A computational approach to evaluating curricular alignment to the united nations sustainable development goals

Philippe Lemarchand[1,2,3], Mick McKeever[4],
Cormac MacMahon[5]* and Philip Owende[6]

[1]School of Electrical and Electronic Engineering, Technological University Dublin, Dublin, Ireland, [2]Dublin Energy Lab, Technological University Dublin, Dublin, Ireland, [3]MaREI SFI Centre for Energy, Climate and Marine, Dublin, Ireland, [4]School of Informatics and Engineering, Technological University Dublin, Dublin, Ireland, [5]Sustainability Office, Technological University Dublin, Dublin, Ireland, [6]Academic Affairs, Technological University Dublin, Dublin, Ireland

The United Nations (UN) considers universities to be key actors in the pursuit of the Sustainable Development Goals (SDGs). Yet, efforts to evaluate the embeddedness of the SDGs in university curricula tend to rely on manual analyses of curriculum documents for keywords contained in sustainability lexica, with little consideration for the diverse contexts of such keywords. The efficacy of these efforts, relying on expert co-elicitation in both subject-matter contexts and sustainability, suffers from drawbacks associated with keyword searches, such as limited coverage of key concepts, difficulty in extracting intended meaning and potential for greenwashing through "keyword stuffing." This paper presents a computational technique, derived from natural language processing (NLP), which develops a sustainability lexicon of root keywords (RKs) of relative importance by adapting the Term Frequency−Inverse Document Frequency (TF-IDF) method to a corpus of sustainability documents. Identifying these RKs in module/course descriptors offers a basis for evaluating the embeddedness of sustainability in 5,773 modules in a university's curricula using classification criteria provided by the Association for the Enhancement of Sustainability in Higher Education's (AASHE). Applying this technique, our analysis of these descriptors found 286 modules (5%) to be "sustainability focused" and a further 769 modules (13%) to be "sustainability inclusive," which appear to address SDGs 1, 17, 3, 7, and 15. Whilst this technique does not exploit machine learning methods applied to large amounts of trained data, it is, nevertheless, systemic and evolutive. It, therefore, offers an appropriate trade-off, which faculty with limited analytics skills can apply. By supplementing existing approaches to evaluating sustainability in the curriculum, the developed technique offers a contribution to benchmarking curricular alignment to the SDGs, facilitating faculty to pursue meaningful curricular enhancement, whilst complying with sustainability reporting requirements. The technique is useful for first-pass analyses of any university curriculum portfolio. Further testing and validation offer an avenue for future design-science research.

Introduction

The Anthropocene is accepted as an era in which population growth, affluence and technological advances have yielded unprecedented human capacity to alter Earth's ecosystems (Lewis and Maslin, 2015). As socio-economic activity and earth system indicators inflected, high profile publications have warned of their negative environmental implications (Meadows et al., 1972; Brundtland, 1987; Steffen et al., 2011). Yet, consensus on how to address one such implication, i.e., climate change, only emerged in 2015 when the Paris Accord became the first binding initiative to curtail atmospheric temperature rises (Schurer et al., 2018). A broader agenda, enshrined in the UN SDGs, articulates a complex intertwine between the environment, society and economy (Rockström et al., 2009; Raworth, 2017). Premised on equity, security, prudence, connectivity, and comprehensiveness, it charters an unfolding transformation for societies (Sachs et al., 2019), in which primacy is given to socioecological systems. Hence, sustainable development has emerged as a thoughtful consciousness to our very existence. Whilst various definitions of sustainable development exist, its depiction as development that "meets present needs without compromising the ability of future generations to meet their own needs" is widely accepted, implicit in which intergenerational inequality poses an existential challenge (Brundtland, 1987), 292).

The SDGs envision a future of social justice and economic prosperity within environmental limits, with education as key agent for transformation. Access to quality education is identified as a standalone goal (SDG 4) to facilitate achievement of other SDGs (Vladimirovaa and Le Blanc, 2015). SDG 4 is premised on educational purposes being fulfilled. In the emerging socio-ecological paradigm, these purposes broaden from an economic focus to a wider lens encompassing global citizenship and environmental stewardship (Kioupi and Voulvoulis, 2019). SDG 4.7 mandates higher education institutions to prepare graduates for this paradigm, and is measured by the extent to which Education for Sustainable Development (ESD) is mainstreamed. Consequently, it can be argued that integration of sustainability in curricula is a key step in implementing effective ESD. There have been many studies to evaluate such integration (Shriberg, 2002; Lozano and Peattie, 2009) and, since 2015, there is growing momentum for using SDGs as indicators of progress (Albareda-Tiana et al., 2018; Brugmann et al., 2019; Chang and Lien, 2020). Despite this momentum, SDG inclusion in university curriculum is in its infancy and some criticize the SDGs for being far too broad and even too aspirational for learning purposes (Janoušková et al., 2018).

In recognizing the benefits to localizing SDGs in learning outcomes, Technological University Dublin (TU Dublin) aims to ensure all its programmes include sustainability as a learning outcome (TU Dublin, 2019). Implicit is an expectation of material alignment between constituent academic courses and SDGs. At TU Dublin, a course is typically referred to as a "module" and a programme of study typically comprises multiple modules. In acknowledging that some modules may be focused on sustainability, whereas others may contribute indirectly to SDGs, the Association for the Enhancement of Sustainability in Higher Education's Sustainability Tracking and Assessment Rating System (AASHE-STARS) includes an interpretive framework for categorizing modules (AASHE, 2015). Yet, interpretive validity relies on expertise and norming. Whilst expert interpretation can be validated by statistical means, it would be problematic for a large curricular portfolio underpinned by over 5,000 modules. One approach to addressing this issue is to curate a lexicon of keywords that guides interpretation (Bolden and Moscarola, 2000). Keywords are used with varying degrees of success to detect sustainability in university documentation (Armitage et al., 2020; Pinto, 2021). Yet, as with any keyword search, there can be gaps between detecting keyword prevalence and reliably extracting the context in use. Applying methods used in NLP offers a way to bridge this gap by identifying keywords and phrases and relating them to overlapping topics. NLP has been used in a broad set of sustainability contexts, such as for evaluating alignment of corporate sustainability reports to the SDGs, for analyzing SDG-related policy discourse and for informing environmental, social and governance (ESG) ratings (Chen et al., 2021; Smith et al., 2021).

This paper aims to explore the application of a statistical model of root keywords (RKs) to evaluate the embeddedness of the SDGs in module/course descriptors and to categorize modules using objective criteria on the basis that compliance is determined or strategy for enhancement is enabled. The underpinning study interrogated an inventory of 5,773 modules on offer at TU Dublin in 2019, across a spectrum of levels on Qualifications Quality Ireland's (QQI) National Framework of Qualifications (QQI, 2021). The study develops a systematic method for identifying the prevalence of SDGs in module/course descriptors, then correlates it with module categorization based on AASHE-STARS criteria. In support of this research objective, the study sets out to answer two underlying research questions:

RQ1: How can academic programmes identify the most relevant keywords to be used in accurately reflecting content in module/course descriptors to indicate the presence of sustainability components relating to SDGs?

RQ2: How can the associations to SDGs in module/course descriptors be used to evaluate the embeddedness of sustainability in modules and programmes and, by extension, in school, faculty and university curricula?

The paper begins with an outline of the ESD agenda and TU Dublin's approach to achieve compliance. It explores the sustainability reporting frameworks for higher education,

including AASHE-STARS. Based on abductive reasoning (Walton, 2014), it then examines various approaches to evaluating sustainability in taught modules before stepping through the technique for compiling an effective contextual sustainability lexicon for TU Dublin, based on a corpus of databases on sustainability related topics. The paper then outlines how the lexicon is applied to form a database of RKs with numerical representations of sustainability importance (SI). The application of SI scoring to categorize each module based on an AASHE-STARS criteria and to attribute the SDGs that they address is then detailed. Areas for further research to undertake a manual verification of the technique and to explore the technique's potential application as a supplementary tool for curricular transformation and engagement with faculty are proposed.

Materials and equipment

The evolving sustainable development agenda

As humanity becomes more aware of Earth's limitations, it faces greater pressure to pursue a delicate balancing act, involving complex trade-offs and synergies between its social, ecological, and economic systems (Raworth, 2012). Yet, as a concept, sustainability is difficult to articulate, nebulous and self-contradictory (Kuhlman and Farrington, 2010). A pragmatic approach to its operationalization relates to its measurement through indicators and targets. This was the basis for the UN publishing the Millennium Development Goals (MDGs) in 2000, comprising 8 goals and 21 targets. In 2015, UN member states adopted the more ambitious SDGs. Whereas, the MDGs applied to the poorest countries, the SDGs apply to all countries across a spectrum of development challenges, each with its different priorities, contexts and capacities to capture the data needed to track progress. The SDGs represent a time-bound call to eliminate poverty, to protect the Earth and to provide security, setting critical thresholds for natural resources usage, whilst recognizing that poverty elimination is integral to economic development (Lafortune et al., 2018).

The SDGs are structured into 17 goals, 169 targets and 244 indicators in an integrative framework, reflecting interwoven challenges along the three dimensions: economic, social, and environmental (Desa, 2016). Whilst the SDGs are not legally binding, progress toward achieving them is evaluated using composite indices. National governments are expected to establish frameworks for implementation and report progress periodically (Government of Ireland, 2018). The UN Sustainable Development Solutions Network (UNSDSN),

the European Commission (EC), and the Bertelsmann-Stiftung (BS) have since published reports monitoring SDG implementation.

Implementation of the SDGs is intended to be a whole-of-society enterprise based on a systems approach to problem-solving (Yang and Cormican, 2021). Countries present their data at the UN's High-level Political Forum (HLPF) through Voluntary National Reviews (VNRs), highlighting actions and challenges. Two of the more complex challenges relate to data and methodological considerations. To address these challenges, the UN has published data resources, handbooks, and metadata for each indicator. Some indicators lack standardized methodologies, whilst some countries lack capacity to produce official data. There are also sustainability issues not addressed by the SDG framework, e.g., how to measure spill-over effects from one country's SDG action into another.

Education for sustainable development

Whilst SDG 4 recognizes education's transformative potential for achieving the SDGs, the idea that ESD could be an engine for transformation by disrupting deeply-rooted epistemology, pre-dates the SDGs. For example, the UN conference on the Human Environment in 1972 articulated the first steps toward ESD in higher education (Calder and Clugston, 2003). The subsequent Talloires Declaration raised the profile of sustainability in universities. The UN Decade of Education for Sustainable Development (DESD) invited universities to explore issues such as poverty, equality, climate change, and circularity by reorienting learning and teaching toward sustainable development principles.

Whilst ESD has been criticized for lacking clear aims (Leal Filho et al., 2015), it is now accepted that it seeks to develop the competencies needed to contribute to a sustainable future. Although the initial focus of ESD was on a separate provision, it has since shifted to integrating sustainability in the curriculum, so that students can understand how their respective disciplines contribute. Therefore, understanding the embeddedness of sustainability in the curriculum can offer guidance on how to improve contextual approaches to ESD. Whilst studies on implementing sustainability in the curriculum are integral to ESD scholarship, there are few such studies mapping this implementation (Weiss and Barth, 2019). Now that ESD has been made an explicit target (SDG 4.7), its priority in higher education has risen (Longhurst et al., 2021). On one hand SDG 4 is a goal in itself and, on the other, it is also a vehicle for the societal transformation needed to achieve other SDGs. This dual-purpose necessitates a new education model that integrates the SDGs in learning (Rieckmann, 2017).

Implementing education for sustainable development in ireland

Ireland's National Strategy on Education for Sustainable Development, aims to equip "learners with the relevant knowledge, key dispositions and skills and the values that will motivate and empower them throughout their lives … for a more sustainable future" (Government of Ireland, 2014), 3). Higher Education and Research is identified as a priority area, which encompasses: cross-disciplinary learning, curriculum with discipline specific ESD content, up-skilling lecturers to deliver ESD curriculum, and a whole institution approach to ESD. The most common response has been to offer both new sustainability focused programmes as well as integrating sustainability into existing modules. The curriculum affects other priority areas within the national strategy, including professional development for teachers to deliver the curriculum [e.g., (Harvey et al., 2021)]. To meet national reporting requirements for SDG 4.7 and to address ESD challenges in higher education, Ireland's Higher Education Authority (HEA) is tasked with reporting the extent of ESD across the Irish higher education sector. Thus, universities in Ireland are increasingly required to report on the number of modules that address sustainable development. Yet, this seemingly innocuous requirement disguises the major reporting complexities for higher education providers.

Sustainability reporting in higher education

Ever more onerous reporting is becoming synonymous with an evolving sustainability agenda. Whilst many sustainability reporting tools exist, they often lack indicators relevant to higher education. Therefore, a raft of tools has emerged specific to sustainability in universities (Holst et al., 2020), their basic premise being to evaluate progress, to set targets and to communicate their sustainability efforts to university stakeholders. Available examples include: Assessing Integration of Sustainability in Higher Education (AISHE) (Lambrechts and Ceulemans, 2013); Campus Sustainability Assessment Framework (CASF) (Cole, 2003); merging Global Reporting Initiative (GRI) indicators with indicators for higher education (Lozano, 2006); Sustainability Tool for Assessing Universities' Curricula Holistically (STAUNCH) (Glover et al., 2011); Learning in Future Environments index (LIFE) (Macgregor et al., 2019); and the uncertainty based DPSEEA-Sustainability index Model (Waheed et al., 2011).

The tools outlined above emphasize sustainability integration across four key functions, namely; education, research, community and operations. Therefore, universities that publish sustainability reports, typically include their sustainability curriculum as a central component (Kosta, 2018). For example, CSAF assesses curricular offerings based on the proportion of modules with some sustainability content and the number of students taking those modules. However, common issues, such as loose definitions and lack of standardization problematizes the benchmarking of sustainability in the curriculum. Hence, sustainability has yet to permeate throughout university curricula and the presence of learning outcomes specific to sustainability is still relatively uncommon (Lozano, 2010).

Evaluating sustainability in the curriculum

Many universities use curriculum management systems to store module/course descriptors, which provide easily accessible data for advancing sustainability in the curriculum. Much of the current thinking in evaluating sustainability in the curriculum focuses on combining text-based methods with interpretative approaches. This combination produces quantitative and qualitative data that can facilitate faculty and student reflection (Tierney et al., 2015). One of the standards to accommodate this approach is the AASHE-STARS self-reporting framework, which, in 2021, facilitated 550 institutions across 14 countries to report on a diversity of sustainability efforts. STARS deploys a credit system, scoring universities over five categories, namely: Academics; Operations; Engagement; Planning & Administration; Innovation and Leadership. It is considered to be amongst the most balanced tools for sustainability tracking and reporting (Kamal and Asmuss, 2013).

Figure 1 shows the AASHE-STARS categories and subcategories. The Academics (AC) category directly relates to SDG 4, comprising two sub-categories—Curriculum and Research. One of the eight components under the Curriculum sub-category, Academic Modules/Courses (AC1), evaluates module content covering sustainability. Under AC1, *Sustainability-focused* modules have a primary and explicit focus on sustainability or on understanding or solving one or more major sustainability challenges. Modules can provide [i] introductory knowledge of sustainability, or [ii] the application of sustainability, or [iii] the skills and knowledge directly connected to understanding or solving one or more major sustainability challenges, e.g., Climate Science, Renewable Energy or Green Chemistry. *Sustainability-inclusive* modules focus on topics other than sustainability but explicitly include a sustainability component. For example, a foundational module in sociology or chemistry might be useful to practitioners of sustainability but it would not be considered inclusive of sustainability unless the concept of sustainability or a sustainability challenge was specifically integrated. To calculate the commensurate credit, universities need to report a range of descriptive statistics derived from the absolute number and percentage of sustainability-focused and sustainability-inclusive

modules for undergraduate and postgraduate programmes, along with a description of the methodology used to determine such descriptive statistics.

The core challenge for universities offering large number of modules lies in their ability to categorize their modules appropriately, which can be a labor-intensive operation and be subject to interpretive bias. Conceivably, the SDGs could be used to ease this process in an objective manner if modules were to be systematically identified with and/or assigned for handling of specific SDGs. Curricular surveys that map modules to the SDGs, have already been attempted by some universities (Brugmann et al., 2019). The UNSDSN offers guidelines on how to undertake mappings (UNSDSN Asia Pacific, 2017). A synthesis of relevant literature identifies core competencies for sustainability (Wiek et al., 2011). Nottingham Trent University and National University of Kaohsiung have implemented their mappings on their own curriculum management systems so that SDGs to which each of the modules relate are visible to both faculty and students (Chang and Lien, 2020).

Text based approaches

Keywords searches of large collections of texts have been used to identify sentiment, themes and topics in documents (Mishra et al., 2021). These corpus-based approaches have, in recent times, been contextualized to SDG related research (Meschede, 2020). Since 2018, Elsevier has generated structure query language (SQL) searches to help universities track their research contributions to the SDGs. They have been subsequently enhanced by machine learning in 2020 (Rivest et al., 2021). The same approach could apply to identifying modules aligned to the SDGs, the main contextual difference being that module/course descriptors are typically brief summaries of the expected learning outcomes and the associated learning activities, including assessment strategies, hence are typically much shorter than research papers. Given the drive toward standardization of curriculum formats in Europe post Bologna, there is significant uniformity in module/course descriptor formats. A typical module/course descriptor in the TU Dublin's Programme Module Catalog (PMC) is provided in the Supplementary material file to this paper.

The PMC and the module/course descriptor, in particular, are designed to accurately reflect the content coverage of each module, and often inform prospective students about the module, hence, may be the basis of decision to enroll. Students, for example, could use the module title to predict the topics of informational text but, in some cases, the titles can be idiosyncratic and contextualized to the programme of study and, therefore, do not have much predictive value. Supplementary Table S1 in the Supplementary material file illustrates a sample module descriptor. The "Overview" section of a module/course descriptor typically offers a succinct outline

of the purpose of the module and is usually written from the expected outcome perspective of the instructor. Learning outcomes are statements of what students are expected to be able to do or know on successful completion, demonstrating knowledge, understanding or competence. A list of learning outcomes is typically preceded by some variant of the phrase:
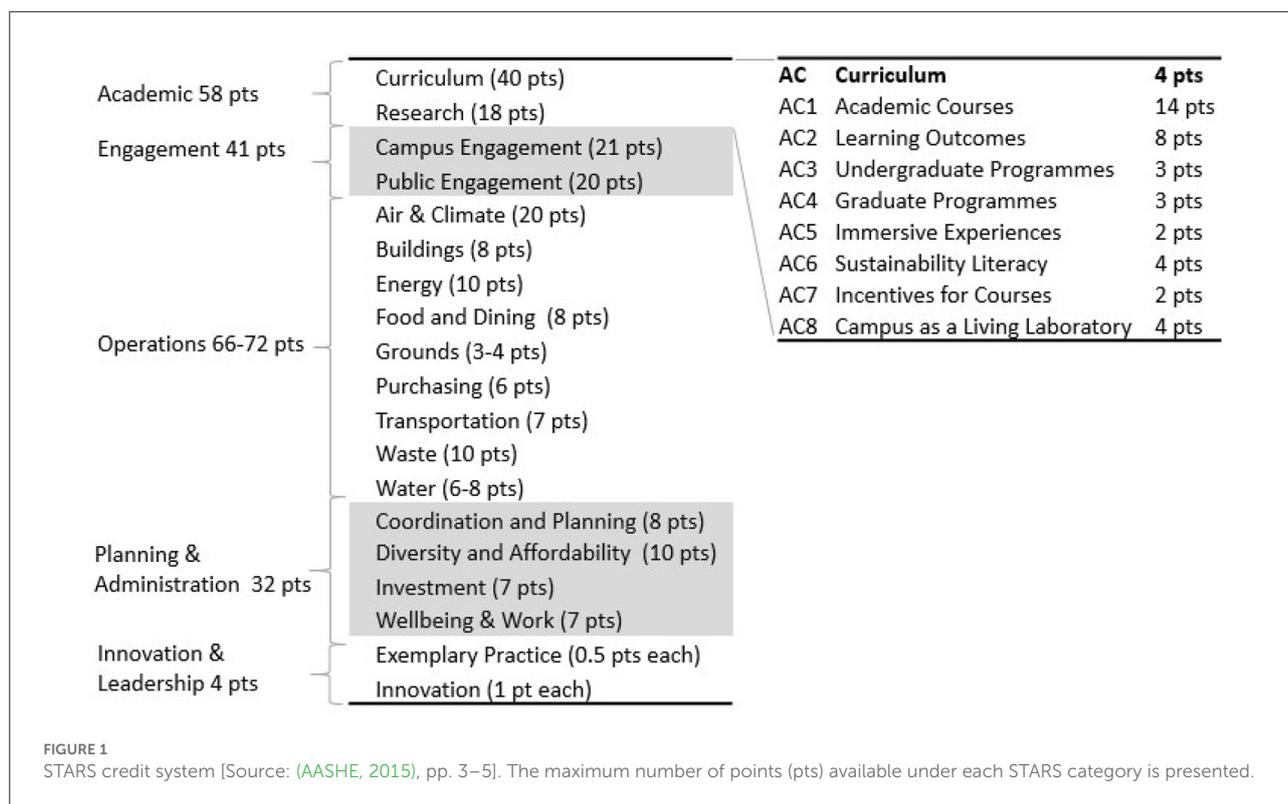
> "On successful completion of this module, students will be able to…".

Learning outcomes commence with an action verb to specify behavior. Different verbs are used to demonstrate different levels of learning (Stanny, 2016). The content section in a module/course descriptor outlines the topics to be covered, and the section is typically written as an abbreviated list of headings and keywords. A comprehensive module/course descriptor will show constructive alignment between learning outcomes, learning activities and assessments that validate learning.

The final part of the module/course descriptor outlines the learning resources, such as core textbooks, websites and digital media. The important point is that module/course descriptors are, themselves, abbreviated documents, which, if analyzed for keywords, may not yield much insights into coverage of the SDGs. A deeper dive into modules with faculty members and students, using qualitative approaches, would be a key step in the discovery process. Notwithstanding, to identify the degree of sustainability without mobilizing large-scale human resources, a systematic and automated approach to analyzing modules descriptors for explicit reference or linking to the SDGs is required. A single word can have multiple meanings, so extracting the context is at the heart of the challenge. A search for keywords, such as "sustainable," may well be insufficient to detect the explicit reference to sustainable development.

Benchmarking with natural language processing

One approach to identifying topics in texts is by using NLP techniques, such as Latent Dirichlet Allocation (LDA) (Sutherland et al., 2020), which annotates relevant entities and relations identified in the text, or by applying machine learning algorithms, such as word embedding (Word2Vec), which captures the context of a word in a document based on semantic and syntactic similarity with other words (Sato et al., 2020). Such methods require a large corpus of relevant documents and computer science expertise. Open domain NLP demonstrations are available online such as Watson, DBpedia Spotlight and Amazon Comprehend. For example, entering the *Environmental Services and Materials* module/course overview (see Annexe 1 in Supplementary material), which has a distinct focus on sustainable design, into the Watson demonstration

FIGURE 1
STARS credit system [Source: (AASHE, 2015), pp. 3–5]. The maximum number of points (pts) available under each STARS category is presented.
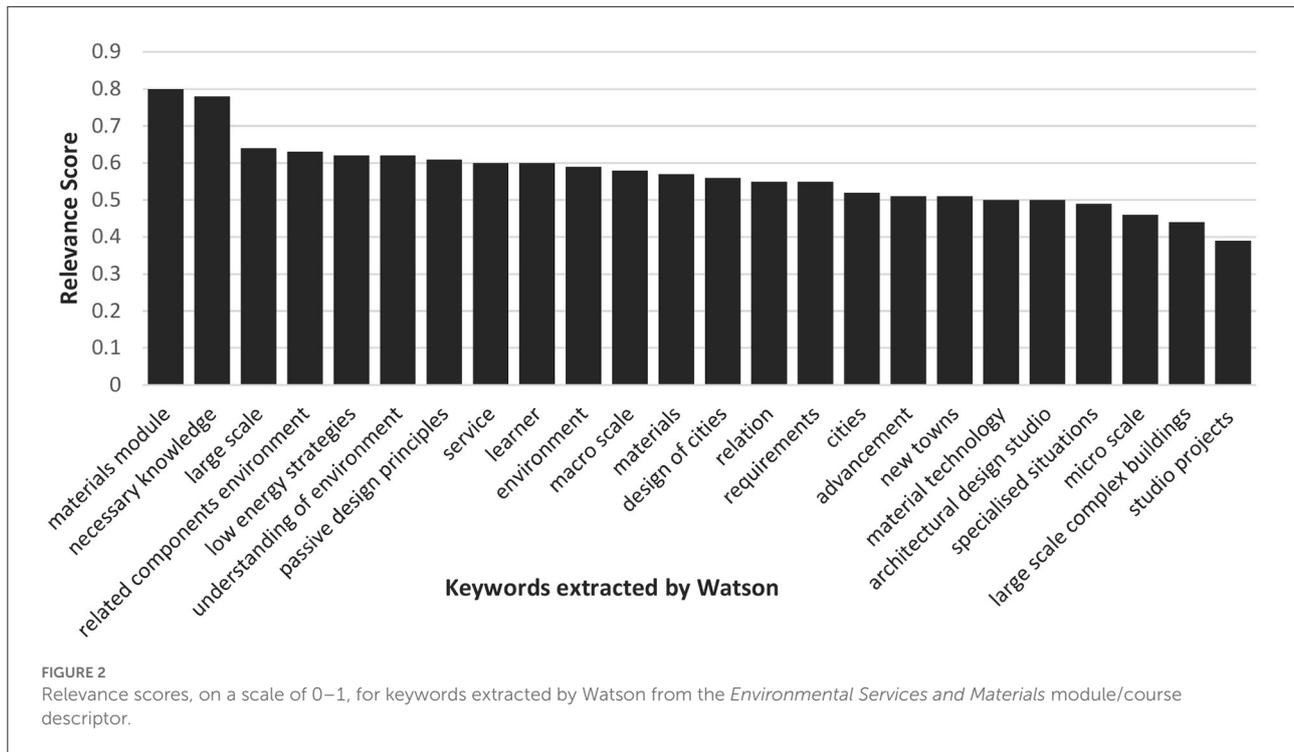
tool, yields keywords ranked by relevance from 0 (for non-relevant keywords) to 1 (for keywords of top relevance). The module/course overview comprises 107 words. Figure 2 shows that Watson returns 25 keywords, representing a large portion of the text (>23%), with relevance scores between 39 and 81%. Many of the returned keywords are composites, most of which, using synonyms or the same individual keywords (such as "material," "environment," and "cities"), can be grouped together as illustrated in the left column of Table 1. Watson's output is then compared with a simpler counting of keywords, after discarding stop words, such as articles, pronouns, conjunctions and prepositions (e.g., to, at, for, of, a, etc.), synonyms and words with the same derivation (e.g., "service" and "servicing") are regrouped and the number of occurrences is aggregated. The middle and right columns in Table 1 show groups of individual keywords with more than one occurrence, organized according to the decreasing number of occurrences.

The keyword hierarchy based on the counting method shows a close match to keyword relevance generated by Watson. There is a small difference in ranking for the word "service," which ranks in 7th position among the grouped keywords returned by Watson but in 3rd/4th position in the word counting method. As the title of the module is *Environment Services and Materials*, one might conclude that the word counting method yields a better relevance of this word. Moreover, the word "introduce," whilst taking a significant position in the counting method, does

not appear relevant for Watson but illustrates that the module includes the important foundational knowledge, as discussed by AASHE-STARS. The keyword "low energy strategies" with 63% relevance according to Watson does not appear as significant in the word counting method as "energy" and "strategies" appear only once in the module/course descriptor. In this case, Watson's result is beneficial as such keywords could still highlight relevance to SDGs.

Watson also identified higher level concepts that were not explicitly referred to in the *Environmental Services and Materials* module/course descriptor, and which included the following, in increasing order of importance, given between brackets: sustainability (0.67), real estate (0.71), energy conversion (0.71), architecture (0.74), construction (0.85), house (0.87), knowledge (0.89), architectural engineering (0.93). Although "Sustainability" scores lowest, it was recognized despite the word "sustainable" or a derivate not appearing in the module/course descriptor. The "Sustainability" concept was most likely linked to keywords such as "environment," "holistic low energy," and "passive design." However, the open domain version of Watson did not directly identify any SDGs that the module potentially addresses. Consequently, if a topic or theme, such as an SDG, is narrowly defined, then the open domain version will not find it. The search protocols would require a domain specific dictionary or trained data. An important first step in developing trained data is the

FIGURE 2
Relevance scores, on a scale of 0−1, for keywords extracted by Watson from the *Environmental Services and Materials* module/course descriptor.

recognition of SDG-relevant content among vast quantities of textual data.

Some researchers have developed ontologies from documents connected to the SDGs using manual extraction or semi-automated approaches to identifying key terms linked to a high-level concept, e.g., Buttigieg et al. (2015) and Bautista-Puig and Mauleón (2019). Others have used supervised machine learning so that such trained models can be used to predict associated data (e.g., (Pincet et al., 2019; Sovrano et al., 2020). Often, ontologies are not comprehensive and, therefore, fail to capture the totality of SDG related discourse. Supervised machine learning models often struggle with out-of-corpus cases. Unsupervised machine learning may be promising, but has yet to be applied beyond a small corpora of texts. However, the available labeled data for SDGs often comes from highly specific homogeneous corpora, such as UN reports, running the risk that such models will capture corpus-specific features that might negatively affect performance on out-of-corpus data. As a compromise, a relative new initiative, SDG Pathfinder, seeks to deploy unsupervised machine learning to identify underlying structures in the data, which were then manually mapped to labels (OECD, 2020). This approach yields a meaningful list of topics linked to the SDGs and a model to detect these topics in new texts, hence, to associate new documents with the SDGs. Each approach to developing trained data has inherent weaknesses. One attempt to overcome these weaknesses has been the OSDG project (Pukelis et al., 2020), which has developed a large ontology of circa. 14,000 keywords from prior

research, which were then mapped to topics from Microsoft Academic. However, there is no analytical tool dedicated to systematically categorizing modules/courses according to the SDGs that they are designed to address.

Methods

Analyzing the sustainability of modules

In this study, 5,773 modules in TU Dublin's PMC were interrogated to evaluate their sustainability components, hence, their relevance and potential contribution to a sustainability focused curriculum. The approach considered the need for a replicable method that could be used to monitor future module enhancements toward achieving ESD and its potential applicability to the higher education context. It was also envisaged that, once validated, the method could be extended to any organization aiming to qualify the sustainability aspects of their core business. The key objectives of the method described in this paper were:

(1) To identify the "sustainability-focused" and "sustainability-inclusive" modules aligned to the AASHE-STARS criteria for measuring the sustainability performance of academic modules.
(2) To distinguish the core SDGs addressed by any identified modules with a sustainability component.

TABLE 1 Comparing returned keywords and percentage relevance from IBM Watson NLU demonstration tool (IBM Watson) to individual keywords and synonyms found in the *Environmental Services and Materials* module/course overview, ordered by the number of occurrences.

Grouped keywords returned by Watson with percentage relevance given in brackets	Keywords	Occurrences
Materials module (0.81); m*aterials* (0.55); *material technology* (0.51)	Material	5
Necessary *knowledge* (0.77); *understanding* of environment (0.63); *learner* (0.59)	Knowledge, learner, understanding	5
Large *scale* (0.68); m*acro scale* (0.58); micro *scale* (0.47); larger *scale* complex building (0.46)	Scale	4
Service (0.56)	Service, servicing	4
Related components *environment* (0.67); understanding of *environment* (0.63); *environment* (0.59)	Environment	3
Holistic low energy strategies (0.62)		3
Passive *design* principles (0.61); *Design* of cities (0.54)	Design	3
	Introduce	3
Design of *cities* (0.54); *cities* (0.52); *new towns* (0.51)	Cities, town	3
Larger scale complex *building* (0.46); complex *building* (0.44)	Building	2
Larger scale *complex* building (0.46); *complex* building (0.44)	Complex	2

Available online at: https://natural-language-understanding-demo.ng.bluemix.net/ (accessed June 23, 2022).

Figure 3 provides a step-by-step summary of the method. The first step was to develop a lexicon of keywords representative of each SDG. A core challenge stemmed from each SDG overlapping or interlinking to other SDGs, meaning that, a sustainability lexicon will usually apply to multiple SDGs. Lists of keywords on the topic of sustainability are available but provided without their SDG context. A non-exhaustive list of documents addresses SDG topics but it is the 17 UN SDGs, themselves, that bear the core text that is the reference for this and further work. Therefore, the initial sustainability lexicon comprised keywords extracted from the 169 targets and 247 indicators describing the SDGs and related sources (details are included in Supplementary Table S2 of the Supplementary material file accompanying this submission). From the compiled list of keywords, root keywords (RKs), e.g., "sustainab" for "sustainable," "sustainably," and "sustainability" were identified. The process identified 135 RKs, from the 17 UN SDGs, and a further 146 RKs from the World Health Organization (WHO) lexicon.

The second step required a method to analyze the available module descriptor texts in MS Excel spreadsheet tables. PMC databases can easily be exported in CSV and MS Excel files formats. To ensure RKs found were distinct from other words, case sensitive searches and entire word searches were prefixed with "#" and "$" characters, respectively. Examples: $#CO2, $#HIV, $aids, $air, $clean, $gas, $land, $safe, $safeguard, $sea, $unsafe, and $wind. Without the characters, the search algorithm will count unwanted words starting with that RK.

In analyzing each module/course descriptor, the number of occurrences of each RK was registered.

The third, and possibly the most crucial, step was to identify the SI scoring for each RK and their association with SDGs. To associate RKs to specific SDGs, the importance of each RK toward sustainability is given by its RK_{SI} score, calculated using an adapted version of term frequency–inverse document frequency (TF-IDF), given by Equation (1). TF-IDF aims to attribute a weighted value of terms or document frequency to evaluate the relevance of a word (Mishra and Urolagin, 2019). Tabulating the targets and indicators from the 17 UN SDGs, the SI score for each RK was calculated by counting the number of times the RK appears in targets and indicators and dividing that number by the number of unique SDGs in which the RK appears.

$$RK_{SI} = \frac{No.\ of\ times\ the\ RK\ appears\ in\ the\ targets\ and\ indicators}{No.\ of\ SDGs\ in\ which\ the\ RK\ appears} \quad (1)$$

Figure 4 shows the SI scores of RKs and their association to the SDGs. From the 281 RKs defined, 218 RKs were found in the 17 SDGs with a sustainable importance RK_{SI}, ranging from 1 to 15 with an average of 2.32. Each RK was associated to each SDG in which the word appeared. Across all 218 RKs, the RKs could be detected from 1 to all 17 SDGs with an average occurrence of 3.72 SDGs, and variance of 11.05 SDGs. It was found that 78 RKs associated to 4 or more SDGs. The number of RKs found in each SDG is plotted in Figure 4C. The sum of the sustainable importance attributed to RKs associated to each SDG was then
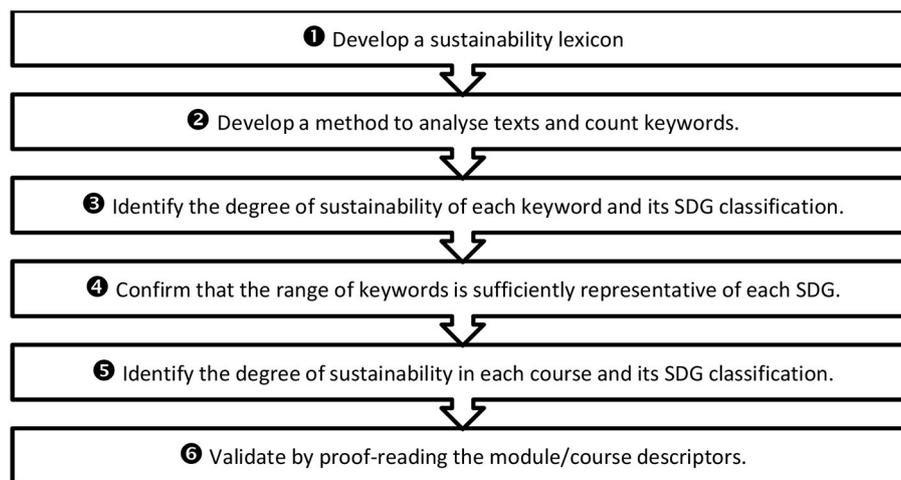
FIGURE 3
Process for analyzing sustainability in module/course descriptors at TU Dublin. Steps 1 to 5 are detailed in this paper.
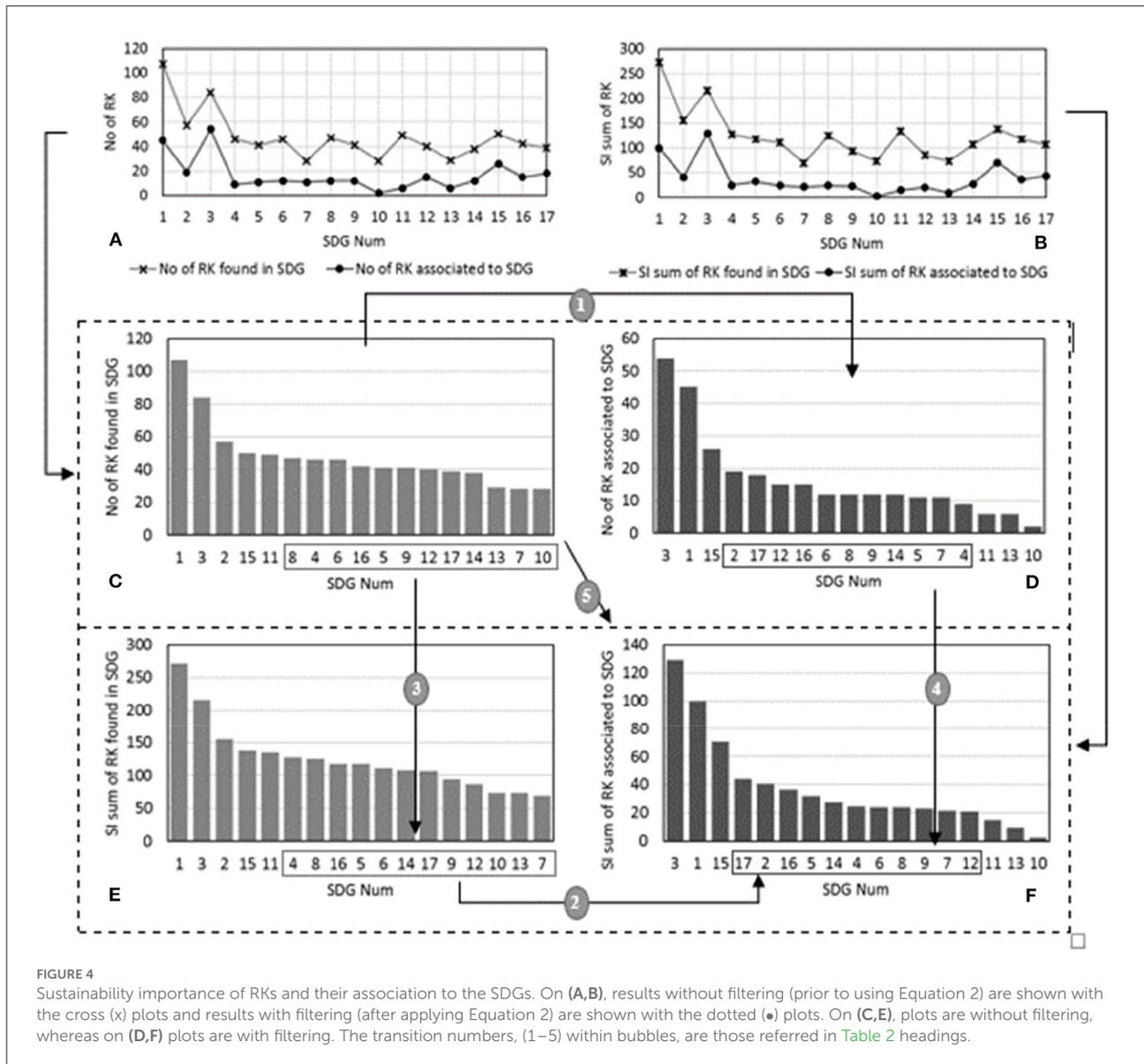
plotted in Figure 4E. To avoid having RKs being associated with too many SDGs and hence diluting their meaningful attribution, a filtering criterion, given by Equation (2), defined the minimum number of occurrences an RK needed to appear in an SDG to be associated to that SDG. To apply this filtering criterion, the maximum number of RKs was subtracted by the standard deviation of RKs across all SDGs. The outcome value was then rounded down to the nearest integer and added to 1 to define the filtering criterion.

$$Min.\ No.\ of\ RKs\ to\ appear\ in\ an\ SDG\ to\ be\ associated\ to\ it =$$
$$rounddown\left([RK]_{Max} - [RK]_{Stdev}, 0\right) + 1 \qquad (2)$$

The application of this filtering results in reducing the number of SDGs associated with RKs from 1 to 4 with an average occurrence in 1.31 SDGs and variance of 0.36 SDGs. This highlights the selective association of RKs to typically 1 or 2 SDGs. From the 218 RKs found in the SDGs, 166 RKs were associated to a single SDG, 38 RKs are associated to 2 SDGs, 13 RKs are associated to 3 SDGs and only 1 RK was associated to 4 SDGs. Figures 4D,F plot the number of RKs associated to SDGs and the ΣRK_{SI} values for RKs associated to each SDG, applying the filter. In Figures 4A,B, results without filtering (prior to using Equation 2) are shown with the cross (x) plots and results with filtering (after applying Equation 2) are shown with the dotted (•) plots. Figure 4A depicts the number of RKs founds (without filter) and associated (with filter). Similarly, Figure 4B depicts the ΣRK_{SI} found in SDGs (without filter) and associated to SDGs (with filter). The trend in the curves is maintained but the SDG position order is shuffled according to the method employed (Count of RKs v ΣRK_{SI}) and the filtering. In Figures 4B,C, plots are without filtering, whereas in

Figures 4D,F plots are with filtering. The transition numbers, (1–5) within bubbles, are those referred in Table 2 headings.

The fourth step was to confirm that the RKs selected were representative of each SDG. Variation in SDG positions, according to the decreasing number of RKs found and associated to each SDG as well as to the decreasing sum of RK_{SI} values found and associated to each SDG, highlights the "richness" of distinguishable RKs in the SDGs. Table 2 highlights the variations in the SDG positions, further analyzed in Figure 4. Transition (1), from (C) to (D), shows that the filter reduced the number of RKs associated to each SDG by more than half. SDG position shuffling and standard deviation (stdev) in Table 2 illustrate the strong influence of the filtering technique. This is also seen in transition (2), from (E) to (F), when examining the change in SDG positions using RK_{SI} values before and after the filter is applied. Without applying the filter, transition (3), from (C) to (E), shows that the position of SDGs with a larger number of RKs did not change if the RK_{SI} value was used. Therefore, the order of importance of SDGs given by the defined RK appeared to be influenced by the number of RKs initially found in each SDG. However, comparing stdev for transitions (3) and (4), the SDG positioning appeared to be less influenced by the number of RKs associated to each SDG than by the filter. Applying the filter, transition (4), from (D) to (F), SDG position shuffling was emphasized. Combining the filter with the RK_{SI} result in transition (5), from (C) to (F), the SDG position shuffling matched transition (1), illustrating the strong influence of the filter. The values plotted in (F) indicate the knowledge on the quantity and quality of RKs assigned to each SDG. A high value typically indicates a high number of meaningful RKs associated to an SDG. A low number, as seen for SDGs 10, 13, and 11, suggests a limited number of RKs associated to these

FIGURE 4
Sustainability importance of RKs and their association to the SDGs. On **(A,B)**, results without filtering (prior to using Equation 2) are shown with the cross (x) plots and results with filtering (after applying Equation 2) are shown with the dotted (•) plots. On **(C,E)**, plots are without filtering, whereas on **(D,F)** plots are with filtering. The transition numbers, (1−5) within bubbles, are those referred in Table 2 headings.

SDGs, and that these RKs have limited ability to distinguish documents addressing these SDGs.

The initial counts of RKs found in (C) for SDG 10 "Reduced inequality" and SDG 13 "Climate action" were low and, as filtering was applied, it appeared that most of the RKs could be better attributed to other SDGs. SDG 11 initially has a higher count of RKs associated than those for SDG 10 and 13. Nonetheless, the filter repositioned it to the third least interpreted SDG. It is notable that SDG 12 was highly ranked in (D) but is then lower ranked due to difficulty in associating RKs to that SDG. In other words, RKs employed in SDG 12 appear to have more significant meaning than in other SDGs. SDG 7, initially lowly ranked, achieves a higher rank by applying

the filter. The uneven distribution of the plots illustrates that the lexicon of RKs was richer for some SDGs than for others. The SI scoring highlights the meaningfulness of some RKs to specific SDGs but is insufficient to even the distribution. Hence, to ensure the sustainability of a module is detected and attributed the correct SDGs, the list of RKs needs to be further developed, in particular for SDGs 10, 13, and 11. It was suspected that the lack of distinguishable RKs for these SDGs stemmed from their "generic" nature where, for example, SDG 13 "climate action" integrates ideas found in many other SDGs. Similarly, SDG 10 "sustainable cities and communities" integrates ideas found in SDGs 1–3, 6–10, 12, and 15–17. This implies that module/course descriptors tackling one or more of these SDGs

TABLE 2 Variation of the SDG position based on the number of RK detected and their sustainable importance with and without the filter criteria.

| | Variation of the SDG ordered position… | | | | |
| | …from Found to Associated to… | | …from No. of RKs to SI sum of RKs… | | |
SDG #	(1) No. of RK	(2) SI sum of RK	(3) without filter	(4) with filter	(5) …from No. of RKs found to SI sum of RKs with filter.
1	−1	−1	0	0	−1
2	−1	−2	0	−1	−2
3	+1	+1	0	0	+1
4	−7	−3	+1	+5	−2
5	−2	+2	+1	+5	+3
6	0	0	−2	−2	−2
7	+3	+4	−1	0	+3
8	−3	−4	−1	−2	−5
9	+1	+1	−2	−2	−1
10	0	−2	+2	0	0
11	−10	−10	0	0	−10
12	+6	0	−2	−8	−2
13	−1	0	−1	0	−1
14	+3	+3	+3	+3	+6
15	+1	+1	0	0	+1
16	+2	+2	+1	+1	+3
17	+8	+8	+1	+1	+9
St. Dev:	4.130	3.710	1.372	2.849	4.130
Average:	0.000	0.000	0.000	0.000	0.000

The transition number given in Figure 4 is provided between brackets in the column headings (below).

could be attributed to SDG 10 and 13, i.e., the extent to which they could be attributed to "generic" SDGs was affected by RK overlap.

The fifth step was to identify modules with sustainable components and their SI scoring to identify their relevant SDGs. To determine the sustainability importance of a module and its relevant SDGs, in order of importance, the RKs associated to each SDG, after applying the filter, were searched for in the text. When an RK was found, its RK_{SI} value was added to its previous value for the respective SDGs with which the RK is associated, given by Equation (3).

$$SI\ Module_{SDG} = \sum_{SDG} RK_{SI} \quad (3)$$

To complete the analysis of TU Dublin's inventory of modules, the descriptive statistics for the sustainability importance of modules were summarized as: average SI Module$_{SDG}$ score of 14.39, standard deviation of 11.90, minimum of 0, and maximum of 90.83. These statistics allowed for categorization of sustainability importance (SI) of each module using the

definition of the importance range (IR), where f is a scaling factor from −4 to 4, given by Equation (4).

$$\frac{Max}{2} + (f - 0.5)^* stdev < IR \leq \frac{Max}{2} + (f + 0.5)^* stdev \quad (4)$$

$$Cumulative\ \% = \frac{\sum_{SI = "No\ RK\ Found"}^{SI = x} N(SI)}{N_{Total}} \quad (5)$$

Modifications were made to the "inexistent" sustainability range with the lower value equal to the Min (0) and to the "beyond" sustainability range, where the upper value was equal to the Max (90.83). Table 3 lists the importance range and the number of modules in that range. The cumulative percentage of modules matching a sustainable importance level or lower, presented in Table 3, is defined by Equation (5) with N(SI) the number of modules of Sustainable Importance SI and N_{Total}, the total number of modules analyzed. Figure 5 shows the resulting distribution of modules across the importance ranges. As a first approximation, we consider a module containing the RK "sustainab" in its description or syllabus was defined as a "sustainability-focused." From the total of 5,773 modules analyzed, 130 contained the word "sustainab." A statistical

analysis of sustainability importance (SI) of these modules was repeated, yielding the following results: average SI Module$_{SDG}$ score of 31.89, standard deviation of 17.24, minimum of 4.2, and maximum of 90.83.

The sixth and final step is to provide a validation of this computational technique in which its results are compared with those from expert interpretation on a random sample of modules/courses across the AASHE-STARS course categories. Only module/course descriptors approved by TU Dublin's quality assurance office were included in our adapted TF-IDF analysis. In-depth testing and validation will be the subject of future design-science research to develop the technique further.

Discussion

Table 4 offers a statistical definition of the AASHE STARS criteria for "sustainability-focused" modules and "sustainable-inclusive" modules. However, considering the significance of the word "sustainab," one might also choose to include modules as being "sustainability-focused" when the RK "sustainab" appears. Such a consideration requires further analysis.

The classification "Nothing" applied to 1,817 modules with a maximum SI Module$_{SDG}$ value less than or equal to 14.64 and, hence, were discarded from being "sustainability-focused" or "sustainability-inclusive" unless they were already defined as "sustainability-focused" by containing the RK "sustainab." Following the IR definition from Table 4, 234 modules were found with a maximum SI Module$_{SDG}$ value within the IR range for "sustainability-focused" and 908 modules were found to be within the IR range for "sustainability-inclusive." However, if modules containing the RK "sustainab" were classified as "sustainability-focused" then this lowered the number of modules found in "nothing" and "sustainability-inclusive," whilst increasing the number of "sustainability-focused" modules by 80. The resulting number of modules in each classification with the statistical results on the number of RKs used for each category is summarized in Table 5.

The average number of RKs in modules deemed as "sustainability-focused" was 15.9. The standard deviation of 6.9 illustrates a large distribution of RKs, partly due to modules containing the "sustainab" RK with a low number of RKs. "Sustainability-inclusive" modules had an average of 9.2 RKs with a small standard deviation of 2.5, illustrating a good selection that distinguished "sustainability-focused" from "sustainability-inclusive." One third of modules had RKs that did not appear in SDG targets and indicators. With a low average number and standard deviation of RKs used, it is most likely that only a small number of modules could be classified as "sustainability-focused" or "sustainability-inclusive." Yet, many of these modules could turn out to be sustainable-focused or sustainability-inclusive if documents containing these RKs were associated to SDGs. This would require large corpora

of documents and structured data analyzed with artificial intelligence or complex semantic algorithms, beyond the scope of this study.

For each "sustainability-focused" and "sustainability-inclusive" module, the list of RKs found was saved and the module is added to the list of modules associated to each SDG along with its SI Module$_{SDG}$ value. Due to the association of RKs to one or more SDGs, modules can be associated to multiple SDGs, in some cases as many as 15. Applying the filter and selecting SDGs for which the SI Module$_{SDG}$ value is within the defined importance range for "sustainability-focused" and "sustainability-inclusive" enabled association of modules up to a maximum of 5 SDGs. This analysis enabled the identification of the SDGs that were most and least addressed by each module. Figure 6 shows the number of modules associated with each SDG. The majority of modules appear to be associated to SDG 1 "no poverty," SDG 17 "sustainable partnership" and SDG 3 "health and wellbeing." The SDGs that appeared to be moderately addressed include SDG 7, 15, 6, 2, 4, 12, and 8 whilst the remaining SDGs had minimal consideration. The large number of modules associated to SDGs 1 and 17 may be due to the generic nature of the topics covered by these SDGs. This warrants further investigation. In essence, some of these modules may have wrongly associated medium to high ΣRK_{SI} values for these SDGs. A large number of RKs (between 25 and 55 RK) are associated to SDGs 1 and 3. A medium number of RKs, between 15 and 25 RK, are associated to SDGs 2, 15 and 17 as shown in Figure 4D. The number of RKs together with their respective SI values result in high ΣRK_{SI} values using Equation (3), varying between 40 and 120, attributed to SDG 3, 1, 15 and 17 as shown in Figure 4F. Whilst one might expect a high ΣRK_{SI} value to indicate good coverage of quality RKs for an SDG, it could also mean that too many RKs with low to medium SI values have been wrongly associated to an SDG. In other words, a "generic" SDG will cover many topics with, potentially, a large range of associated RKs, making it difficult to distinguish this SDG from others. Referring back to the previous example with SDG 10, some RKs that could have been correctly associated to SDG 10 may have been wrongly associated to SDG 1 and/or 17 because the adapted TF-IDF method used gives priority to these associations.

Having knowledge that several TU Dublin modules were geared to SDG13 "Climate action," the results suggest that there may be an issue with the distribution of RKs used across all SDGs or their applicability to several SDGs. To understand the origin of this issue, further analysis considered the ΣRK_{SI} associated with each SDG as plotted in Figure 4F, which represent the quantity and quality of sustainable RKs assigned to each SDG. The poor quality of distinguishable RKs for the "generic" SDGs 10 and 13 resulted in no modules being associated to them, but instead to more lexically defined SDGs. SDG 14 was associated to a single module. In Figure 4F SDG 14 is averagely ranked with an RK$_{SI}$ > 20 implying that there was a good understanding of

TABLE 3 Categorization of module sustainability importance by importance range (IR).

| Sustainability importance	f	Importance range		No. of modules	% of modules	Cumulative %
		$Max/2 + (f - 0.5)^* stdev$	$Max/2 + (f + 0.5)^* stdev$			
No RK Found				2,799	48.48	48.48
Inexistent	−4	0	3.75	439	7.60	56.09
Very Low	−3	3.75	15.65	1,489	25.79	81.88
Low	−2	15.65	27.55	688	11.92	93.80
Low+	−1	27.55	39.46	222	3.85	97.64
Average	0	39.46	51.37	86	1.49	99.13
Medium	1	51.37	63.28	36	0.62	99.76
High	2	63.28	75.19	9	0.16	99.91
Very High	3	75.19	87.09	4	0.07	99.98
Beyond	4	87.09	90.83	1	0.02	100.00
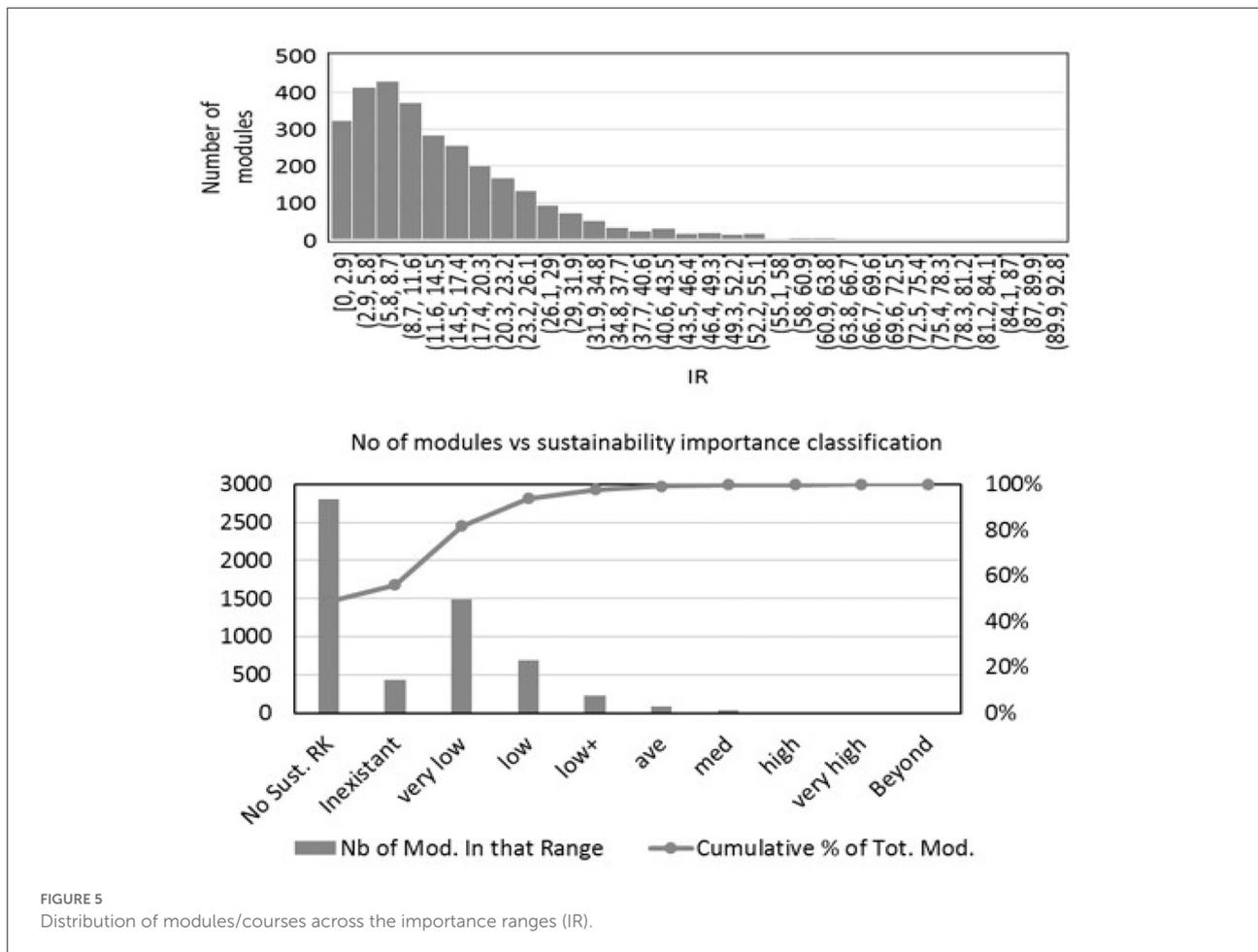
FIGURE 5
Distribution of modules/courses across the importance ranges (IR).

meaningful RKs in that SDG. This also implied that SDG 14 was not being addressed by TU Dublin modules. SDGs 7 and 12 also returned $RK_{SI} > 20$ as seen in Figure 4F, were found to be the fourth and ninth most studied SDGs in TU Dublin, respectively. Such high rankings lead to a conclusion that $RK_{SI} > 20$ is a reasonable threshold to effectively attribute modules to their most significant SDGs. Table 6 summarizes the number of RKs being used in modules containing, and absent of "sustainab."

The results demonstrate that modules with the word "sustainab" tended to use an enriched sustainability vocabulary

TABLE 4 Statistical definition of "sustainability-inclusive" and "sustainability-focused" modules based on SI values.

Definition	Literal Range	Importance range		No. of modules
		Min.	Max.	
Nothing	0 < Value < (Average −Std.)	0	14.6	1,817
Sustainability-inclusive	Value ≥ (Average −Std.) < Average	14.65	31.89	908
Sustainability-focused	Value ≥ Average	31.89	90.81	234

TABLE 5 Classification of modules with RK descriptive statistics.

Classification	No. of modules	% of modules	No. of RKs used			
			Min	Max	Average	Std. Dev.
Sustainability-focused	314	5.44	1	41	15.94	6.49
Sustainability-inclusive	848	14.69	4	19	9.19	2.50
RK not in SDG	1,797	31.13	1	12	3.45	1.88
No RK found	2,814	48.74				

FIGURE 6
Number of "sustainability-focused" and "sustainability-inclusive" modules associated to each SDG.
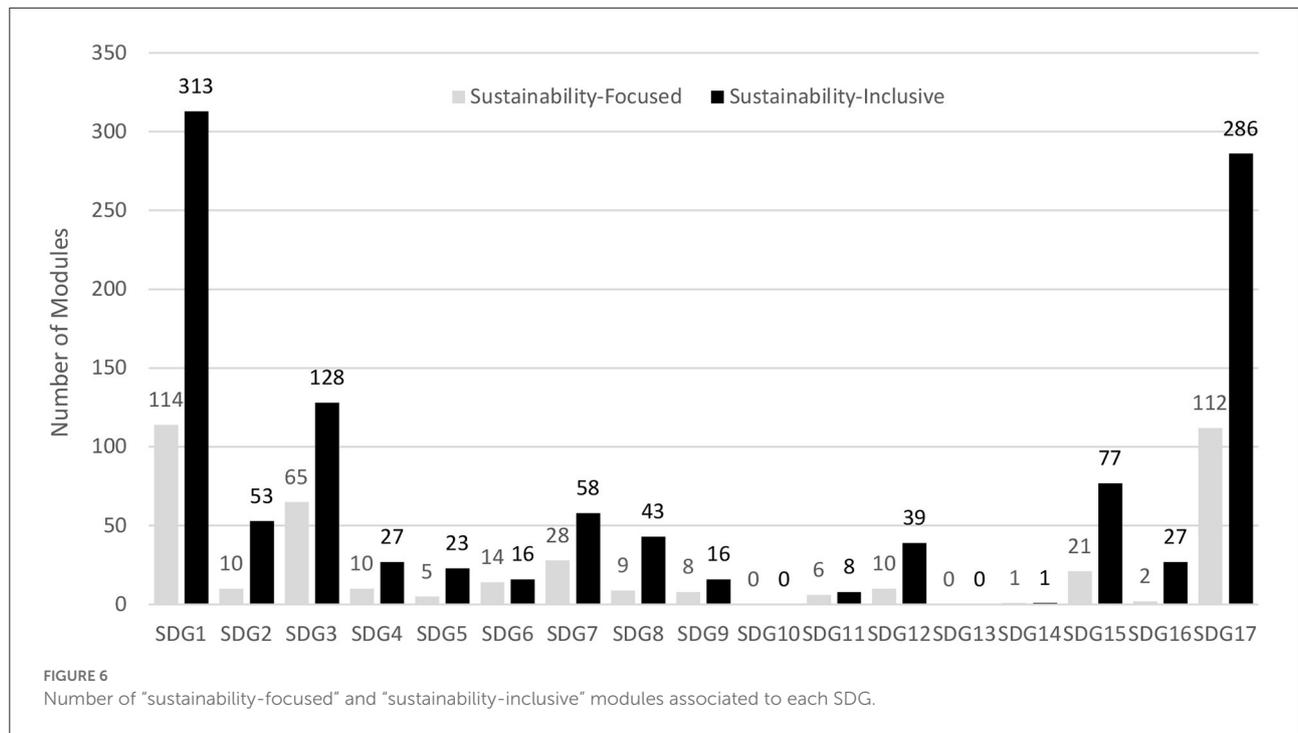
with more than twice the number of RKs found in modules that do not use "sustainab." However, some modules do not use "sustainab," yet have as much or more enriched vocabulary than modules using the "sustainab," thereby more being likely to be "sustainability-focused." To determine the probability for "sustainability-focused" and "sustainability-inclusive" modules using the average number of distinguishable RKs, the following criteria were applied.

Criteria:

Probability is "low" if: $0 < N \leq 50\% \times N_{average}$, i.e., if $1 \leq N \leq 6.56$

Probability is "med" if: $50\% \times N_{average} < N \leq 100\% \times N_{average}$, i.e., if $6.56 \leq N \leq 13.13$

Probability is "high" if: $N > N_{average}$, i.e., if $N > 13.13$

Table 7 shows that, from the 130 modules initially defined as "sustainability-focused," another 208 modules

TABLE 6 Number of RKs for modules containing or absent of the "sustainab" RK.

	No. of RKs used	
	Max	Average
Modules containing the RK "sustainab"	41	13.13
Modules that DO NOT contain the RK "sustainab"	35	6.09
Across all modules containing detected RK words	41	6.40

do not contain the "sustainab" RK and also have a high probability of being sustainability-focused. Modules that did not contain the "sustainab" RK and had a medium number of RKs in their description and syllabi were then most likely to be "sustainability-inclusive" modules. Cross referencing "low," "med," and "high" probability tags to the previously defined "sustainability-focused" and "sustainability-inclusive" modules, it was found that 286 modules (5% of all modules analyzed) could be defined as "sustainability-focused" with a medium to high level of confidence. Taking into account modules containing the RK "sustainab" and having a low number of RKs, it is also concluded that 769 modules (13.32% of all modules analyzed) could be defined as "sustainability-inclusive" modules with a medium to high level of confidence. Figure 7 summarizes the statistical categorization of TU Dublin modules offered at its city locations based on AASHE-STARS criteria. It indicates that another 1,847 modules contained sustainable RKs but could not be significantly associated to SDGs. Overall, it appears that "sustainability-focused" modules contained 16 ± 6 RKs while "sustainability-inclusive" modules contained 9 ± 2 RKs.

Concluding remarks

Based on the sustainability importance of root keywords (RKs), a sufficient lexicon was constructed to recognize modules that appear to address 14 of the SDGs. A list of over 1,200 potential RKs was initially identified as synonyms, antonyms and words associated with sustainability. This list of RKs was reduced to 135 RKs, from the 17 UN SDGs, and a further 146 RKs from the World Health Organization (WHO) lexicon and the relative frequency of RKs was applied to identify their sustainability importance. Adapting TD-IDF, the computational technique was extended to attribute modules to SDGs by analyzing the frequency of occurrence and, hence, the sustainability importance of RKs, recognized for their applicability to specific SDGs. The approach was implemented using a Curriculum Sustainability Interrogator (CSI) tool, based on automated macros in Microsoft Excel that mined the texts

within an individual module/course descriptor. Statistical definitions of "sustainability-focused," "sustainability-inclusive" modules (as defined by the AASHE STARS framework) were determined based on distinguishing the core SDGs addressed by modules with a sustainability component and their respective ΣRK_{SI} scores.
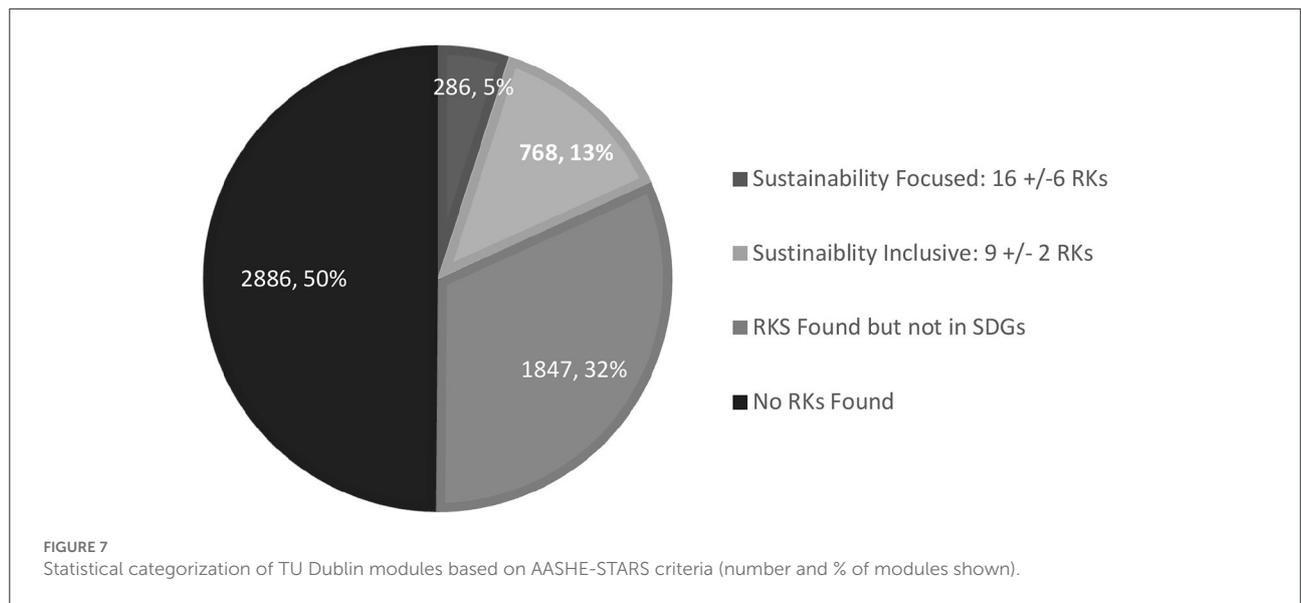
Among the 5,773 modules analyzed from TU Dublin's PMC, the technique identified 286 modules (~5%) as "sustainability-focused" modules with a medium to high level of confidence, and 769 modules (~13%) as modules with "sustainability-inclusive" with a medium to high level of confidence. Modules with a sustainable component appear to primarily address SDGs 3, 7 and 15 with a high level of confidence as well as the "generic" SDGs 1 and 17 with a lower level of confidence as modules may have been wrongly associated to such SDGs. However, some inherent limitations of the computational technique were noted in the analysis. For example, SDGs 10 and 13 are arguably "generic" SDGs in that they use a vocabulary that is more significantly attributed to other SDGs. Consequently, modules associated to SDGs with a vocabulary similar to SDG 10 and 13 could then be inappropriately associated to SDG 10 and 13 by default. Conversely, the lexicon of RKs must be enriched for SDG 11. Outcome from this study does not claim that the approach can be a substitute for meaningful engagement with faculty and students on how sustainability is to be embedded within their modules [e.g., (Freeman et al., 2021)]. However, it is recommended that a validation of the CSI tool's accuracy (step 6) is the primary opportunity for further research, in which subject-matter experts and sustainability experts would manually co-interpret a random selection of module/course descriptors for alignment to SDGs and for AASH-STARS categorization for purposes of reconciliation with the results from the CSI tool. Although machine learning methods based on large amounts of trained data were not used in this analysis, this study offers an avenue for future research focused on the application of advanced NLP to evaluating sustainability in the curriculum.

As university curriculum portfolios become more agile and interconnected, automating mappings of module/course descriptors, such as to the SDGs will become all the more important in guiding students through their learning experiences. Overall, this paper

TABLE 7 Probability analysis of "sustainability-focused" modules and "sustainability-inclusive" modules.

	$N =$ number of RKs detected				
	High $N > 13$	**Med** $7 < N \leq 13$	**Low** $1 \leq N \leq 6$	$N = 0$	**Total**
Number and (%) of modules containing "sustainab"	49 (0.85%)	53 (0.92%)	28 (0.49%)	0 (0.00%)	130 (2.25%)
Number and (%) of modules not containing "sustainab"	208 (3.60%)	823 (14.26%)	1,812 (31.89%)	2,800 (48.50%)	5,643 (97.75%)
Total number and (%) of modules	1,432 (24.81%)	1,116 (19.33%)	509 (8.82%)	2,716 (47.05%)	5,773 (100%)

5.85% : most likely to be **sustainability-focused modules**

14.26% : most likely to be **sustainability-inclusive modules**

	High $N > 13$	**Med** $7 < N \leq 13$	**Low** $1 \leq N \leq 6$	$N = 0$	**Total**
Sustainable-focused	206	80	28	0	314
Sustainability-inclusive	51	690	107	0	848
Nothing	0	106	1,706	2,799	4,611
Total	257	876	1,841	2,799	5,773

	High $N > 13$	**Med** $7 < N \leq 13$	**Low** $1 \leq N \leq 6$	$N = 0$	**Total**
Sustainability-focused	3.57%	1.39%	0.49%	0.00%	5.44%
Sustainability-inclusive	0.88%	11.95%	1.85%	0.00%	14.69%
Nothing	0.00%	1.84%	29.55%	48.48%	79.87%
Total	4.45%	15.17%	31.89%	48.48%	100.00%

4.95% : most likely to be **sustainability-focused modules**

13.32% : most likely to be **sustainability-inclusive modules**

The top table shows the number and percentage of modules containing or not the "sustainab" RK organized with the probability (high to low) of being sustainable. The middle table shows the number of modules found in each AASHE-STARS category. The bottom table shows the percentage of modules in each AASHE-STARS category with the probability of being what they are said to be.

FIGURE 7
Statistical categorization of TU Dublin modules based on AASHE-STARS criteria (number and % of modules shown).

has established that some RKs are more important than others, particularly so when mining textual data in abbreviated documents, such as those from academic module descriptors.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frsus.2022.909676/full#supplementary-material

References

AASHE (2015). *Stars® Technical Manual Version 2.0*. Philadelphia, PA: American Association for the Advancement of Sustainability in Higher Education.

Albareda-Tiana, S., Vidal-Raméntol, S., and Fernández-Morilla, M. (2018). Implementing the sustainable development goals at University level. *Int. J. Sustain. Higher Educ.* 19, 473–397. doi: 10.1108/IJSHE-05-2017-0069

Armitage, C. S., Lorenz, M., and Mikki, S. (2020). Mapping scholarly publications related to the Sustainable Development Goals: do independent bibliometric approaches get the same results? *Quantitative Sci. Stud.* 1, 1092–1108. doi: 10.1162/qss_a_00071

Bautista-Puig, N., and Mauleón, E. (2019). Unveiling the path towards sustainability: is there a research interest on sustainable goals? *Int. Confer. Scientometrics Informetrics* 17, 2770–2771. Available online at: https://dblp.org/rec/conf/issi/Bautista-PuigM19

Bolden, R., and Moscarola, J. (2000). Bridging the quantitative-qualitative divide: the lexical approach to textual data analysis. *Soc. Sci. Computer Rev.* 18, 450–460. doi: 10.1177/089443930001800408

Brugmann, R., Côté, N., Postma, N., Shaw, E. A., Pal, D., and Robinson, J. B. (2019). Expanding student engagement in sustainability: using SDG-and CEL-focused inventories to transform curriculum at the University of Toronto. *Sustainability* 11, 530. doi: 10.3390/su11020530

Brundtland, G. H. (1987). Our common future-call for action. *Environ. Conserv.* 14, 291–294. doi: 10.1017/S0376892900016805

Buttigieg, P. L., McGlade, J., and Coppens, L. (2015). "Clarifying terms in the SDGs: representing the meaning behind the terminology," in *2nd Meeting of the Inter-Agency Expert Group on Sustainable Development Goal Indicators* (New York, NY: United Nations Statistics Division).

Calder, W., and Clugston, R. M. (2003). International efforts to promote higher education for sustainable development. *Plann. Higher Educ.* 31, 30–44. Available online at: https://www.researchgate.net/publication/265108977_International_Efforts_to_Promote_Higher_Education_for_Sustainable_Development

Chang, Y. C., and Lien, H. L. (2020). Mapping course sustainability by embedding the SDGS inventory into the university curriculum: a case study from national university of Kaohsiung in Taiwan. *Sustainability* 12, 4274. doi: 10.3390/su12104274

Chen, M., Mussalli, G., Amel-Zadeh, A., and Weinberg, M. O. (2021). NLP for SDGs: measuring corporate alignment with the sustainable development goals. *J. Impact ESG Invest.* 2, 61–81. doi: 10.3905/jesg.2021.1.035

Cole, L. (2003). *Assessing sustainability on Canadian University campuses: development of sustainability assessment framework* (Doctoral dissertation). Royal Roads University, Victoria, BC, Australia.

Desa, U. N. (2016). *Transforming Our World: The 2030 Agenda for Sustainable Development*. Resolution A/RES/70/1 adopted by the General Assembly on 25 September 2015. Available online at: https://stg-wedocs.unep.org/handle/20.500.11822/11125

Freeman, O., Duffy, D., McAlpine, A., Nolan Dr, C., McMahon, C., and Walsh, L. (2021). Exploring the impact of authentic assessment on sustainability literacy through reflective and action-oriented tasks: a roundtable podcast. *Irish J. Acad. Prac.* 9, 7. doi: 10.21427/CSYC-2271

Glover, A., Peters, C., and Haslett, S. K. (2011). Education for sustainable development and global citizenship: an evaluation of the validity of the STAUNCH auditing tool. *Int. J. Sustain. Higher Educ.* 12, 125–144. doi: 10.1108/14676371111118192

Government of Ireland (2014). *National Strategy for Education for Sustainable Development, 2014 - 2020*. Dublin: Department of Education and Skills.

Available online at: https://www.gov.ie/en/publication/02952d-national-strategy-on-education-for-sustainable-development-in-irelan/ (accessed May 21, 2022)

Government of Ireland (2018). *The Sustainable Development Goals National Implementation Plan 2018-2020*. Dublin: Department of Communications, Climate Action & Environment. Available online at: https://www.gov.ie/en/publication/7cde9f-the-sustainable-development-goals-national-implementation-plan-2018-/ (accessed May 21, 2022)

Harvey, J., Dodd, D., Deegan, C., Freeman, O., MacMahon, C. H., and Williams, H. (2021). Cultivating a Community of Practice model to support and encourage innovative T&L practices to engage practitioners and enhance student success. *Irish J. Acad. Prac.* 9, 3. doi: 10.21427/8TG2-QF02

Holst, J., Brock, A., Singer-Brodowski, M., and de Haan, G. (2020). Monitoring progress of change: implementation of Education for Sustainable Development (ESD) within documents of the German education system. *Sustainability* 12, 4306. doi: 10.3390/su12104306

Janoušková, S., Hák, T., and Moldan, B. (2018). Global SDGs assessments: helping or confusing indicators?. *Sustainability* 10, 1540. doi: 10.3390/su10051540

Kamal, A. S., and Asmuss, M. (2013). Benchmarking tools for assessing and tracking sustainability in higher educational institutions: identifying an effective tool for the University of Saskatchewan. *Int. J. Sustain. Higher Educ.* 14, 449–465. doi: 10.1108/IJSHE-08-2011-0052

Kioupi, V., and Voulvoulis, N. (2019). Education for sustainable development: a systemic framework for connecting the SDGs to educational outcomes. *Sustainability* 11, 6104. doi: 10.3390/su11216104

Kosta, K. (2018). "Sustainability curriculum in UK university sustainability reports," in *Implementing Sustainability in the Curriculum of Universities*, ed W. L. Filho (Cham: Springer), 79–97.

Kuhlman, T., and Farrington, J. (2010). What is sustainability? *Sustainability* 2, 3436–3448. doi: 10.3390/su2113436

Lafortune, G., Fuller, G., Moreno, J., Schmidt-Traub, G., and Kroll, C. (2018). *SDG Index and Dashboards Detailed Methodological Paper*. New York, NY: Bertelsman Stiftung and Sustainable Development Solutions Network.

Lambrechts, W., and Ceulemans, K. (2013). "Sustainability assessment in higher education: evaluating the use of the auditing instrument for sustainability in higher education (AISHE) in Belgium," in *Sustainability Assessment Tools in Higher Education Institutions*, eds S. Caeiro, W. L. Filho, C. Jabbour, and U. M. Azeiteiro (Cham: Springer), 157–174.

Leal Filho, W., Manolas, E., and Pace, P. (2015). The future we want: key issues on sustainable development in higher education after Rio and the UN decade of education for sustainable development. *Int. J. Sustain. Higher Educ.* 16, 112–129. doi: 10.1108/IJSHE-03-2014-0036

Lewis, S. L., and Maslin, M. A. (2015). Defining the anthropocene. *Nature* 519, 171–180. doi: 10.1038/nature14258

Longhurst, J., Kemp, S., Mori, K., Hack, C., and Baughan, P. (2021). *Education for Sustainable Development Guidance*. Advance HE and QAA, United Kingdom. Available online at: https://www.advance-he.ac.uk/knowledge-hub/education-sustainable-development-guidance (accessed January 28, 2022).

Lozano, R. (2006). A tool for a graphical assessment of sustainability in universities (GASU). *J. Cleaner Prod.* 14, 963–972. doi: 10.1016/j.jclepro.2005.11.041

Lozano, R. (2010). Diffusion of sustainable development in universities' curricula: an empirical example from Cardiff University. *J. Cleaner Prod.* 18, 637–644. doi: 10.1016/j.jclepro.2009.07.005

Lozano, R., and Peattie, K. J. (2009). "Developing a tool to audit curricula contributions to sustainable development," in *Sustainability at Universities - Opportunities, Challenges and Trends*, Vol. 31, ed W. Leal Filho (Frankfurt: Peter Lang Publishing Group). p. 179–194.

Macgregor, C. J., Connell, A., O'Conor, K., and Sagar, M. (2019). "Using the learning in future environments (LiFE) index to assess James Cook university's progress in supporting and embedding sustainability," in *Sustainability on University Campuses: Learning, Skills Building and Best Practices*, eds W. L. Filho and U. Bardi (Cham: Springer Nature), 147–164.

Meadows, D. H., Meadows, D. L., Randers, J., and Behrens, I. I. I., W. W. (1972). *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. New York, NY: Universe Books.

Meschede, C. (2020). The sustainable development goals in scientific literature: a bibliometric overview at the meta-level. *Sustainability* 12, 4461. doi: 10.3390/su12114461

Mishra, R. K., and Urolagin, S. (2019). "A sentiment analysis-based hotel recommendation using TF-IDF approach," in *2019 international conference on computational intelligence and knowledge economy* (Dubai: ICCIKE), 811–815.

Mishra, R. K., Urolagin, S., Jothi, J. A., Neogi, A. S., and Nawaz, N. (2021). Deep learning-based sentiment analysis and topic modeling on tourism during Covid-19 pandemic. *Front. Computer Sci.* 3, 775368. doi: 10.3389/fcomp.2021.775368

OECD (2020). *SDG Pathfinder Tool*. Available online at: https://sdg-pathfinder.org/about (accessed March 14, 2022).

Pincet, A., Okabe, S., and Pawelczyk, M. (2019). *Linking Aid to the Sustainable Development Goals - A Machine Learning Approach*. OECD Development Co-operation Working Papers, No. 52. Paris: OECD Publishing.

Pinto, P. T. (2021). *The Sustainable Development Goal (SDGs) Words 'Poverty' and 'Sustainability' in Brazilian Research: A Preliminary Thematic Corpus-Based Analysis*. Cadernos de Linguística, no prelo. Available online at: https://interab12.abralin.org/

Pukelis, L., Puig, N. B., Skrynik, M., and Stanciauskas, V. (2020). OSDG–open-source approach to classify text data by UN Sustainable Development Goals (SDGs). *arXiv preprint arXiv:2005.14569*. doi: 10.48550/arXiv.2005.14569

QQI (2021). *National Framework of Qualifications*. Available online at: https://www.qqi.ie/what-we-do/the-qualifications-system/national-framework-of-qualifications (accessed March 16, 2022).

Raworth, K. (2012). *A Safe and Just Space for Humanity: Can We Live Within the Doughnut?* Oxford: Oxfam.

Raworth, K. (2017). A doughnut for the anthropocene: humanity's compass in the 21st century. *Lancet Planetary Health* 1, e48–e49. doi: 10.1016/S2542-5196(17)30028-1

Rieckmann, M. (2017). *Educación para los Objetivos de Desarrollo Sostenible: Objetivos de aprendizaje*. Paris: UNESCO Publishing.

Rivest, M., Kashnitsky, Y., Bédard-Vallée, A., Campbell, D., Khayat, P., Labrosse, I., et al. (2021). Improving the scopus and aurora queries to identify research that supports the United Nations Sustainable Development Goals (SDGs) 2021. *Mendeley Data* 2. doi: 10.17632/9sxdykm8s4.4

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E. F., et al. (2009). A safe operating space for humanity. *Nature* 461, 472–475. doi: 10.1038/461472a

Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., and Rockström, J. (2019). Six transformations to achieve the sustainable development goals. *Nat. Sustain.* 2, 805–814. doi: 10.1038/s41893-019-0352-9

Sato, S., Hashimoto, T., and Shirota, Y. (2020). "Evaluation for ESD (Education for Sustainable Development) to achieve SDGs at University," in *2020 11th International Conference on Awareness Science and Technology (iCAST)* (Qingdao: IEEE), 1–6.

Schurer, A. P., Cowtan, K., Hawkins, E., Mann, M. E., Scott, V., and Tett, S. F. B. (2018). Interpretations of the Paris climate target. *Nat. Geosci.* 11, 220–221. doi: 10.1038/s41561-018-0086-8

Shriberg, M. (2002). Institutional assessment tools for sustainability in higher education: strengths, weaknesses, and implications for practice and theory. *Higher Educ. Policy* 15, 153–167. doi: 10.1016/S0952-8733(02)00006-5

Smith, T. B., Vacca, R., Mantegazza, L., and Capua, I. (2021). Natural language processing and network analysis provide novel insights on policy and scientific discourse around Sustainable Development Goals. *Sci. Rep.* 11, 1–10. doi: 10.1038/s41598-021-01801-6

Sovrano, F., Palmirani, M., and Vitali, F. (2020). "Deep learning based multi-label text classification of UNGA resolutions," in *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance* (Athens), 686–695.

Stanny, C. J. (2016). Reevaluating bloom's taxonomy: what measurable verbs can and cannot say about student learning. *Educ. Sci.* 6, 37. doi: 10.3390/educsci6040037

Steffen, W., Persson, Å., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., et al. (2011). The anthropocene: from global change to planetary stewardship. *Ambio* 40, 739–761. doi: 10.1007/s13280-011-0185-x

Sutherland, I., Sim, Y., Lee, S. K., Byun, J., and Kiatkawsin, K. (2020). Topic modeling of online accommodation reviews via latent dirichlet allocation. *Sustainability* 12, 1821. doi: 10.3390/su12051821

Tierney, A., Tweddell, H., and Willmore, C. (2015). Measuring education for sustainable development: experiences from the University of Bristol. *Int. J. Sustain. Higher Educ.* 16, 507–522. doi: 10.1108/IJSHE-07-2013-0083

TU Dublin (2019). *Realising Infinite Possibilities. Strategic Intent 2030*. Technological University Dublin. Available online at: https://www.tudublin.ie/media/TU-Dublin-2030.pdf (accessed January, 16, 2022).

UNSDSN Asia Pacific (2017). *Getting Started With the SDGs in Universities: A Guide for Universities*. Higher Education Institutions and the Academic Sector. Available online at: https://resources.unsdsn.org/getting-started-with-the-sdgs-in-universities (accessed December 16, 2021).

Vladimirovaa, K., and Le Blanc, D. (2015). *How Well are the Links Between Education and Other Sustainable Development Goals Covered in UN Flagship Reports?: A Contribution to the Study of the Science-Policy Interface on Education in the UN System (October 2015)*. Available online at: https://www.un.org/esa/desa/papers/2015/wp146_2015.pdf (accessed March 23, 2022).

Waheed, B., Khan, F. I., Veitch, B., and Hawboldt, K. (2011). Uncertainty-based quantitative assessment of sustainability for higher education institutions. *J. Cleaner Prod.* 19, 720–732. doi: 10.1016/j.jclepro.2010.12.013

Walton, D. (2014). *Abductive Reasoning*. Tuscaloosa: University of Alabama Press.

Weiss, M., and Barth, M. (2019). Global research landscape of sustainability curricula implementation in higher education. *Int. J. Sustain. Higher Educ.* 20, 570–589. doi: 10.1108/IJSHE-10-2018-0190

Wiek, A., Withycombe, L., and Redman, C. L. (2011). Key competencies in sustainability: a reference framework for academic program development. *Sustain. Sci.* 6, 203–218. doi: 10.1007/s11625-011-0132-6

Yang, L., and Cormican, K. (2021). The crossovers and connectivity between systems engineering and the sustainable development goals: a scoping study. *Sustainability* 13, 3176. doi: 10.3390/su13063176