VISUALIZING 17 YEARS OF CDIO INFLUENCE VIA BIBLIOMETRIC DATA ANALYSIS

Alexandra Meikleham, Ron Hugo
Department of Mechanical and Manufacturing Engineering, University of Calgary

Aldert Kamp
Delft University of Technology, Delft, the Netherlands

Johan Malmqvist
Department of Industrial and Materials Science, Chalmers University of Technology, Gothenburg, Sweden

ABSTRACT

Bibliometric data analysis has gained popularity in recent years as an efficient means of visualizing multi-dimensional indicators of influence in communities of practice (Youtie & Shapira, 2008). Such an approach has been used to map emerging fields of research such as synthetic biology and nanotechnology (Shapira, Kwon, & Youtie, 2017; Youtie & Shapira, 2008). Using this approach, one can track citation and social network data over time to develop a deeper understanding of the influence of the CDIO initiative on engineering education publications since its inception (i.e., the past 17 years). In this paper, bibliometric data analysis will be used to examine how publications on the CDIO Initiative have evolved. Visualizations are presented using an open-source visualization tool, VOSViewer, and used to understand geographic distribution and co-authorship. A word frequency and co-occurrence analysis has been used to analyze title and abstract data over the same time period. Geographic author network analysis reveals continued growth in regional collaborations over the past seventeen years. Co-authorship by author name reveals a core community of researchers, which has diverged over time into dispersed collaboration groups. Word co-occurrence analysis of title and abstract data from Scopus reveals that design-implement and project-based learning activities have been the central topic of CDIO-related engineering education literature over this time period. An analysis of the terms “faculty competence” and “learning assessment” indicates that these topics are comparatively under-served in the literature, representing fertile research topics for practitioners. The benefit of this research is to provide insight to past development areas and opportunities for growth in the CDIO Initiative.

INTRODUCTION

Since 2000, the CDIO Initiative has grown from a collaboration between four institutions in two countries into a world-wide organization comprised of over 130 institutions. As the initiative continues to mature and expand it can be healthy to reflect on the evolution of the organization by conducting a historical review from a variety of perspectives. This can provide an understanding of past impact and identify new areas for future direction and influence. While the CDIO online library presents a rich source of bibliometric data waiting to be mined, this body of knowledge could also be considered to reflect more internally-focused dialogue. In
reflecting on future directions for the initiative, another valuable perspective could be to better understand how the initiative has influenced publishing and external communities of practice. Both approaches provide different forms of valuable information that, when taken together, can better support decision-making. For this analysis, we have probed how literature that discusses engineering education and CDIO appearing in the world’s largest database of peer-reviewed literature, Scopus (and for some analyses Web of Science), has evolved since 2002.

**VOSViewer**

VOSViewer is open-source software that uses a mapping technique called visualization of similarities. It aims to locate items in a dataset using a 2-dimensional visual, with distances between items reflecting their relatedness. The distance between any two items reflects the similarity of the items as accurately as possible, and the smaller the distance between two items the stronger the relation. This is achieved by minimizing the weighted sum of the squared distances between all pairs (van Eck & Waltman, 2011). The size of circles in a visual indicates the total number of co-occurrences, with larger circles indicating more occurrences. Colours are also used to indicate clusters (or communities), which are terms or items that are related to one another.

**METHODOLOGY**

**Data Set**

Scopus and Web of Science were used to conduct our literature search; these databases do not include proceedings available from the online library on CDIO.org, and therefore represent a mutually exclusive dataset. Scopus by Elsevier, is the world’s largest database of peer-reviewed literature. It provides access to scientific journals, books and conference proceedings in a diverse set of fields. Web of Science is another database which covers publications from as early as 1900 and contains over 90 million records. These two databases were chosen to create the data set for this analysis as they best represent a broader set of publication sources external to the CDIO initiative. Both websites are designed to provide easy export of .csv files for use in bibliometric and altmetric analyses – including title, author, and abstract fields. The VOSViewer software was also built to import and easily analyze Web of Science and Scopus corpus files, making the analysis more repeatable and consistent.

**Search Terms**

A combination of “engineering education” and “CDIO” were used as search terms for this analysis. Searching “engineering education” in Scopus returned over 700,000 results. Adding CDIO as an additional search term in all fields limited the results to 1,355 hits. In Web of Science a search on “engineering education” in topic returned 25,000+ results; adding CDIO as a co-search term in ALL fields returned 216 items. A database was then created combining the results from Scopus and Web of Science, and after removing duplicates, 1,453 unique entries were found from the period 2002-2018. The search criteria do not imply that relevant records resulted from only CDIO researchers or CDIO community members. The choice of search criteria was to reveal influence by the CDIO initiative in that the publications contain mentions to both engineering education and CDIO, which can provide some indication to broader publication trends which reference to both themes.
VOSViewer Analysis

The program is currently built to handle files from a number of popular databases, such as Web of Science and Scopus, however each database formats their citation data in a slightly different manner. It is possible to combine data from multiple databases into the corpus of one standard accepted format, however the citation data must all be homogenous for the program to work properly. For example, if Scopus formats author names as Doe, J.A. and another database formats its authors as JA Doe these will be treated as separate entities. This renders the process of analyzing data from multiple different sources simultaneously a prohibitive process for large datasets.

In deciding which database to use in this analysis it was therefore necessary to take an either-or approach between using external databases or using the CDIO.org database, as citation data is formatted differently in each. It was decided that the focus of this paper would be on trends in CDIO publishing in external databases such as Scopus and Web of Science as these are directly compatible with VoSViewer; CDIO.org library data has therefore been left for future analysis.

The analysis was broken up into several time steps, and cumulative data was used at each time step. For example, 2002-2010 analysis includes the same data from 2002-2007, with 2008-2010 added in. The justification for this was that we were interested in visualizing how the CDIO initiative has potentially influenced external network development over time. As networks are human relationships and do not dissolve after the year a paper is authored, we found it more interesting to visualize over time how the network changed as new relationships were added (rather than only looking at new relationships formed in the time span). Clusters, or colours, are considered to be communities of similarly grouped items (van Eck & Waltman, 2011).

Two types of analyses were conducted:

- Analysis 1 - Co-authorship links based on country of affiliation for over 41 countries. This analysis was to visualize how networks of authors across diverse geographies have collaborated on publications associated with engineering education and CDIO, and how these networks have evolved over time.
- Analysis 2 - Co-authorship links based on author linkages for over 3000 authors with co-authorship. This analysis shows which authors wrote papers together. The full counting method was used, where each co-author was counted as a full author for the paper. Another visualization technique is available using fractional or normalized counting, which is utilized to provide a more consistent basis on which to compare different fields of study (Waltman & van Eck, 2013). Since we were less concerned about scaled impact, and more about community trends within only one field, full counting was used. Only Scopus results were used for this particular analysis as the formatting between the two databases was not consistent and Scopus provided a richer dataset than Web of Science.

Word Frequency Analysis

Word frequency and co-occurrence analysis based on title and abstract data from Scopus (1293 unique entries) was completed to track trends in idea clusters (themes) in that database over time. A corpus of title and abstract data was created for the year divisions: 2002-2007; 2002-2010; 2002-2012; 2002-2014; 2002-2016; 2002-2018 (cumulative data) as well as 2002-
A visualization of frequency of publication types by year and database can be found in Figure 1 and Figure 2.
An aggregation of frequency of publication by year across the two databases is shown in Figure 3.

A log-linear and log-log plot of the number of publications vs year number (from 1-17) was completed and are shown in Figure 4 and Figure 5.
The log-log graph reveals that publication frequency underwent a power-law growth between 2000-2011, indicated by the approximately linear portion of the log-log graph. Since 2011, publication growth has reached a relatively steady state (flat region), indicating that CDIO has permeated into engineering education literature to its maximum extent in the absence of any major changes in its approach. The apparent outlier on the graph coincides with 2018 (year 17 on this graph) for which there are currently only a few publications.

**Analysis 1 – Country Co-authorship Analysis**

The initial analysis was performed on a year-by-year incremental basis, however, the most visually dynamic ranges were found to correspond to years that included changes in publication rate. With this, 2002 to 2007 formed the initial duration, and then this duration was extended to include years 2010, 2012, 2014, 2016, and 2018. The database file was imported into VOSViewer and co-authorship analysis by country was chosen. This analysis treats countries in the author affiliation field as co-occurring. A threshold value of a minimum of 3 documents to be published by a particular country was used for this analysis. The number of countries with connections that met this criteria for each age range are shown below.

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Number of Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002-2007</td>
<td>3 countries</td>
</tr>
</tbody>
</table>

2002-2010 | 10 countries were found with co-authorship links, 2 clusters were formed

2002-2012 | 14 countries were found with co-authorship links, 4 clusters were formed

2002-2014 | 28 countries were found with co-authorship links, 4 clusters were formed
The general trend observed in this analysis is that the community of researchers publishing on engineering education and CDIO-related topics has expanded geographically over time. From 2002-2007, the geographic network as defined in this analysis consisted of three countries: Sweden, United States, and the United Kingdom. Since then an additional 38 countries have engaged in collaborations in this area. In 2012 China emerged as a significant contributor in this ecosystem. The relatively flat graphic for 2002-2012 implies that China was less integrated.
with the other major contributors in this region. In 2014 China became much more integrated into the community with links to many more countries. The 2002-2018 visual demonstrates a community consisting of China, Spain, Finland, France, Mexico and Serbia, and another community consisting of United States, Russian Federation, Sweden, Brazil, Portugal and others. These findings support those of a recent report by Graham (2018) in which a global shift of engineering education leadership has been observed from west to Asia and South America. The 2017 14th CDIO Conference welcomed 35 countries, which is approximately equivalent to the number of countries currently represented in this database. There is an opportunity for future research to compare whether there are any differences between these networks and the CDIO library data, and how this subset of literature differs from the broader community established in engineering educational literature in general. Any differences in community trends could represent opportunities or areas for collaboration or engagement.

**Analysis 2 - Co-Authorship analysis by author**

The results from Scopus were imported into VOSViewer, and analyzed for co-authorship by name, using full-counting. Since Scopus and Web of Science export their author names in different formats, only Scopus was used as it contained more items. This visualization was challenging to create as it required significant data processing. Initial analysis yielded unexpected results as many influential authors were not visible in the output. Upon further inspection it was found that there were many authors with similar names but very different affiliations. For example, the author name Wang Y was found to be affiliated with 10 different institutions. A manual cleaning of the data was then conducted with iterations in VOSViewer until names with multiple affiliations no longer impacted the analysis. Names were coded with numbers for each new affiliation, as it was assumed that each author name with a different affiliation was a new individual. We recognize that there may have been some authors removed from the analysis who may have had more than 1 affiliation, however it was not possible to identify which combination that would have been in cases where names were affiliated with three or more institutions. We therefore opted to take a consistent approach to reduce bias in the analysis. The minimum number of documents for an author to be included in the analysis was 1 publication, however authors must have engaged in a collaboration with another author to appear in these visualizations.
2002-2007 | 135 authors, 35 relevant items

2002-2010 | 544 authors, 500 most relevant chosen, 42 relevant items
The findings from this analysis indicate that five major communities of researchers have been established over the past seventeen years. The relatively flat, dispersed visual indicates that these communities publish relatively independently from one another. Future iterations of this analysis could further probe what the demographics of each of these networks are and whether there are thematic differences in interests within each of these groups.

**Analysis 3 – Word Frequency Analysis of Title and Abstract Data**

An analysis of single word frequency of each of the six time periods, both cumulative and non-cumulative, revealed that the top four most frequently used words in all years were engineering, design, students, and education. This finding is not surprising and an indication that the core of the focus of the literature has not changed much over the last sixteen years, and that students, education, and design remain important themes. These findings become more revealing when two- and three-word phrases are also analyzed.

A rank-order and a log-log plot of the approximately 300 top-ranking two-word phrases is shown in Figure 6. Word frequency for 2-word phrases between 2002-2018 revealed that by far the most mentioned phrase was “engineering education”, and the second most mentioned phrase was “based-learning”.

*Proceedings of the 14th International CDIO Conference, Kanazawa Institute of Technology, Kanazawa, Japan, June 28 – July 2, 2018.*
Figure 6. Rank order and log-log rank vs frequency plot of two-word phrases from 2002-2018.

The rank vs frequency (non-log) plot appears to follow Zipf’s law (Zipf, 1949), which states that any text corpus will approximately follow a power law. This is further confirmed by taking the log-log of the data, which shows a roughly linear trend which is to be expected for a power-law relationship.

To create a word cloud, the term “engineering education” had to be removed. It represented 8% of all words in the corpus and could not be drawn since its weight was comparatively too high to the other words. Visualizing the log-value did not provide enough resolution between the frequencies and the word cloud lost dimensionality. A word cloud of the remaining top two-word phrases was then created and is shown in Figure 7.

Figure 7. Word cloud of 2-word phrases from titles and abstracts 2002-2018.
The word cloud is a useful tool to quickly gain insight into important phrases in the corpus. The reader can easily see that “Project Based”, “Design”, “Learning Outcomes”, and “Engineering Students” are the most popular terms. Each reader may notice different relationships and trends based on their own personal experience and biases on what they may find to be important. It is important to note that the word cloud was presented not to stand alone and be interpreted individually but rather as one data point across the entire analysis to be used to promote discussion. In the future, it could be presented next to others from the CDIO.org library or compared to the broader corpus of engineering education literature. In this paper it will be utilized in tandem with a more detailed frequency analysis of themes.

To get a better understanding of whether publishing trends on particular engineering disciplines followed the same trends as practice, rank of frequencies of two-word engineering discipline phrases in the corpus was conducted. The ranking of mention of engineering disciplines in the literature was then compared to findings from available survey data of CDIO practitioners presented in Malmqvist, Hugo, & Kjellberg (2015). The 2015 survey asked CDIO practitioners which disciplines they practiced CDIO in, and a ranking of the disciplines applying CDIO were presented. Our findings from the word frequency analysis found that mechanical engineering ranked first in number of mentions, followed by electrical engineering, then aerospace, computer science, civil engineering, chemical engineering and industrial engineering respectively. These rankings roughly accorded with rankings from the 2015 practice survey (Malmqvist et al., 2015). There were two exceptions - aerospace engineering, which represented a disproportionately high frequency of mentions, and industrial engineering which represented a disproportionately low number of mentions in comparison to the survey data. The remainder of the disciplines followed the same rank-order, indicating literature themes roughly follow the trends observed from practice (or at the very least it indicated that those publishing on their practice parallel the sample who responded to the 2015 survey). Future analysis could investigate why there are comparatively so few papers published on applications of CDIO to industrial engineering, and may represent an intervention area for the CDIO community if there are barriers that prevent publishing in this area, for example. Additionally, biomedical engineering and biological engineering were present and this discipline ranked last behind the previously mentioned disciplines, however frequency of mention have reduced to approximately 40% of the 2002-2007 value, indicating less focus in the literature over time. Emerging disciplines, such as nanotechnology engineering and robotics engineering were barely present in the literature, perhaps reflecting barriers to publishing or practicing CDIO in these areas, however this requires further investigation. A more comprehensive analysis for relevant word combinations in these areas should be conducted in the future, as emerging fields their themes and vocabulary are less homogeneous and it may be more difficult to quantify their presence by a word frequency analysis than more mature disciplines.

The remainder of the analysis focused on better understanding which CDIO standards were being represented in this corpus of literature. For this analysis the “essential” (Edström & Kolmos, 2014) CDIO standards were examined: context, learning outcomes, integrated curriculum, design-implement experiences, integrated learning experiences, enhancement of faculty competence (also: faculty training, faculty development, tutor training), learning assessment. In the corpus, design-implement was merged with design-implementation as these phrases both had similar number of mentions. While “context” is on the list of “essential” CDIO standards, at this level of analysis it was difficult to tell whether its use was related directly with reference to CDIO standards, or whether it was used in a sentence with some other meaning. The word “context” was therefore excluded, as a more detailed linguistic
analysis would be required to better parse out when it has been used in a way that is relevant and was left for future work.

In 2002-2007 “design implement operate” was a top-ranked phrase, and remained a high-ranking phrase among the literature, indicating that authors find it a useful approach for their teaching and learning tasks. The standards analyzed and their frequencies across all years (total) were compiled in a treechart and visualized in Figure 8.

![Treechart of frequency of essential CDIO standards phrases appearing in literature (2002-2018)](image)

**Figure 8.** Frequency of essential CDIO standards phrases appearing in literature (2002-2018).

While this treemap demonstrates critical trends with respect to the CDIO standards, it doesn’t show the full picture. The theme “project based learning”, which is not a CDIO standard, went from being 21st most important from 2002-2007 to consistently ranking among the top two most mentioned phrases in the literature between 2008 and 2018 (shared with design implement). It was included in the treechart in Figure 9.
A more detailed picture can be drawn by examining the time-based trends associated with these themes. A stacked, normalized bar chart representing year range and relative frequency of mention of each theme is shown in Figure 10.

Figure 9. Treechart of total relative frequency of CDIO essential standards mentions in the literature 2002-2018, including Project Based Learning.

Figure 10. Relative frequencies of CDIO essential standard phrase mentions in literature by year range.
The analysis revealed that the bulk of the research in external literature is centered around techniques for teaching – papers discussed student learning and teaching practice reflecting some "essential" CDIO standards, but not all of them. "Learning outcomes" remained important across all the years, with peak mentions in 2013-2014. The phrases “integrated curriculum” and “integrated learning” have had very few mentions in the literature over the past 17 years, with the frequency of mentions remaining relatively stable. “Design-implement”, “design-implementation” and “design implement operate” have continued to be integral themes, though peak mentions occurred in 2011-2012 and have declined by about 50% since. Enhancement of faculty competence, faculty training, faculty development, tutor training have had very little reference in the literature. Learning assessment also has had very few relative mentions in the past, which is a gap given the equal emphasis one would expect it to have with learning outcomes in a constructively aligned curriculum (Anderson, Krathwohl, & Bloom, 2001). These themes (faculty competence and assessment) could represent ripe areas for future work.

A main benefit of the CDIO approach is that it provides a framework for educational reform that is holistic (Crawley et al., 2014). A key question that therefore arises from these findings is: why does the literature not appear to reflect this? A key component to CDIO is the assumption of an approach to learning that is “solution-independent” (Edström & Kolmos, 2014). These findings, however indicate that practitioners have primarily focused on application of one particular solution. Perhaps this is an indication that project-based learning is the solution that best solves the question of how to teach CDIO standards, but this debate is far from settled (Beddoes, Jesiek, & Borrego, 2010; Kirschner & Clark, 2006). If project-based learning continues to represent an increasing proportion of the literature related to CDIO, CDIO faces the risk of diluting its unique value proposition. An over-emphasis on project-based learning could lead CDIO to become synonymous with the community of practice of project-based learning; while the two have many synergies, there are critical differences (Edström & Kolmos, 2014) which should be maintained and supported.

In 2002-2007 the three-word phrase “reform engineering education” was ranked first in a tie with the phrase “design implement operate”. The use of the phrase has dropped off quite remarkably, with no mentions in 2017-2018 literature whatsoever. There were zero responses for the phrases “faculty training” or “tutor training” across all years. The phrase “program evaluation” dwindled from a top rank (15) to very little importance in the literature whatsoever (rank 1679 in 2017-2018).

These findings appear to indicate that while the CDIO initiative has impacted the external literature landscape particularly in the realm of increasing publications on design-implement and project-based learning activities, discussions on systemic and institutional change have not gained the same relative momentum. PJBL and active learning design-build activities are excellent engagement mechanisms, but without sustained, deliberate, holistic stakeholder engagement, these activities may not live on in the curriculum once the instructor or faculty offering them in their course has moved on. Support for further discussion on faculty development and learning assessment is a critically important factor in the continued sustainability of the CDIO initiative, therefore a future emphasis in these areas should be maintained.

One could argue that mentions of engineering education reform have decreased, while annual publications have increased (Figure 3), particularly on innovative pedagogical practice (Figure 7-Figure 10), because reform has truly occurred. In this highly unlikely case, however, then what place does an educational reform framework play in this ecosystem as we move forward? Another interesting trend, however, is that annual publications initially followed a power-law
growth, and now appear to have reached a steady-state. As CDIO matures as an organization it will be critical to better understand how it can remain relevant within an ever-evolving community of practice. It is critically important to remember that “engineering education reform” is not an end state; how we understand and implement engineering education reform is rather constantly evolving based on the expansion of knowledge and experience. As practitioners seek to embed new content, accommodate increasingly lifelong learners, and above all, incorporate the benefits of blended learning through off- and on-campus education, engineering education reform initiatives like CDIO must and will continue to adapt.

CONCLUSION

The findings in this paper are meant to support discussion and decision-making for the initiative by visualizing the evolution of CDIO influence in the field of engineering education since 2000. 1426 distinct records from Scopus and Web of Science relating to the search terms “engineering education” and “CDIO” were used for the analysis. Co-authorship analysis was completed for country and author name to visualize publication networks in this ecosystem. A word co-occurrence analysis was conducted to visualize how ideas in this realm are related to one another. The geographic network of research collaborations has expanded over time from three to 38 countries. From 2012 to 2014 China emerged as a collaborator, becoming more integrated in the network over time. Co-authorship analysis by name revealed a set of core collaborators that existed throughout the years, however these communities became more and more isolated over time. Word frequency analysis found that external literature has placed a great deal of emphasis on learning approaches and interventions, however considerably less comparative discussion has occurred on other CDIO standards such as learning assessment and faculty competence. Findings from this analysis could indicate areas that may require heavier emphasis by CDIO leaders, for example as themes around meetings or conferences, or through the support of research. Researchers and practitioners can also use these findings to guide their topics of focus for future inquiry. By supporting further dialogue in the currently under-served standards as indicated by this analysis, CDIO is more likely to maintain its distinct niche, and therefore relevance, within the engineering education ecosystem. Like so many disciplines in the 21st century, the field of engineering education reform is on the brink of major change. The shift of the centre of gravity of engineering education leadership from the western world to Asia and South America (Graham, 2018) will mean new challenges and opportunities for the CDIO initiative. The changing position of university (and its education) in society will also be an interesting consideration as CDIO continues to expand its global influence, and relevance. At the very core of the CDIO initiative’s holistic approach lies the tools and the abilities to adapt to the ever-evolving needs of the ecosystem’s many stakeholders. We expect CDIO’s commitment to continuous improvement and the community’s openness to self-reflection will allow it to adapt to the ever-changing changing world.

REFERENCES


BIOGRAPHICAL INFORMATION

Alexandra Meikleham is an MSc student in Mechanical engineering at the University of Calgary in Calgary, Canada. Her current research focuses on human-centered design and on engineering curriculum development.

Ron Hugo is Professor of Mechanical and Manufacturing Engineering and the Chair in Engineering Education Innovation in the Schulich School of Engineering. His research interests are in the areas of experimental fluid dynamics, energy systems, and engineering education.

Aldert Kamp is the Director of Education for TU Delft Faculty of Aerospace Engineering and Co-director of the CDIO-initiative. He is deeply involved in the rethinking of higher engineering education with a horizon of 2030. He is the author of the report “Engineering Education in a Rapidly Changing World – Rethinking the Vision for Higher Engineering Education” and fervent blogger about “Adapting Engineering Education to Change”.

Johan Malmqvist is a Professor in Product Development and Dean of Education at Chalmers University of Technology, Gothenburg, Sweden. His current research focuses on information management in the product development process (PLM) and on curriculum development methodology.

Corresponding author
Alexandra Meikleham
University of Calgary
2500 University Dr. NW
Calgary, Alberta T2N1N4
1-514-567-8488
alexandra.meikleh1@ucalgary.ca

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License.