Venters, Colin, Seyff, Norbert, Becker, Christoph, Betz, Stefanie, Chitchyan, Ruzanna, Duboc, Leticia, McIntyre, Dan and Penzenstadler, Birgit

Characterising Sustainability Requirements: A New Species, Red Herring, or Just an Odd Fish?

Characterising Sustainability Requirements: A New Species, Red Herring, or Just an Odd Fish?

Colin C. Venters
University of Huddersfield
Huddersfield, UK
c.venters@hud.ac.uk

Norbert Seyff
University of Zurich
Zurich, Switzerland
seyff@ifi.uzh.ch

Christoph Becker
University of Toronto
Toronto, ON, Canada
christoph.becker@utoronto.ca

Stefanie Betz
Karlsruhe Insitute of Technology
Karlsruhe, Germany
stefanie.betz@kit.edu

Ruzanna Chitchyan
University of Leicester
Leicester, UK
rc256@leicester.ac.uk

Leticia Duboc
State Univ. of Rio de Janeiro
Rio de Janeiro, Brazil
leticia@ime.uerj.br

Dan McIntyre
University of Huddersfield
Huddersfield, UK
d.mcintyre@hud.ac.uk

Birgit Penzenstadler
CSU Long Beach
Long Beach, CA, USA
birgit.penzenstadler@csulb.edu

***Abstract*—Requirements articulating the needs of stakeholders are critical to successful system development and key to influencing their long-term effects. As the concept of sustainability has entered the discourse of a number of software-related computing fields, so has the term 'sustainability requirement'. However, it is unclear whether sustainability requirements are and should be different from how we already understand software requirements. This paper presents the results of a corpus-assisted discourse analysis study that explored the concept of sustainability requirements in order to understand how the term is being used in software and requirements engineering and related fields. The results of this study reveal that the term 'sustainability requirement' is generally used ambiguously and reveals significant segmentation across different fields. Our detailed analysis of selected influential papers highlights the segmented use of the term and suggests key focus questions that need to be addressed to establish a shared operative understanding of the term.**

I. Introduction

Modern societies are increasingly dependent on complex software and software systems, which underlie almost every aspect of day to day living from transportation, finance, education, retail, communication, governance, healthcare and fitness, entertainment and leisure, defense and security etc. [1].

Requirements are the foundation of all software products [2]. Failure to produce a set of software requirements that satisfies the primary needs of its users and other stakeholders can result in significant economic consequences [3]. As a result, the field of Requirements Engineering (RE) plays a critical role in software system development and is considered to be the key leverage point for practitioners who want to design sustainable software-intensive systems [4].

There are many ways to characterise sustainability [5], but it is generally defined as 'the capacity to endure' [6]. To increase the understanding and tangibility of this abstract concept, Tainter [7] suggests reflecting on four points when thinking about sustainability: What should be sustained? For whom? For how long? At what cost? As part of the concept of sustainable development, a widely adopted characterization proposed by Brundtland [8] emphasizes its focus on 'meeting the needs of the present without compromising the ability of future generations to meet their own needs'. This perspective emphasizes the view that development has effects that lie outside the system to be developed. The word 'need' is central to this definition and includes a dimension of time, present and future.

While sustainability is difficult to define as it needs context and (social) structure [9], it has become clear that the concept of sustainability requires simultaneous consideration of several interrelated dimensions [4]. Consequently, we characterize it using the following five dimensions: (1) *Environmental* refers to the responsible use and stewardship of natural resources. (2) *Economic* focuses on assets, capital and added value, which includes wealth creation, prosperity, profitability, capital investment, income, etc. (3) *Individual* covers individual freedom and agency. (4) *Social* is concerned with societal communities (groups of people, organizations) and the factors that erode trust in society. (5) *Technical* relates to the ability to maintain and evolve artificial systems over time.

The concept of sustainability has emerged as a topic of interest in different areas of computing such as artificial intelligence, high-performance computing, human-computer interaction, software engineering (SE), and requirements engineering. While there have been attempts to understand how sustainability is perceived in the practice of software engineering and how sustainability can become an inherent part of software engineering practice [10], consensus on what sustainability means in the field of software and requirements engineering is still emerging [11], [12].

The emergence of sustainability included the introduction of the term 'sustainability requirement' in software and requirements engineering literature with a number of approaches being proposed for eliciting, modelling, managing, and capturing reusable knowledge with regards to sustainability requirements [13], [14], [15], [16], [17].

However, it is unclear how this term is being defined, applied and understood. The term is often not formally defined but associated with one or more of the dimensions of sustainability such as environmental, economic, or social. Where definitions are proposed, it is unclear how these foster the shared understanding of the term, and whether they suggest that sustainability raises new or different requirements on software systems than other concerns such as security or safety [18], [19]. As a consequence, the different branches of computing interested in understanding and addressing sustainability issues use the term sustainability requirement in a variety of diametrically opposed ways resulting in fragmentation and the replication of effort in addressing common issues and themes. This fragmentation remains one of the main barriers and challenges for the field of computing making progress in addressing sustainability [20], [21]. Rather than seeking broad conformity of definitions, the aim should be to clarify how the terms are used by different communities in order to have a shared understanding [22].

To foster a shared understanding, this paper characterises how the term 'sustainability requirement' is being defined and used in the field of computing with a specific focus on software and requirements engineering. Thus, the research presented here was motivated by the question, how does current research construct the notion of 'sustainability requirements' through published work? To answer this question, we conducted a systematic search for relevant literature and performed a combined quantitative and qualitative analysis using corpus linguistics techniques on an identified body of literature. Our results allow us to better understand the meaning attributed to the term 'sustainability requirement' through its use in literature and highlight opportunities to integrate the segmented perspectives.

Section II provides background to sustainability in software engineering and to requirements. Section III presents the design of our study, which combines corpus linguistics and qualitative content analysis. Section IV discusses findings from concordance analysis, Section V investigates influential papers in detail, and Section VI examines the implications of the findings. The paper concludes in Section VII with recommendations for next steps toward a more comprehensive and consistent perspective on sustainability requirements.

II. Background

In order to lay the foundations for our work, we highlight the emergence of the term 'software sustainability' and its relationship to requirements and outline the use of the term 'requirement' as it is generally understood in the fields of software and requirements engineering.

A. Sustainability and Software Systems

From a software engineering perspective, a number of definitions and interpretations have emerged in relation to 'software sustainability', which range from a simple measure of time [23] to that of an emergent property [24]. Despite these diametrically opposed positions, software sustainability is increasingly being described in the literature as a first-class, non-functional requirement or software quality [18]. However, this has generally been made without explicit reference to the characteristics or qualities that sustainability would be composed of. Further research is needed to actually confirm or refute this. Nevertheless, it is argued that **maintainability** and **evolution** of the software artefact are key enablers to achieving long-living software [25]. As such, it has been argued that software sustainability as a composite, non-functional requirement is a measure of a number of quality attributes of a systems [11]. While this addresses the technical sustainability of developing long-living software, sustainability applies to both a system and its wider contexts [12]. However, the results of a study to explore perceptions and attitudes towards sustainability highlighted that requirements engineering practitioners' tended to have a narrow understanding of the concept of sustainability with a focus on environmental and economic sustainability [10].

B. Requirements in RE

The term requirement has been in use in the field of software engineering since the 1960's [26]. Software requirements express the needs and constraints placed on software that contributes to addressing a real-world problem. At its most basic, a software requirement is a property that the software system must exhibit. However, it is acknowledged that the term 'requirement' is not used consistently in the software industry [27].

The field of requirements engineering distinguishes between an actual need and a statement that represents that need, both of which are referred to as a 'requirement'. Additionally, requirements are often classified into two broad categories: functional requirements (FR) or non-functional requirements (NFR). In very simplistic terms, functional requirements describe what a system should 'do' and non-functional requirements describe what and how a system should 'be'.

A range of different definitions have been proposed for the term functional requirement. For example, Kotonya and Sommerville [28] define a functional requirement as simply being related to the systems functionality. Lauesen [29] extends this definition to include not only specifying the functions of the system, but how it records, computes, transforms, and transmits data. However, Sharma and Biswas [30] highlight that there are different granular levels of functional requirements that can be distinguished.

While there is a broad consensus regarding how to define the term 'functional requirements', the definition of 'non-functional requirements' (NFR) has proven to be more controversial within the software and requirements engineering communities. NFR are often classified by quality attributes such as performance, maintainability, interoperability etc., a number of which are enshrined in standards such as ISO/IEC 25010:2011 [31]. However, Glinz [32] suggests a faceted classification of requirements that removes the concept of a coarse separation into 'functional' or 'non-functional'.

C. *Observations*

In our Introduction we highlight the rising attention that is paid to concerns related to sustainability in fields close to software engineering, and the key role that requirements play in this area. It is through requirements that the various and often diverging perspectives of specific disciplines and domains in the context of specific software development are expressed and articulated; it is through requirements that the success criteria for development are established; and it is in this understanding that the sustainability effects of software-intensive systems begin [4]. However, the involved disciplines' perspectives have not been linked coherently yet, so that uses and meanings of the term sustainability vary considerably [20], [21]. The consequence is that contradictory recommendations are given for practice, and the scope of these is often unclear. Sustainability is used in different contexts, which give rise to different interpretations, but these may not be made explicit. This causes confusion and misunderstandings of what it means to 'develop sustainable software' in practice [10], and difficulties to establish clear research scopes that would help individual areas of research to articulate their role in software sustainability [20], [12].

III. Methodology

The overarching research question guiding our analysis is the following: *How does current research construct the notion of sustainability requirements through published work?* To answer it, we conduct a corpus-assisted discourse analysis [33]. Corpus analysis techniques were used on the whole data set, while qualitative text analysis was used to arrive at a more nuanced view and support the interpretation of how the terms identified through quantitative analysis are situated and used. We describe how we constructed the corpus and then outline how quantitative and qualitative analysis methods complemented each other in discourse analysis.

A. *Corpus Development*

We constructed a corpus by searching for articles containing the term 'sustainability requirement*[1]' in ACM Digital Library[2], IEEE Xplore[3], ScienceDirect[4], and Scopus[5]. The initial search of these sources retrieved a combined total of one hundred and fifty one (151) papers that were published between 2000-2016. Table 1 shows the source, search parameters - fields searched, date range, and subject area - and the number of retrieval results.

Source	Fields	Range	Subject Area	Result
ACM DL	All	1947-	-	20
IEEE Xplore	All	1988-	-	26
ScienceDirect	All	1823-	Comp. Sci.	34
Scopus	All	1960-	Comp. Sci.	71

Table I: Corpus Sources

[1]A wildcard (*) was used to identify plurals.
[2]ACM DL: http://dl.acm.org/
[3]IEEE Xplore: http://ieeexplore.ieee.org/
[4]ScienceDirect: http://www.sciencedirect.com
[5]Scopus: https://www.scopus.com/

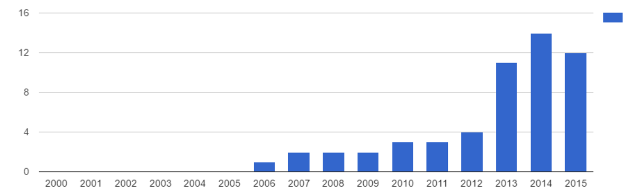

Figure 1: Number of publications and use of the term 'sustainability requirements' over time.

The removal of duplicates resulted in a set of one hundred and twenty seven articles (127). The corpus was further refined to remove papers that were considered (a) out of scope[6], (b) did not contain the term 'sustainability requirement*' in the full body of the text i.e. pointed to references that contained the term, (c) the full text of the article was not available, or (d) were conference or workshop proceeding volumes. This process resulted in a corpus consisting of fifty nine (59) articles from the field of computing. Figure 1 plots these papers over time (excluding 2016) and shows that the use of 'sustainability requirement' is relatively recent.

B. *Data Analysis Methods*

Corpus linguistic techniques and qualitative content analysis were employed iteratively to analyse the corpus.

1) Corpus Linguistics: Corpus linguistics is an approach to investigate language use that employs both quantitative and qualitative methods to aid in the analysis of large bodies of machine-readable text in order to uncover and interpret linguistic patterns [33]. A corpus is a collection of machine-readable text that is authentic and sampled in a way that is representative. In order to reliably establish how frequently a word or phrase occurs, statistical analysis of the frequency data is performed.

For the purposes of this study, we employ *Concordance Analysis*, a linguistic analysis technique routinely used for the exploration of a corpus [33]. The approach employs a range of techniques including word frequency analysis; keyword analysis; the generation of n-grams (i.e. repeated sequences of words); and collocation. We analysed the corpus using these standard methods, of which the most relevant for our purposes was collocation analysis. AntConc[7] was used to generate collocates and carry out concordance analysis.

The corpus was searched for collocates of **sustainability** and **sustainability requirement*** with a minimum frequency setting of twenty (20) and a collocate window span of four (4) words either side of the node word or phrase (4L to 4R). Four words either side of the node word is a generally accepted setting within corpus linguistics as there is little evidence to suggest that larger window spans produce reliable

[6]A paper was considered out of scope if it did not fall into the field of computing e.g. urban energy conservation and thermal performance [34], domestic building energy efficiency [35], an educational competence framework in the field of sustainable manufacturing [36] etc.
[7]AntConc: http://www.laurenceanthony.net/software

results [33] although Mason [37] points out that not all words influence their environment in the same way. Collocational strength can be tested by calculating an Mutual Information (MI)[8] score as collocational strength is a more reliable measure than raw frequency. A minimum frequency of twenty (20) was set in order to avoid low frequencies affecting its reliability. MI scores of three (3) and above are indicative of strong collocation. While collocational behaviour can be assessed purely by reference to the frequency of collocates [38], it is more reliable to calculate the statistical significance of collocates. This is because simple frequency is not necessarily indicative of collocational strength.

Since it is widely acknowledged that concordance analysis alone is limited in interpreting and understanding a corpus of literature, this is complemented with interpretive perspectives. In order to explore linguistic phenomena in more detail and interpret the use of terms across specific influential works, qualitative methods were employed to aid in the interpretation of linguistic patterns within their immediate context and to provide deeper insights.

2) Qualitative Content Analysis: Within the framework of Qualitative Content Analysis [39], we used a mixed approach of inductive category development and deductive category application [40], [41].

In the process of analysing how the term 'sustainability requirement' has been used, we identified common themes and differences across various disciplines. We did so through a triangulation process whereby sub-groups consisting of three co-authors read the articles individually, assigned each to a subject area and identified additional facets such as the dimensions of sustainability covered per article. Where discrepancies in classifications arose, the sub-group discussed this to reach consensus. The result allowed the corpus to be segmented into six subject areas for more detailed analysis. Citation counts on Google Scholar[9] were then used to identify the highest cited articles within each area.

In the second stage, these were analysed in detail to identify emerging themes and characteristics. An initial codebook was created and agreed upon by all co-authors of this paper, which was updated with each following coding activity. We extracted explicit definitions on sustainability and implicit motivations, and identified classifications of sustainability requirements, as well as key entities, concepts, dimensions, relationships, motivations, and actors. Six sub-groups were formed consisting of two co-authors per subject category. Each member of the sub-group read the articles individually and then coded the text with conceptual categories relevant to sustainability and sustainability requirements. This process allowed each member to work independently as well as peer-review each others analysis. A web-based text analysis tool[10] was used to support the coding and review.

[8]The Mutual Information (MI) score is a measure of the strength of association between words x and y. MI is calculated on the basis of the number of times the pair of tokens are observed together versus the number of times the tokens are observed separately.

[9]GoogleScholar:http://scholar.google.com

[10]SaturateApp: http://www.saturateapp.com

Rank	Freq.	Freq (L)	Freq. (R)	MI	Token
1	20	17	3	6.59	theme
2	123	70	53	6.24	dimension
3	163	73	90	6.14	dimensions
4	30	20	10	5.91	improving
5	103	12	91	5.83	debt
6	34	31	3	5.78	integrate
7	70	58	12	5.78	generic
8	230	195	35	5.72	social
9	21	21	0	5.61	integrating
10	22	20	2	5.58	corporate

Table II: Top 10 Collocates of the term 'sustainability'

3) Presentation of results: In the following sections, we will move from an overview of quantitative results over the entire corpus and their interpretation (Section IV) to a set of selected papers that we discuss in depth based on qualitative coding (Section V). This leads to a synthesis in Section VI which highlights salient observations and draws conclusions.

IV. Sustainability in Context

The following sections introduce the results of the concordance analysis for the terms **sustainability** and **sustainability requirements**. We present the results of the collocates analysis and discuss the significance of the findings for each of the terms individually.

A. Collocates of Sustainability

There are 178 statistically significant collocates types and 16588 collocate tokens of sustainability. Table II shows the ten most significant collocates based on MI scores.

The most frequent collocate of sustainability is **theme**. In 17 of the 20 occurrences, it is the head of a noun phrase that is post-modified by sustainability (e.g. 'the theme of corporate sustainability'). 16 of these instances are definite noun phrases (i.e. pre-modified by the definite article the), which in pragmatic terms gives them the status of existential presuppositions; that is, a definite reference to *'the theme of sustainability'* presupposes the existence of sustainability, thereby naturalising and legitimising the concept. The key point to note here is that sustainability is never introduced as a questioned concept.

Sustainability is thought of in various **dimensions**, which is ranked second and third. The concordances for these collocates reveal that where they occur as a head noun, the adjectives that pre- or post-modify it are *environmental*, *economic*, *social*, *individual*, *technical*, *human* and *general*.

The fourth ranked collocate, is **improving** (i.e. 'improving sustainability'), which functions to cast sustainability as a scalar rather than an absolute concept. The significance of this is that it reveals an assumption that sustainability is not a binary concept, i.e. that there is no simple distinction to be made between sustainability and *unsustainability*. This is supported by the presence of *more* as a strong collocate (MI 4.86) of the related word *sustainable*. 49 of the 56 instances of 'more' as a collocate occur as a bigram (more sustainable). The implicature in such usage is that any approach which is *'more' sustainable* remains, by default, unsustainable.

concerns	policies	environmental	aware
explicit	considering	aspect	goals
individual	economic	concern	indicators
consideration	technical	aspects	objectives
requirements	engineering	goal	modelling

Table III: Frequently collocated terms 11-30 (MI: > 4.7)

In position 5 is **debt**. All but 5 occurrences are as a pre-modified noun phrase (sustainability debt). A search of the 1.9 billion-word GloWbE corpus[11] reveals that 'debt' has an overwhelmingly negative semantic prosody as a consequence of its habitual collocates (its top collocate in GloWbE is *crisis*). The importance of this is that when the phrase *sustainability debt* is used, it is likely to invoke negative perceptions of the concept of sustainability generally.

In position 6 and 9 is **integrate** and **integrating**. Both refer to the practice of trying to incorporate the concept of sustainability into e.g. software design, with the consequent implications that this is desirable and has not, to date, been achieved. **Generic** (rank 7) refers primarily to a 'generic sustainability model' and **social** (rank 8) refers to one of the dimensions of sustainability. In position 10 is **corporate**, which also functions as a pre-modifier of sustainability. What is significant here is that this implies the existence of *forms of sustainability* that are not corporate. This begs the question why *corporate sustainability* should be any different from sustainability in other fields.

Table III lists the next 20 statistically significant collocates. It illustrates the emerging language used to discuss sustainability requirements in software-related disciplines and the need to develop explicit indicators that represent the multi-dimensional aspects across stakeholder groups.

It is interesting to see in which context these terms surface. For example, a sample of 10 (out of 37) occurrences of 'explicit' shown in Table IV illustrates how authors see the need to advocate for the theme of sustainability through emphasizing the need to state it explicitly.

While collocational strength as measured by an MI score is more reliable than a simple frequency measure, it is still worthwhile looking at collocates by frequency in order to see which of the dimensions of sustainability is mentioned most often in the corpus. In the frequency list of collocates, the first of the dimensions to be listed is **environmental** in position 11 (f=304) followed by **social** in position 8 (f=230), **economic** in position 15 (f=115), **technical** in position 23 (f=100), and **individual** in position 16 (f=73).

The overview of highly ranked collocates reveals a number of interesting features of usage in the corpus. Sustainability is naturalised as a concept as a result of functioning as the head of definite noun phrases which give rise to existential presuppositions. It is also is scalar in nature, which suggests that 100% sustainability is unachievable. The phrase **sustainability debt** hints at this without stating it explicitly.

[11]Global Web-based English: http://corpus.byu.edu/glowbe/

KWIC	File
open creation process with an explicit focus on sustainability.	Becker_2015
designed in complience with explicit strategic sustainability policies?	Bolis_2016
management - can be undertaken with explicit support for sustainability.	Chitchyan_2015
architectural design decisions and sustainability requirements explicit.	Ojameru-aye_2016
thereby created an explicity support of sustainability at RE'13.	Penzen-stadler_2013

Table IV: Concordance of 'explicit'

Rank	Freq.	Freq(L)	Freq(R)	Score	Collocate
1	48	0	48	4.54	requirement
2	247	13	234	4.19	requirements
3	22	3	19	3.11	patterns
4	13	12	1	1.98	technical
5	11	11	0	1.34	social
6	11	9	2	0.76	environmental
7	19	7	12	0.73	we
8	59	37	22	0.54	for
9	10	5	5	0.52	engineering
10	18	4	14	0.31	software

Table V: Collocates of the term 'sustainability requirement*'

B. Collocates of Sustainability Requirement

There are 10 statistically significant collocates of **sustainability requirement***, shown in Table V, with a total number of 29 collocate types and 1045 collocate tokens.

The top collocate of **sustainability requirement*** is requirement. This is because what follows the use of **sustainability requirement*** tends to be a description that adds a level of specificity to what has just been stated. The **requirements** collocate functions in a similar way. Analysing the concordance of 'sustainability requirements*' reveals pre-modifying adjectives before the target phrase (e.g. 'technical sustainability requirements') that imply that there are various sub-categories of sustainability requirement. This is seen in collocate 4, **technical**, collocate 5, **social** (i.e. 'social sustainability requirements') and collocate 6, **environmental**. The token **patterns** is ranked highly because it turns up in the phrase sustainability requirement patterns, as revealed by the concordance plot. However, that plot also reveals that it is not well dispersed across the corpus and turns up in just three papers. The reason that **we** appears as a collocate of sustainability requirement* is perhaps because use of the phrase sustainability requirement* necessitates some discussion of what authors mean when they use it. Inevitably, then, the plural pronoun is invoked. It seems unsurprising that **engineering** and **software** are collocates given that sustainability requirement* is discussed in the context of software and requirements engineering. **For** is ranked 8 in the collocates list. This appears to be because when people discuss sustainability requirements, they do so with reference to their *purpose* (e.g. 'sustainability requirements for software systems') or with reference to what is needed in order to achieve them (e.g. 'goal refinement as a checklist for sustainability requirements').

	Sys Eng	SW Eng	Req Eng
environmental	9 (5.7)	6 (5.1)	19 (4.9)
economic	-	16 (4.7)	15 (5.1)
social	6 (6.1)	17 (4.7)	10 (5.2)
individual	-	8 (5.0)	25 (4.7)
technical	-	23 (4.4)	26 (4.7)

Table VI: Sustainability dimensions across research areas

Area	Most cited (count)	2nd most cited (count)
Sys. Eng.	Büyüközkan [43] (135)	Brent [44] (135)
SW Eng.	Naumann [45] (110)	Mahmoud [46] (29)
Req. Eng.	Makropoulos [47] (196)	Penzenstadler [48] (46)
Inf. Sys.	Lutze [49] (30)	Pinder [50] (17)
ICT	Harmon [51] (118)	Ancillotti [52] (95)
Ergo	Radjiyev [53] (12)	Bolis [54] (1)

Table VII: Most influential papers per research area

C. Summary

The analysis highlights the context in which sustainability requirements arise and the key themes that pervade the corpus we analyse. Different dimensions have relevance, and sustainability needs to be improved through integrated perspectives that bridge specific aspects and generic models. However, it is worth noting that the time dimensions arises neither in the collocates nor in concordance analysis; virtually no words related to this dimension surface. It remains unclear how the different areas covered in this analysis articulate these specific concerns. As such, we perform a detailed content analysis on selected articles, supported by additional statistics, to characterise how and where sustainability requirements are discussed.

V. Analysis

A. Areas and themes

From the overall corpus, we identified tendencies in the characterization of sustainability requirements that aligned with the publications' research domains. Therefore, we categorized the corpus according to subject areas to conduct further qualitative analysis. Six overall research areas appear in the corpus: (1) Information Systems (IS, 8 papers), (2) Information and Communication Technologies (ICT, 4 papers), (3) Ergonomics (Ergo, 2 papers), (4) Systems Engineering (Sys Eng, 18 papers), (5) Software Engineering (SW Eng, 10 papers) and (6) Requirements Engineering (Req Eng, 19 papers). The full set of codes used for the analysis as well as the raw data is provided at [42]. All six groups were analysed in detail and summarised below.

B. Definitions

A number of the papers explicitly refer to the Brundtland [8] definition of 'sustainable development' or to the lexical definition of sustainability as the *capacity to endure or maintain*. However, the term sustainability is often not formally defined but simply used without specific explanation. Where no explicit definition exists, its use can be understood in relation to one or more of the dimensions of sustainability – environmental, economic, individual, social, or technical. However, while the *environmental* dimension is most often referenced, association with different dimensions varies significantly across the areas. Table VI characterizes the coverage of the key dimensions of sustainability in the discussions across the identified subject areas. Each cell provides the collocate rank, followed by the MI score in brackets. The awareness of all areas clearly is highest in the general discourse of requirements – this corresponds to the role of requirements in translating concerns of the system development context into actionable specifications of system requirements, including system qualities and constraints.

Similarly, the vast majority of the papers provided no explicit definition or clarification of 'sustainability requirement', but simply use the term in isolation. Most instances were implicitly related to one or more of the dimensions of sustainability, or considered a software quality, i.e., non-functional requirement. In only two of the corpus articles the term sustainability requirement was formally defined. Roher and Richardson [15] defined a sustainability requirements as 'requirements that may be used to specify system behavior (e.g. requirements that will reduce a system's energy consumption) as well as to influence the users' behavior (e.g. the system incentivizes sustainable actions)'. In contrast, Huber, Hilty, and Glinz [55] defined a sustainability requirement as 'a requirement for a sustainable software system which concerns sustainability'. While the second definition remains generic, the first one essentially points towards a subset of functional requirements and towards user influence. Consequently, neither one provides tangible guidance towards how to work with this type of requirement.

C. Influential perspectives

To complement these high-level views, we identified the two highest-cited articles in each of the six areas, shown in Table VII. We analysed these to identify the key motivators and actors in the discourse on sustainability requirements, and characterise how they define and describe the dimensions and their relationships. Table VIII summarises salient characteristics of the use of sustainability requirements in the most influential papers across the identified six domains.

Given the core focus of this article, we focus our detailed discussion of individual papers on the areas of Systems Engineering, Software Engineering, and Requirements Engineering. Table IX illustrates focal areas beyond these dimensions by providing a list of the highest ten collocates and the range of their MI score per area. Dimensions are given in bold. The three areas are discussed in detail below.

a) Systems Engineering: Both Büyüközkan et al. [43] and Brent et al. [44] are **motivated** by the goal to optimize systems considering sustainability issues. They argue that economic expectations such as reducing cost are **key concepts** relevant in supply chain management and health care waste management, respectively. Büyüközkan et al. [43] point out that recently environmental consciousness has become an important focus in their domain. Both papers **define** sustain-

Area	Key concepts	Motivation	Main actors	Sustainability requirement context
IS	Cost effectiveness Process improvement Process structuring	Improve cost effectiveness of process, aiming for cost reduction.	Business, Regulators, Customers	Metrics and controls context, "such as operating and capital cost, safety, energy cons., waste gen., efficiency"
ICT	Optimisation of IT infrastructure, Green computing, Environmental sustainability, Sustainability of IT services, Longevity of energy systems	Improved resource and energy efficiency of ICT	Customers, employees, business partners, NGOs	Environmental sustainability related to energy consumption and performance
SW Eng	Software development process models	Environmental impacts of ICT	Software developers, administrators, users	Implicit non-functional qualities
Sys Eng	Optimize systems considering sustainability issues	Economic expectations and environmental consciousness	All stakeholders in context, noting they have varying background	Sustainability requirements have to be communicated
Ergonomics	Multi-dimensional understanding with economic, social, and environmental	Economic and business-strategic aspects, human factors	Wide range of stakeholders, including all designers	Environmental context and long life cycles
RE	Multi-dimensionality of sustainability, Interdependence of dimensions, Trade-offs, General models of sustainability	Make sustainability more tangible, Make related goals explicit, Assess sustainability	Decision making households and/or software professionals, regulators	Multiple dimensions and trade-offs: 'Achieve acceptable level of service (...), have min. impact on natural env., be socially and economically acceptable'

Table VIII: Coverage of sustainability aspects in influential papers of key areas

	Sys Eng	SW Eng	Req Eng
1	Orientation	Aware	Improving
2	Indicators	Concerned	index
3	safety	Considering	Incorporating
4	requirement	Concept	integrating
5	connectivity	Dimensions	advocate
6	**social**	Debt	dimension
7	requirements	**environmental**	dimensions
8	performance	**individual**	chair
9	**environmental**	Define	systematically
10	network	Metaphor	**social**
MI	(8.4,5.1)	(5.6,4.9)	(5.6,5.2)

Table IX: Top collocates across selected research areas

ability. Büyüközkan et al. [43] argue that sustainability is "using resources to meet the needs of the present without compromising the ability of future generations to meet their own". Brent et al. [44] refer to the Brundtland definition [8] and highlight that the concept of sustainability and sustainable development can be understood intuitively, but it remains difficult to express it in concrete, operational terms. However, both Brent et al. [44] and Büyüközkan et al. [43] agree that sustainable development is about achieving environmental, economic, and social welfare for present as well as future generations.

Both share the view that stakeholders have to communicate **sustainability requirements**. In [43], the customer requirements represent the sustainability requirements of a supply chain. However, neither provides a formal definition of the term sustainability requirement.

Büyüközkan et al. [43] argue that key **actors** are all stakeholders in companies benefiting from efficient supply change management. Brent et al. [44] argue that the variety in backgrounds of stakeholders such as health care personnel and patients needs to be adequately considered in requirements elicitation activities.

b) Software Engineering: Both Naumann et al. [45] and Mahmoud and Ahmad [46] are **motivated** by the concern that *ICT has a negative impact on the environment* due to its increasing resource and power consumption, and aim to improve resource or energy efficiency of ICT. For example, Naumann et. al. [45] argue that it is not clear whether the resource and energy savings through ICT overbalance the resource and energy consumption by ICT.

The key **concepts** in both these papers are *software development process models* with a strong emphasis on *Green Computing* and *Environmental Sustainability* as well as *sustainability assessment metrics*. While both papers acknowledge that the *impacts of ICT on sustainable development* should also include *human and social sustainability*, their focus is environmental. Neither defines **sustainability** directly. Instead, both cite a definition of *sustainable software* as 'software that meets its (realistic) sustainability objectives, expressed in terms of direct and indirect impacts on economy, society, human beings, and environment that result from its definition and deployment' [56].

Similarly, the term **sustainability requirement** is not formally defined, though implicitly these requirements are considered to be non-functional qualities. In [46], these are specific to environmental concerns, while in [45] they relate to other identified dimensions as well.

The key **actors** here are software developers, administrators, and software users. Mahmoud and Ahmad [46] argue that in addition to supporting all the all stakeholders from developers to users in creating, maintaining, and using software in a 'more environmentally sustainable' way, the role of software itself in maintaining and optimizing energy usage in ICTs must be considered. However, Naumann et. al. [45] argue that a 'green and sustainable' software product can only be achieved if the developing organization is aware of negative and positive impacts on software development. In order to enable the various stakeholders to recognize these impacts, it is necessary to institutionalize their assessment and recognition in the applied software development processes.

c) Requirements Engineering: While Penzenstadler and Femmer [48] is concerned with providing a generic reference model for decomposing sustainability goals along different dimensions, Makropoulos et. al., [47] analyses the specific case of sustainable urban water management. The **motivation** for both approaches is to *make sustainability more tangible to allow its assessment and to make sustainability goals explicit*.

Neither paper **defines** sustainability or sustainability requirement, but both give examples of such requirements for their application domains. For instance, Makropoulos et. al., [47] state that *'sustainability of the urban water environment Water management in new residential developments is tasked with the requirement to achieve acceptable levels of service, not overburden existing infrastructure, have minimal impacts on the natural environment and to be, at the same time, socially and economically acceptable.'*

A key **concept** here is the *multi-dimensionality of sustainability*, noting the environmental, social, economic, and technical dimensions, with Penzenstadler adding an extra *individual* dimension. *Interdependencies* between the dimensions are emphasized and lead to *trade-offs* between related requirements, such as lowest possible energy usage versus lowest possible water usage for washing machines [47]. Both papers present *general models* of sustainability that are to be contextualized in order to *assess sustainability*.

The **actors** in these papers differ. While Makropoulos notes households and politicians, Penzenstadler notes various roles in the software engineering profession (e.g., quality engineer, architect, etc.). Yet, these can be generalised as *decision making individuals* who act under the constraints of country- or company-wide *legislations and regulations*.

d) Summary: The analysis illustrates that 'sustainability requirements' as a term surfaces in a range of domains, yet signifies distinct aspects, considerations and key concepts. In combination with the detailed analysis, this suggests that:

- Systems Engineering is focused on *measurements and indicators* of sustainability requirements in specific dimensions, without comprehensive coverage of the relevant dimensions.
- Software Engineering is at a stage where *awareness* is built and *consideration* is given to the *concerns and concepts* of sustainability. The discussion also acknowledges that sustainability debt is already present due to the way software is engineered.
- Requirements Engineering is ahead of other disciplines in that it is primarily focused on *systematically integrating* requirements for different *dimensions* and looking at ways to measure such requirements.

VI. Discussion

The various disciplines use the term 'sustainability requirement' often without defining and clearly contextualizing it. Instead, it is implicitly situated in their domain, focusing on particular aspects. For example, environmental engineering focuses on the capacity of the environment to endure, while software maintenance and evolution focus on the capacity of a (technical) software system to endure. This is not clearly distinguished from the effects that systems under development have on the capacity of their surroundings to endure. Only in RE does a clear articulation of this distinction arise. Time is conspicuously absent from the discourse. Even while some papers reference definitions that clearly articulate long-term outlooks ('capacity to endure') and needs (Brundlandt), none of the words associated with this future outlook appear collocated. In returning to our research questions, we structure the following discussion according to the key questions raised by Tainter [7].

Scope: what to sustain? – Few examples demonstrated an explicit attention to the object or scope that *sustainability* was meant to refer to; the term was introduced without clarifying *what was meant to be sustained*. This can lead to significant confusion between the capacity *of* a software system to endure; the capacity of the containing system, which often can be considered best from a socio-technical perspective, or sometimes a cyber-physical system; and what we could call the sustainability effects of these systems in their wider environment. Each research area covers certain dimensions, but without clarity on this scope, it remains unclear how these perspectives could converge. Since these perspectives will cause entirely different requirements, their relationship is crucial.

Time: For how long? – It was striking to note that no statistically significant collocates surfaced in the concordance analysis that pointed to the time dimension of sustainability. While it played an explicit role in two individual articles [21], [57], it is not sufficiently represented in the corpus to pass the threshold of relevance. The few articles that mention it highlight the need to better understand it. This omission highlights the need to identify how time should be articulated and considered in SE.

Stakeholders: For whom? – Each domain identifies particular actors and other stakeholders, with all of them trying to be inclusive. Systems engineering names 'all stakeholders' in context, IS and ICT talk about business and customers, and SE and RE talk mainly about software engineers and developers. However, none of them state how to specifically act upon that in the sense of how to relate to those stakeholders or how to make sure they are all included. Few consider the role of the organization or the wider system of key decision and policy makers as agents of change. As such, the awareness of the role of stakeholder analysis for sustainability has not yet led to actionable guidance for sustainability design practice.

Resources: At what costs? – The publications within the corpus acknowledge the economic costs in moving towards sustainable development and solutions. The main concern that arises in ICT, IS and Sys. Eng. is improved cost efficiency of the process of interest within a given business context. Furthermore, regulatory compliance of the IT systems within a business is discussed. However, there is no comprehensive perspective of sustainability over time that encompasses dimensions beyond monetary cost and benefits.

a) Limitations: We discuss four threats to validity: construct validity, internal validity, external validity, and reliability.

Construct Validity. The choice of search terms, which narrowed the search to those articles that make explicit mention of 'requirements', seems warranted given the broad use of that term. Exclusion criteria identified that some articles were only tangentially related, which were then excluded. Apart from the focus on sustainability as multi-dimensional and time-related, which arises from its definitions, the constructs emerged inductively through quantitative and qualitative analysis.

Internal Validity. The use of corpus linguistics is a common technique used to establish the context of term's use in a body of literature. Resources limited an in-depth reading of all articles, which caused us to focus on highly-cited papers. Since this choice impacts how representative the particular findings about these papers are for the overall body of literature, our analysis uses these articles not to substantiate the argument, but primarily to illustrate the use of terms across disciplines.

External Validity. Despite the significant increases in size, what has not changed in corpus-building is the need for corpora to constitute balanced, representative samples of the population from which they are drawn. Representativeness is key in order for the results generated from corpus analysis to be generalizable beyond the specific sample from which they are taken. The corpus data we select to explore a research question must be well matched to the research questions. To ensure this was the case, the sample was compiled by combining results from multiple literature sources to ensure broad coverage.

Reliability. To mitigate the threat that the outcome is affected by the interpretation of the researcher, we applied researcher triangulation to ensure correctness of the findings. Each article was coded independently by two researchers. Conflicts were discussed until consensus was reached. The detailed analysis of selected articles was conducted in pairs of two researchers, with the same procedure. Finally, a significant part of the corpus contains work of the authors. Concordance analysis helped to create a distance to the work that is analysed.

VII. Conclusions and Outlook

This article characterizes the emerging use of sustainability requirements in software-related disciplines to clarify its context, salient meanings, and opportunities for convergence. To understand how the identified research constructs the notion of sustainability requirements, the paper has pursued a corpus-assisted discourse analysis approach.

To answer our research question, *how does the current research construct the notion of sustainability requirements through published work?*, we identified the context in which the term arises, characterised differences across subfields of computing, and then focused on representative examples in the fields closest to software engineering.

The analysis showed that the term *sustainability requirement* may be considered as a *red herring* in the sense that it is constructed in a way that suggests it is different in the way from how we understand requirements in general.

The results also suggest that the term sustainability requirement is *not a new category*, but the sustainability dimensions help understand the level of abstraction of the source of the requirement and they allow us to think more broadly about the sustainment of the system and its wider impact.

The article thus provides the evidence to back previous claims that there must be a discussion to foster the shared understanding of the term 'sustainability requirement' in the field of computing.

The conclusions highlight the need for (1) establishing common ground , perhaps based on the Karlskrona manifesto [12], (2) actionable empirical work to provide tangible evidence of how sustainability requirements play out in specific contexts, (3) systematic approaches that enable clearly scoped sustainability cost-benefit analysis over time for concrete systems development projects.

Acknowledgements

Part of this work was supported by NSERC through RGPIN-2016-06640, FAPERJ 210.551/2015 and CNPq. 210.551/2015.

References

[1] F. P. Deek, J. A. M. McHugh, and O. M. Eljabiri, *Strategic Software Engineering: An Interdisciplinary Approach*. Auerbach Publications, 2005.
[2] I. Sommerville and P. Sawyer, *Requirements engineering: a good practice guide*. John Wiley & Sons, 1997.
[3] I. Sommerville, *Software engineering*, 7th ed. Addison Wesley, 2004.
[4] C. Becker *et al.*, "Requirements: The Key to Sustainability," *IEEE SW Sp. Issue: The Future of Software Engineering*, vol. 33, no. 1, pp. 56–65, Jan-Feb 2016.
[5] D. H. Meadows, J. Randers, and D. L. Meadows, *The limits to growth: the 30-year update*. Chelsea Green Publishing Company, 2004.
[6] "Oxford English Dictionary Online, 2nd edition," http://www.oed.com/, July 2003.
[7] J. A. Tainter, "Social complexity and sustainability," *Journal of Ecological Complexity*, no. 3, pp. 91–103, 2006.
[8] G. H. Brundtland and UN World Commission on Environment and Development, *Our common future*. Oxford University Press, 1987.
[9] J. L. Ramsey, "On not defining sustainability," *Journal of Agricultural and Environmental Ethics*, vol. 28, no. 6, pp. 1075–1087, 2015.
[10] R. Chitchyan *et al.*, "Sustainability design in requirements engineering: Theory and practice," in *Proc ICSE*, 2016.
[11] C. C. Venters *et al.*, "The blind men and the elephant: Towards an empirical evaluation framework for software sustainability," *Journal of Open Research Software*, vol. 2, no. 1, 2014.
[12] C. Becker *et al.*, "Sustainability Design and Software: The Karlskrona Manifesto," in *Proc. ICSE*, 2015.
[13] J. Cabot *et al.*, "Integrating sustainability in decision-making processes: A modelling strategy," in *31st ICSE*. IEEE, 2009, pp. 207–210.
[14] M. Mahaux, P. Heymans, and G. Saval, "Discovering sustainability requirements: an experience report," in *Intl. Working Conf. REFSQ*, 2011, pp. 19–33.

[15] K. Roher and D. Richardson, "Sustainability requirement patterns," in *Requirements Patterns (RePa), 2013 IEEE Third Intl Workshop on*, 2013, pp. 8–11.
[16] ——, "A proposed recommender system for eliciting software sustainability requirements," in *Workshop USER*, 2013, pp. 16–19.
[17] S. A. Kocak, G. I. Alptekin, and A. B. Bener, "Integrating environmental sustainability in software product quality," in *Workshop Re4SuSy*, vol. 1416. CEUR, 2015, pp. 17–24.
[18] B. Penzenstadler *et al.*, "Safety, security, now sustainability: The nonfunctional requirement for the 21st century," *IEEE Software*, vol. 31, no. 3, pp. 40–47, 2014.
[19] G. Rodriguez-Navas, L. Duboc, S. Betz, R. Chitchyan, B. Penzenstadler, and C. C. Venters, "Safety vs. sustainability design: Analogies, differences and potential synergies," in *Workshop on RE4SuSy*. CEUR-WS, August 2015.
[20] C. C. Venters *et al.*, "Software sustainability: The modern tower of babel," in *Workshop RE4SuSy*, 2014.
[21] C. Becker, "Sustainability and longevity: Two sides of the same quality?" in *Workshop RE4SuSy*, 2014.
[22] B. Knowles, L. Blair, and S. Walker, "Toward a sustainability lexicon and pattern language?" in *CHI'14 workshops*, 2014.
[23] H. Koziolek, "Sustainability evaluation of software architectures: a systematic review," in *ACM SIGSOFT Conf. QoSA and ISARCS*. ACM, 2011, pp. 3–12.
[24] D. S. Katz *et al.*, "Summary of WSSSPE1," *J of Open Research Software*, vol. 2, no. 1, July 2014.
[25] Z. Durdik *et al.*, "Sustainability guidelines for long-living software systems," in *ICSM*, Sept 2012, pp. 517–526.
[26] B. Boehm, "A view of 20th and 21st century software engineering," in *Proc ICSE*, vol. 2006. ACM, 2006, pp. 12–29.
[27] K. E. Wiegers, *Software requirements*, 2nd ed. Microsoft Press, 2004.
[28] G. Kotonya and I. Sommerville, *Requirements engineering: processes and techniques*. Wiley, 1998.
[29] S. Lauesen, *Software requirements: styles and techniques*. Addison-Wesley, 2002.
[30] R. Sharma and K. K. Biswas, "Functional requirements categorization grounded theory approach," in *Intl Conf on Evaluation of Novel Approaches to Software Engineering*, April 2015, pp. 301–307.
[31] "ISO/IEC 25010:2011: Systems and software engineering. Systems and software quality requirements and evaluation (SQuaRE)," 2011.
[32] M. Glinz, "Rethinking the notion of non-functional requirements," in *Third World Congress for Software Quality*, Sep. 2005, pp. 55–64.
[33] T. McEnery and A. Hardie, *Corpus linguistics: method, theory and practice*. Cambridge University Press, 2012.
[34] S. Baimu, S. Baimu, and L. Jiaping, "Urban climatology, energy conservation and thermal performance of lhasa city," in *Intl Conf on Energy and Environment Technology*, vol. 1, 2009, pp. 497–500.
[35] J. Li, J. p. Liu, and Y. Liu, "Research on building heating energy option of lhaza and development energy management tools," in *Intl Conf on Management and Service Science*, Sept 2009, pp. 1–4.
[36] G. Cerinšek and S. Dolinšek, "In a search for competent engineers... competence framework in the field of sustainable manufacturing," in *Global Engineering Education Conference*, 2011, pp. 865–870.
[37] O. Mason, "Parameters of collocation: the word in the centre of gravity," in *Corpora Galore: Analyses and Techniques in Describing English*, J. M. Kirk, Ed. Rodopi, 1999, pp. 267–280.
[38] B. Louw, "Irony in the text or insincerity in the writer? the diagnostic potential of semantic prosodies," in *Text and Technology: In honour of John Sinclair*. John Benjamins, 1993, p. 157.
[39] S. Elo and H. Kyngäs, "The qualitative content analysis process," *Journal of Advanced Nursing*, vol. 62, no. 1, pp. 107–115, 2008.
[40] P. Mayring, "Qualitative Content Analysis," in *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, vol. 1, no. 2, 2000.
[41] J. Saldaña, *The Coding Manual for Qualitative Researchers*. Sage, 2012, no. 14.
[42] "Literature analysis data set," http://tinyurl.com/h8dz4s3.
[43] G. Buyukozkan and C. Berkol, "Designing a sustainable supply chain using an integrated analytic network process and goal programming approach in quality function deployment," *Expert Systems With Applications*, vol. 38, no. 11, pp. 13 731–13 748, 2011.
[44] A. C. Brent *et al.*, "Application of the analytical hierarchy process to establish health care waste management systems that minimise infection risks in developing countries," *Europ Journ of Operational Research*, vol. 181, no. 1, pp. 403–424, 2007.
[45] S. Naumann *et al.*, "The greensoft model: A reference model for green and sustainable software and its engineering," *Sustainable Computing: Informatics and Systems*, pp. 294–304, 2011.
[46] S. S. Mahmoud and I. Ahmad, "A green model for sustainable software engineering," *Intl Journal of Software Engineering and Its Applications*, vol. 7, no. 4, pp. 55–74, 2013.
[47] C. K. Makropoulos *et al.*, "Decision support for sustainable option selection in integrated urban water management," *Environmental Modelling and Software*, vol. 23, no. 12, pp. 1448–1460, 2008.
[48] B. Penzenstadler and H. Femmer, "A generic model for sustainability with process- and product-specific instances," in *Workshop GIBSE*. ACM, 2013, pp. 3–8.
[49] P. Lutze *et al.*, "A systematic synthesis and design methodology to achieve process intensification in (bio) chemical processes," *Computers & Chemical Engineering*, vol. 36, pp. 189–207, 2012.
[50] P. Pinder, "Preparing information security for legal and regulatory compliance (sarbanes–oxley and basel ii)," *Information Security Technical Report*, vol. 11, no. 1, pp. 32–38, 2006.
[51] R. R. Harmon and N. Auseklis, "Sustainable it services: Assessing the impact of green computing practices," in *Intl Conf on Management of Engineering Technology*, 2009, pp. 1707–1717.
[52] E. Ancillotti, R. Bruno, and M. Conti, "The role of communication systems in smart grids: Architectures, technical solutions and research challenges," *Computer Communications*, vol. 36, no. 17–18, pp. 1665–1697, 2013.
[53] A. Radjiyev *et al.*, "Ergonomics and sustainable development in the past two decades (1992-2011): Research trends and how ergonomics can contribute to sustainable development," *Applied ergonomics*, vol. 46 Pt A, p. 67, 2015.
[54] I. Bolis, C. M. Brunoro, and L. I. Sznelwar, "Work for sustainability: Case studies of brazilian companies," *Applied ergonomics*, vol. 57, pp. 72–79, 2016.
[55] M. Z. Huber, L. M. Hilty, and M. Glinz, "Uncovering sustainability requirements: An exploratory case study in canteens," in *Intl Workshop on Requirements Engineering for Sustainable Systems*, vol. 1416. CEUR, 2015, pp. 35–44.
[56] G. Lami, F. Fabbrini, and M. Fusani, *Software Sustainability from a Process-Centric Perspective*. Springer, 2012, pp. 97–108.
[57] G. Mussbacher and D. Nuttall, "Goal modeling for sustainability: The case of time," *MoDRE*, pp. 7–16, 2014.