt; and
- *High variety*: design data of attempts from different users can be collected for a large question bank.

At the same time, our framework also satisfies the three design principles set by Schimpf and Goldstein [17], as summarized in Table 3. Using the big data approach, this allows us to enable three affordances for research and education: capturing design activities, context setting and operationalization, and research design scalability [17].

4.2 Contribution to Design Research

With the capability of collecting a large volume of data, the deployment of this framework enables the replication of many research studies from the literature, but at a larger scale. With research methods that utilize machine learning, this framework provides an especially valuable platform to collect large-scale data for training.

4.2.1 Research on Design Product

Given a final CAD model created for a modelling task, there are mainly two types of data that may be interesting for research: the parametric feature list used to construct the model and the model itself. To start, some simple statistical analyses can be performed, such as model comparisons between the first failure and the final submission, the orientation of CAD models with respect to the default planes, and the number, types, and complexity of the features used.

As the length of the feature list grows for large models, the organization of the feature list becomes increasingly important. With the parameters and other metadata (e.g., custom names) of the features, research can be performed around the use of semantics and its capture of design intent [76,77]. Especially for questions that require consecutive modifications to the CAD model, the organization of the feature list likely has an impact on the model's ease of iteration, which theoretically correlates to design efficiency.

While different feature lists also generate different parent-child interdependencies between features, certain modelling strategies can effectively lead to better design robustness and flexibility [12]. The analysis of feature dependencies may also unveil the modelling intents of the CAD model [65]. Since our proposed framework collects feature list data from users working on some common modelling tasks, we enable comparative analyses to test the correlation between different modelling strategies and design efficiency in CAD.

To characterize the modelling strategies used for a design product in CAD, the construction of the model can be represented with a workflow graph, as proposed by Chang et al. [78]. Using our framework, we can retrieve the geometric representation (e.g., point cloud, mesh) of every intermediate state of the final CAD model's feature list regeneration process, represented with graph nodes. This workflow graph can encode all the alternative modelling strategies used for the same modelling task, where the weights of the edges may represent the number of observations on the use of each strategy.

4.2.2 Research on Design Process

In addition to the final product, we get a high-level view of the actions performed by the users as they complete the modelling tasks with a complete list of microversion descriptions collected. These actions include not only the creation of features that are recorded in the final product, but also actions such as the edit and deletion of features as they model. This presents a different view of the modelling strategy, better reflecting users' cognitive decision making during the design process.

Typically, different states of a CAD model throughout its design process can be represented as graph nodes, and the design actions that lead from one state of the model to another state can be characterized as directed graph edges between the nodes. The graph representation of the design process in CAD enables analyses using standard graph-based algorithms. For example, in the product model roadmap established by Jin and Ishino [56], the design process of a product can be presented as a graph to visualize the exploration of the design space. Then, algorithms are proposed to analyze and identify meaningful design operations between states of the product.

Focusing on the sequential actions performed by the users in CAD, various methods of analysis are also available in the literature. For example, it is suggested that different types of users would present different types of behaviors in CAD action usage, where clustering analysis can be used to group users of the same type for analysis [33,59]. Exploring the transitions between consecutive actions, different types of Markov models can be used to characterize design behaviors in a different perspective [33,57,58]. Further, training machine learning models on sequential CAD actions also creates interesting research opportunities such as predicting the next CAD feature or action [60–62], and predicting the ultimate success of a design process at its early stage [79].

Efforts have also been made to visualize the design process by plotting design actions on a time-series [35,80,81]. Visualization of the design process allows both the researchers and the users to better understand and reflect on the modelling process. Meanwhile, it is also useful to develop metrics for analysis with a grounded approach.

4.3 Implication to Design Education

As we publish the productized version of this framework through Onshape's App Store, we provide a platform for practicing parametric modelling in a commercially available CAD software. As a free integrated app to a cloud-CAD platform, there is a minimal barrier of entry (e.g., no need for software downloads and version upgrades, relatively low computer hardware requirements) for learners. While deliberate practice is crucial on the way to becoming an expert in a field [27], all training and practicing can be done asynchronously, without the need of supervision from a teacher. In addition, being able to "rank" higher than other students who also completed the same question in terms of time spent creates some incentives for the students, and this motivational resource can be an important component in long-term deliberate practicing [30].

Other parts of the framework design embed additional educational resources. In most traditional CAD certification exams, students are often provided with 2D manufacturing drawings only, the format of CAD information that is traditionally stored and transferred in industry. Our framework also allows a view-only 3D reference part to the students, without giving out the construction history of the parts. This can extend the practice of parametric modelling skills in CAD to (1) reading manufacturing drawings with expected 3D solid and (2) spatial reasoning skills. Both skills were found to correlate to one's ability to design in CAD [6,7,25]. While enabling engagement with the traditional 2D drawings widely used in manufacturing, we also allow students to engage with model-based definition, an emerging practice in modern product lifecycle management [82].

While we encourage all students to keep working through a question and finish with a successful submission, we also allow a student to give up on a question that they are struggling on, or a question that the student's CAD expertise is not yet prepared for. Additional assistance can be provided by the framework, such as automatically importing the reference part(s) to the user's working document using a derive feature in Onshape. This allows students to compare the two models geometrically and learn from the mistakes that they made. Raising the awareness and triggering the cognitive corrective actions in a timely manner through failures in the learning process can be beneficial to students' ultimate success in education [83].

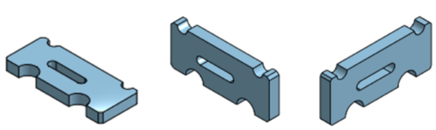
As the data collected from this framework enables design research, some research findings can also directly benefit CAD education. This framework enables the implementation of the concept of crowd learning [19]. With many (unconnected) people performing the same task, crowd learning captures the knowledge, potential, and expertise of the "crowd" of people. For example, with research findings that utilize the collected data from users' question attempts, the optimal modelling strategies for any specific questions can be algorithmically concluded. Presented with graph nodes and edges (see the workflow graph [78] as an example), students can self-reflect on their modelling process and compare it to the most popular modelling workflow. This can be a form of cognitive apprenticeship learning [25], where students can more effectively develop their procedural knowledge in CAD. However, the potential risk of herding in crowd learning should also be noted [84].

4.4 Preliminary Research Analysis

With the productized *CAD Challenges* app publicly launched in the Onshape App Store, we started collecting data as the app was adopted by users. To demonstrate the feasibility of the proposed framework, we present the preliminary analysis from examining the design data collected from one question (with $n = 99$ successful attempts). The selected question was designed to be easy and only require standard parametric sketches and features.

We can visualize the difference in the orientation of each users' design with respect to the default planes with a shaded view captured after every successful question attempt. As shown in Figure 3, with images capturing an isometric view from the same view angle, we observed 82 attempts modeled as shown in Figure 3(a) on the default Top plane, 16 attempts modeled as shown in Figure 3(b) on the default Front plane, and 1 attempt modeled as shown in Figure 3(c) on the default Right plane. Although the choice of base plane did not lead to significant efficiency advantage in this case, its relationship with other design characteristics is worth future investigation.

We also observed that two modeling approaches were used by 88 of the 99 successful attempts analyzed. 17 attempts used only two parametric features (one sketch and one extrude) to complete the challenge, indicating that all the geometry was defined in the initial sketch. 71 attempts used three parametric features (one sketch, one extrude, and one chamfer) to complete the challenge, indicating that their sketch had sharp corners which were modified in 3D with the chamfer feature. Respectively, the two approaches required 17 and 11 minutes on average, a difference that was statistically significant ($t = 2.41, p = .018$). This difference seems to indicate that it is more efficient to modify corners in 3D geometry rather than defining modified corners in 2D sketches.



(a) (b) (c)

Figure 3. Orientations of the design product from different attempts of the same question.

4.5 Limitations and Future Work

There are certain limitations to this proposed framework. As Onshape does not collect any personal information from its users, we cannot directly collect, and hence cannot correlate any of the research results to factors such as the gender, race, age, and academic background of the users. Similarly, we have no access to any factors outside the CAD platform. For example, some users may use hand drafting on papers to aid their dissections of the problem. Any research conducted with this framework will inevitably need to account for these uncontrolled factors in the dataset. However, future work can consider integrating external research tools to the CAD system, capturing additional user behaviors during the design process [18].

Future work can explore the design of open-ended tasks in CAD. Due to the need for model evaluation at submission, all question types that the proposed framework currently supports are closed-ended, which do not necessarily reflect the CAD use of an engineer that focuses on conceptual product design [10]. However, model evaluation is not necessarily required if the purpose of deployment is mainly for research data collection. Similarly, future work can also enable collaboration, a unique feature enabled by Onshape as a cloud-CAD platform, where multiple users can work on the same design. Research in teamwork will further explore the affordance of synchronous collaboration in emerging cloud-CAD platforms [44].

5. CONCLUSION

In this paper, we propose an informatics framework that enables both parametric modelling practice for design education and data collection for design research in CAD. With the productized version of the framework, namely the *CAD Challenges* App, integrated to Onshape, a commercially available cloud-CAD platform, users and students can practice CAD modelling with parametric features, skills that are transferrable across CAD software. Meanwhile, their design data (i.e., the informatics) can be collected asynchronously and unobtrusively for design research. Enabling design research of a potentially large CAD user community fosters the concept of crowd learning, where design behaviors can be studied with a big data approach. Research findings can also benefit the education community, providing valuable learning resources for students to self-reflect on their modelling processes.

While this framework provides a platform for research data collection, it also serves as a scalable template to be adapted for more customized or controlled experiments. For example, researchers can carefully design their CAD modelling questions based on specific research hypotheses. With design data collected using this framework, differences in modelling behaviors and design solutions may be observed from assigning questions to different groups of participants.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support from everyone who provided feedback and assistance to the productization of the framework.

REFERENCES

[1] Stark, R., 2022, "Major Technology 1: Computer Aided Design—CAD," *Virtual Product Creation in Industry*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 113–138.

[2] Mitchell, W. J., 1995, "CAD as a Social Process," *Sixth International Conference on Computer-Aided Architectural Design Futures*, pp. 7–9.

[3] Robertson, D., and Allen, T. J., 1993, "CAD System Use and Engineering Performance," IEEE Trans Eng Manag, **40**(3), pp. 274–282.

[4] Zissis, D., Lekkas, D., Azariadis, P., Papanikos, P., and Xidias, E., 2017, "Collaborative CAD/CAE as a Cloud

Service," International Journal of Systems Science: Operations & Logistics, **4**(4), pp. 339–355.
[5] Gerhard, D., Köring, T., and Neges, M., 2023, "Generative Engineering and Design – A Comparison of Different Approaches to Utilize Artificial Intelligence in CAD Software Tools," *Product Lifecycle Management. PLM in Transition Times: The Place of Humans and Transformative Technologies*, F. Noël, F. Nyffenegger, L. Rivest, and A. Bouras, eds., Springer, Cham, pp. 206–215.
[6] Hsi, S., Linn, M. C., and Bell, J. E., 1997, "The Role of Spatial Reasoning in Engineering and the Design of Spatial Instruction," Journal of Engineering Education, **86**(2), pp. 151–158.
[7] Sorby, S. A., 2009, "Educational Research in Developing 3-D Spatial Skills for Engineering Students," Int J Sci Educ, **31**(3), pp. 459–480.
[8] Chester, I., 2008, "3D-CAD: Modern Technology--Outdated Pedagogy?," Des Technol Educ, **12**(1).
[9] Chester, I., 2007, "Teaching for CAD Expertise," Int J Technol Des Educ, **17**(1), pp. 23–35.
[10] Cross, N., 2004, "Expertise in Design: An Overview," Des Stud, **25**(5), pp. 427–441.
[11] Bodein, Y., Rose, B., and Caillaud, E., 2013, "A Roadmap for Parametric CAD Efficiency in the Automotive Industry," Computer-Aided Design, **45**(10), pp. 1198–1214.
[12] Camba, J. D., Contero, M., and Company, P., 2016, "Parametric CAD Modeling: An Analysis of Strategies for Design Reusability," Computer-Aided Design, **74**, pp. 18–31.
[13] Rynne, A., and Gaughran, W., 2007, "Cognitive Modelling Strategies For Optimum Design Intent In Parametric Modelling (Pm).," *2007 Annual Conference & Exposition Proceedings*, ASEE Conferences, Honolulu, pp. 12.366.1-12.366.15.
[14] Otey, J., Company, P., Contero, M., and Camba, J., 2014, "A Review of the Design Intent Concept in the Context of CAD Model Quality Metrics," *ASEE Annual Conference & Exposition Proceedings*, ASEE Conferences, pp. 24.100.1-24.100.10.
[15] Bodein, Y., Rose, B., and Caillaud, E., 2014, "Explicit Reference Modeling Methodology in Parametric CAD System," Comput Ind, **65**(1), pp. 136–147.
[16] Dinar, M., Shah, J. J., Cagan, J., Leifer, L., Linsey, J., Smith, S. M., and Hernandez, N. V., 2015, "Empirical Studies of Designer Thinking: Past, Present, and Future," Journal of Mechanical Design, **137**(2).
[17] Schimpf, C., and Goldstein, M. H., 2022, "Large Data for Design Research: An Educational Technology Framework for Studying Design Activity Using a Big Data Approach," Frontiers in Manufacturing Technology, **2**.
[18] Sivanathan, A., Lim, T., Ritchie, J., Sung, R., Kosmadoudi, Z., and Liu, Y., 2015, "The Application of Ubiquitous Multimodal Synchronous Data Capture in CAD," Computer-Aided Design, **59**, pp. 176–191.
[19] Kalisz, D. E., 2016, "Crowd Learning: Innovative Harnessing the Knowledge and Potential of People," *Innovative Management Education Pedagogies for Preparing Next-Generation Leaders*, IGI Global, pp. 55–74.
[20] Lang, G. T., Eberts, R. E., Gabel, M. G., and Barash, M. M., 1991, "Extracting and Using Procedural Knowledge in a CAD Task," IEEE Trans Eng Manag, **38**(3), pp. 257–268.
[21] Webster, R., and Ottway, R., 2018, "Computer-Aided Design (CAD) Certifications: Are They Valuable to Undergraduate Engineering and Engineering Technology Students?," Journal of Engineering Technology, **35**(2), pp. 22–32.
[22] Trent, D. L., 2011, "Efficacy of Computer Aided Drafting (CAD) Certifications," Dissertation, Old Dominion University.
[23] Kavakli, M., and Gero, J. S., 2003, "Strategic Knowledge Differences between an Expert and a Novice Designer," *Human Behaviour in Design*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 42–52.
[24] Garikano, X., Garmendia, M., Manso, A. P., and Solaberrieta, E., 2019, "Strategic Knowledge-Based Approach for CAD Modelling Learning," Int J Technol Des Educ, **29**(4), pp. 947–959.
[25] Huang, T. C., Chen, M. Y., and Lin, C. Y., 2019, "Exploring the Behavioral Patterns Transformation of Learners in Different 3D Modeling Teaching Strategies," Comput Human Behav, **92**, pp. 670–678.
[26] Nieusma, D., Malazita, J., Krauss, L., Ukleja, A., and Andrews, T., 2018, "From Learning to CAD to CADing to Learn: Teaching the Command, Strategic, and Epistemic Dimensions of CAD Software," *ASEE Annual Conference & Exposition Proceedings*, ASEE Conferences, Salt Lake City.
[27] Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C., 1993, "The Role of Deliberate Practice in the Acquisition of Expert Performance.," Psychol Rev, **100**(3), pp. 363–406.
[28] Valerio, R., 2023, "The Ultimate CAD Battle: Witnessing Speed 3D Modeling in Record Time" [Online]. Available: https://www.onshape.com/en/blog/speed-3d-modeling-record-time. [Accessed: 09-Feb-2023].
[29] Schnaars, T., "3D CAD Speed Modeling Championship Tournament" [Online]. Available: https://tootalltoby.com/. [Accessed: 09-Feb-2023].
[30] Lehmann, A. C., and Ericsson, K. A., 1997, "Research on Expert Performance and Deliberate Practice: Implications for the Education of Amateur Musicians and Music Students.," Psychomusicology: A Journal of Research in Music Cognition, **16**(1–2), pp. 40–58.
[31] Gero, J. S., and Kannengiesser, U., 2014, "The Function-Behaviour-Structure Ontology of Design," An

Anthology of Theories and Models of Design, pp. 263–283.
[32] Sim, S. K., and Duffy, A. H. B., 2003, “Towards an Ontology of Generic Engineering Design Activities,” Res Eng Des, **14**(4), pp. 200–223.
[33] Rahman, M. H., Schimpf, C., Xie, C., and Sha, Z., 2019, “A Computer-Aided Design Based Research Platform for Design Thinking Studies,” Journal of Mechanical Design, **141**(12), pp. 1–12.
[34] Gopsill, J., Snider, C., Shi, L., and Hicks, B., 2016, “Computer Aided Design User Interaction as a Sensor for Monitoring Engineers and the Engineering Design Process,” Proceedings of International Design Conference, DESIGN, **DS 84**, pp. 1707–1718.
[35] Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., and Saleem, J., 2007, “Engineering Design Processes: A Comparison of Students and Expert Practitioners,” Journal of Engineering Education, **96**(4), pp. 359–379.
[36] Linsey, J. S., Tseng, I., Fu, K., Cagan, J., Wood, K. L., and Schunn, C., 2010, “A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty,” Journal of Mechanical Design, **132**(4), pp. 0410031–04100312.
[37] Zhang, H. Z., Xie, C., and Nourian, S., 2018, “Are Their Designs Iterative or Fixated? Investigating Design Patterns from Student Digital Footprints in Computer-Aided Design Software,” Int J Technol Des Educ, **28**(3), pp. 819–841.
[38] Häggman, A., Tsai, G., Elsen, C., Honda, T., and Yang, M. C., 2015, “Connections Between the Design Tool, Design Attributes, and User Preferences in Early Stage Design,” Journal of Mechanical Design, **137**(7).
[39] Robertson, B. F., Walther, J., and Radcliffe, D. F., 2007, “Creativity and the Use of CAD Tools: Lessons for Engineering Design Education from Industry,” Journal of Mechanical Design, Transactions of the ASME, **129**(7), pp. 753–760.
[40] Marion, T. J., and Fixson, S. K., 2021, “The Transformation of the Innovation Process: How Digital Tools Are Changing Work, Collaboration, and Organizations in New Product Development*,” Journal of Product Innovation Management, **38**(1), pp. 192–215.
[41] Phadnis, V., Arshad, H., Wallace, D., and Olechowski, A., 2021, “Are Two Heads Better Than One for Computer-Aided Design?,” Journal of Mechanical Design, **143**(7), pp. 1–38.
[42] Deng, Y., Marion, T., and Olechowski, A., 2022, “Does Synchronous Collaboration Improve Collaborative Computer-Aided Design Output: Results From a Large-Scale Competition,” *Volume 6: 34th International Conference on Design Theory and Methodology (DTM)*, American Society of Mechanical Engineers.
[43] Stone, B., Salmon, J. L., Hepworth, A. I., Red, E., Killian, M., La, A., Pedersen, A., and Jones, T., 2017, “Methods for Determining the Optimal Number of Simultaneous Contributors for Multi-User CAD Parts,” Comput Aided Des Appl, **14**(5), pp. 610–621.
[44] Marion, T., Olechowksi, A., and Guo, J., 2021, “An Analytical Framework For Collaborative Cloud-Based CAD Platform Affordances,” Proceedings of the Design Society, **1**, pp. 375–384.
[45] Zhou, J. (Janice), Phadnis, V., and Olechowski, A., 2021, “Analysis of Designer Emotions in Collaborative and Traditional Computer-Aided Design,” Journal of Mechanical Design, **143**(2).
[46] Koch, S., Matveev, A., Jiang, Z., Williams, F., Artemov, A., Burnaev, E., Alexa, M., Zorin, D., and Panozzo, D., 2019, “ABC: A Big CAD Model Dataset For Geometric Deep Learning,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9601–9611.
[47] Willis, K. D. D., Pu, Y., Luo, J., Chu, H., Du, T., Lambourne, J. G., Solar-Lezama, A., and Matusik, W., 2020, “Fusion 360 Gallery: A Dataset and Environment for Programmatic CAD Construction from Human Design Sequences,” ACM Trans Graph, **40**(4), pp. 1–24.
[48] Jones, B., Hildreth, D., Chen, D., Baran, I., Kim, V. G., and Schulz, A., 2021, “AutoMate: A Dataset and Learning Approach for Automatic Mating of CAD Assemblies,” ACM Trans Graph, **40**(6), pp. 1–18.
[49] Seff, A., Ovadia, Y., Zhou, W., and Adams, R. P., 2020, “SketchGraphs: A Large-Scale Dataset for Modeling Relational Geometry in Computer-Aided Design.”
[50] Ganin, Y., Bartunov, S., Li, Y., Keller, E., and Saliceti, S., 2021, “Computer-Aided Design as Language,” Adv Neural Inf Process Syst, **34**, pp. 5885–5897.
[51] Willis, K. D. D., Jayaraman, P. K., Chu, H., Tian, Y., Li, Y., Grandi, D., Sanghi, A., Tran, L., Lambourne, J. G., Solar-Lezama, A., and Matusik, W., 2021, “JoinABLe: Learning Bottom-up Assembly of Parametric CAD Joints.”
[52] Fan, H., Su, H., and Guibas, L., 2017, “A Point Set Generation Network for 3D Object Reconstruction from a Single Image,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 605–613.
[53] Subrahmanian, E., and Rachuri, S., 2008, “Special Issue on ‘Engineering Informatics,’” J Comput Inf Sci Eng, **8**(1).
[54] Kuo, Y.-H., and Kusiak, A., 2019, “From Data to Big Data in Production Research: The Past and Future Trends,” Int J Prod Res, **57**(15–16), pp. 4828–4853.
[55] Sung, R., Ritchie, J. M., Rea, H. J., and Corney, J., 2011, “Automated Design Knowledge Capture and Representation in Single-User CAD Environments,” Journal of Engineering Design, **22**(7), pp. 487–503.
[56] Jin, Y., and Ishino, Y., 2006, “DAKA: Design Activity Knowledge Acquisition through Data-Mining,” Int J Prod Res, **44**(14), pp. 2813–2837.
[57] McComb, C., Cagan, J., and Kotovsky, K., 2017, “Capturing Human Sequence-Learning Abilities in

Configuration Design Tasks Through Markov Chains," Journal of Mechanical Design, **139**(9).
[58] McComb, C., Cagan, J., and Kotovsky, K., 2017, "Mining Process Heuristics from Designer Action Data Via Hidden Markov Models," Journal of Mechanical Design, Transactions of the ASME, **139**(11).
[59] Leonardo, K., and Olechowski, A., 2021, "Identifying Computer-Aided Design Action Types From Professional User Analytics Data," *Volume 3B: 47th Design Automation Conference (DAC)*, American Society of Mechanical Engineers.
[60] Rahman, M. H., Yuan, S., Xie, C., and Sha, Z., 2020, "Predicting Human Design Decisions with Deep Recurrent Neural Network Combining Static and Dynamic Data," Design Science.
[61] Rahman, M. H., Xie, C., and Sha, Z., 2021, "Predicting Sequential Design Decisions Using the Function-Behavior-Structure Design Process Model and Recurrent Neural Networks," Journal of Mechanical Design, Transactions of the ASME, **143**(8).
[62] Gao, W., Zhang, X., He, Q., Lin, B., and Huang, W., 2022, "Command Prediction Based on Early 3D Modeling Design Logs by Deep Neural Networks," Autom Constr, **133**, p. 104026.
[63] Raina, A., McComb, C., and Cagan, J., 2019, "Learning to Design From Humans: Imitating Human Designers Through Deep Learning," Journal of Mechanical Design, **141**(11).
[64] Deng, Y., Mueller, M., Rogers, C., and Olechowski, A., 2022, "The Multi-User Computer-Aided Design Collaborative Learning Framework," Advanced Engineering Informatics, **51**, p. 101446.
[65] Cheng, Z., Xie, Y., and Ma, Y., 2018, "Graph Centrality Analysis of Feature Dependencies to Unveil Modeling Intents," Comput Aided Des Appl, **15**(5), pp. 684–696.
[66] Wu, D., Rosen, D. W., Wang, L., and Schaefer, D., 2015, "Cloud-Based Design and Manufacturing: A New Paradigm in Digital Manufacturing and Design Innovation," Computer-Aided Design, **59**, pp. 1–14.
[67] PTC Education, "CAD Challenges," App Store - Onshape [Online]. Available: https://appstore.onshape.com/apps/Utilities/M2YDH5LM52FAL5JIWB7HUCVOFQZDQURENUPHV2Q=/description. [Accessed: 05-Mar-2023].
[68] Deng, Y., Mueller, M., and Shields, M., 2023, "CAD Challenges App," Zenodo [Online]. Available: https://doi.org/10.5281/zenodo.7897262.
[69] "Onshape | Product Development Platform" [Online]. Available: https://www.onshape.com/en/. [Accessed: 05-Feb-2023].
[70] "Onshape Developer Documentation" [Online]. Available: https://onshape-public.github.io/docs/. [Accessed: 26-Feb-2023].
[71] "Certified Onshape Professional" [Online]. Available: https://learn.onshape.com/courses/certified-onshape-professional. [Accessed: 21-Jan-2023].
[72] "SOLIDWORKS Certification Program" [Online]. Available: https://www.solidworks.com/solidworks-certification-program. [Accessed: 21-Jan-2023].
[73] "Autodesk Certification" [Online]. Available: https://www.autodesk.com/certification/all-certifications. [Accessed: 21-Jan-2023].
[74] Baran, I., 2015, "Under The Hood: How Collaboration Works" [Online]. Available: https://www.onshape.com/en/blog/under-the-hood-how-collaboration-works. [Accessed: 26-Feb-2023].
[75] De Mauro, A., Greco, M., and Grimaldi, M., 2015, "What Is Big Data? A Consensual Definition and a Review of Key Research Topics," *AIP Conference Proceedings*, American Institute of Physics Inc., pp. 97–104.
[76] Cheng, Y., He, F., Lv, X., and Cai, W., 2019, "On the Role of Generating Textual Description for Design Intent Communication in Feature-Based 3D Collaborative Design," Advanced Engineering Informatics, **39**(December 2018), pp. 331–346.
[77] Mandorli, F., Borgo, S., and Wiejak, P., 2020, "From Form Features to Semantic Features in Existing MCAD: An Ontological Approach," Advanced Engineering Informatics, **44**(September 2019), p. 101088.
[78] Chang, M., Lafreniere, B., Kim, J., Fitzmaurice, G., and Grossman, T., 2020, "Workflow Graphs: A Computational Model of Collective Task Strategies for 3D Design Software," *Graphics Interface 2020*.
[79] Singh, J., Perera, V., Magana, A. J., Newell, B., Wei-Kocsis, J., Seah, Y. Y., Strimel, G. J., and Xie, C., 2022, "Using Machine Learning to Predict Engineering Technology Students' Success with Computer-Aided Design," Computer Applications in Engineering Education, **30**(3), pp. 852–862.
[80] Xie, C., Zhang, Z., Nourian, S., Pallant, A., and Bailey, S., 2014, "On the Instructional Sensitivity of CAD Logs," International Journal of Engineering Education, **30**(4), pp. 760–778.
[81] Chen, J. Y.-H., 2021, "Development of a Novel Computer-Aided Design Experiment Protocol for Studying Designer Behaviours," University of Toronto.
[82] Alemanni, M., Destefanis, F., and Vezzetti, E., 2011, "Model-Based Definition Design in the Product Lifecycle Management Scenario," The International Journal of Advanced Manufacturing Technology, **52**(1–4), pp. 1–14.
[83] Brito, M. A., and De Sá-Soares, F., 2014, "Assessment Frequency in Introductory Computer Programming Disciplines," Comput Human Behav, **30**, pp. 623–628.
[84] Glazer, J., Kremer, I., and Perry, M., 2015, *Crowd Learning without Herding: A Mechanism Design Approach*.

