

Mining Metrics for Enhancing E-Commerce Systems User Experience

Antonia Stefani

Department of Management Science and Technology, University of Patras, Patras, Greece Email: astefani@upatras.gr

How to cite this paper: Stefani, A. (2022) Mining Metrics for Enhancing E-Commerce Systems User Experience. *Intelligent Information Management*, **14**, 25-51. https://doi.org/10.4236/iim.2022.141003

Received: December 20, 2021 Accepted: January 24, 2022 Published: January 27, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/

CC O Open Access

Abstract

The diversity of e-commerce Business to Consumer systems and the significant increase in their use during the COVID-19 pandemic as a one of the primary channels of retail commerce, has made all the most important the need to measuring their quality using practical methods. This paper presents a quality evaluation framework for web metrics that are B2C specific. The framework uses three dimensions based on end-user interaction categories, metrics internal specs and quality sub-characteristics as defined by ISO25010. Beginning from the existing large corpus of general-purpose web metrics, e-commerce specific metrics are chosen and categorized. Analysis results are subjected to a data mining analysis to provide association rules between the various dimensions of the framework. Finally, an ontology that corresponds to the framework is developed to answer to complicated questions related to metrics use and to facilitate the production of new, user defined meta-metrics.

Keywords

E-Commerce, Web Metrics, Quality Attributes, Data Mining, Association Rules, Evaluation Framework, Taxonomy, Ontology, ISO25010

1. Introduction

The growth that Business to Consumer (B2C) e-commerce systems have experienced in the past few years has triggered the research on the identification of the factors that determine end-user acceptance of such systems [1]. An e-commerce system is a software platform where buyers and sellers interact through web-based services. Accessing content on-line or remotely manage transactions is difficult for novice users which are most of the on-line population today [2]. E-commerce differs from other web applications in that a basic condition of their success is the total involvement of the end-user at almost every stage of the purchasing process [3]. This is not the case in most other web applications.

E-commerce systems are comprised of many components with several configuration parameters that optimize system performance [4]. These parameters include hardware components (routers, firewalls, digital switches, servers, and workstations); software (HTML editors, Java development environments, network user interfaces, browsers, groupware, middleware, and so forth); network elements; other transmission network services (the Internet and virtual private networks). E-commerce systems are heterogeneous, distributed and concurrent and as such, designing for quality is not an easy task. B2C software has several features that make traditional software quality metrics less effective in producing realistic quality measurements. To ensure the high quality of e-commerce systems, rigorous web engineering approaches are needed to help developers to address the complexities of these web applications, as well as to minimize the risk of development, deal with the possibility of change, and deliver applications quickly, based on end-users' requirements [5].

In this work, the work in [6] is extended to present an e-commerce system three-dimensional evaluation framework-based end-user interaction categories, metrics internal specs and quality characteristics as defined by ISO25010 [7]. End-user interaction methods (facets), map the selected metrics to identified B2C processes. Metric's specs (meta-metrics) evaluate the measurement process and the reliability of measurements results provided by the metrics. This is a view of quality from a technical point (e.g., the view of the developer). External quality characteristics provide an end user's point of view to e-commerce systems quality. By combining these views in one framework a combined, metricoriented view of the quality in a system is produced. The framework provides a guideline on what metrics should be used how they should be used and where, when assessing specific parts of an e-commerce system. To this end, association rules mining is used and an ontology acting as knowledge base for inference mechanisms, are presented.

Beginning from the corpus of existing general-purpose web metrics, the first step of our methodology for constructing the framework includes a survey of web metrics that can be specifically applied to e-commerce systems. The survey resulted in a categorisation and qualitative measurement of metrics. To the best of our knowledge, this survey resulted in the classification of the majority of web metrics, and it is unique in its B2C software orientation. This helped not only to gain a spherical view of the field but to identify gaps that need to be filled in. This classification is beneficial to researchers who may wish to carry out a meta-analysis. After the collection and initial categorization, the metrics were categorised using the framework which also includes a taxonomy which identifies internal metric characteristics. A data mining analysis provided a set of association rules between the various dimensions of the framework. The framework answers questions about what metrics are appropriate for evaluating which parts of the system and how they should be used. These are usually questions involving at most two dimensions of the framework. To provide answers to more complex questions involving combinations of dimensions, an ontology that corresponds to the framework was developed. The population of the ontology with the results of the categorisation analysis resulted in a e-commerce web metrics knowledge base. This knowledge base can be used to produce new, user defined meta-metrics, based on special attributes incorporated in the underlying ontology structure.

The contribution of this work is three-fold. Firstly, this research addresses the issue of web metrics in e-commerce systems quality evaluation process. The results should be of great interest to web designers, Information System staff and researchers. Secondly, by explaining the relationship among quality and e-commerce systems' components that influence e-commerce success, the current research aid researchers in further refinement of E-Systems success models in general. Last but not least, the current study provides a framework for applying existing metrics of information systems' success on the e-commerce environment.

This paper is structured as follows: Section 2 presents the theoretical background and the framework; Sections 3 and 4 present the categorization of metrics based on the three dimensions of the framework. Section 5 presents the analysis results, and Section 6 presents the e-commerce web metrics ontology. Finally, the paper concludes in Section 7.

2. State of the Art

2.1. Web Metrics

The literature provides a breadth of different categories of web metrics as an evaluation tool for web engineers. However, none of these metrics or classification systems is specifically targeted to B2C e-commerce. Relevant proposals include methodologies for web quality improvement [8] [9], estimation models [10] [11] [12], usability guidelines [13] and assessment methods [14] and metrics [15] [16].

At the early stages of the web's maturity, a wide range of metrics has been proposed for quantifying web quality attributes [17] [18]. Functional size metrics [19] help in the estimation and evaluation of the software process controlling application quality cost and schedules. Web cost estimation metrics and web size metrics provide a taxonomy of basic concepts of software measurement while there have been proposed classification frameworks for determining how the classified metrics can be applied in the improvement web information access and use [20] [21] [22] [23] [24].

Especially in e-commerce systems, the high quality of services is one way to keep users revisiting the web site; this can be assured when quality is definable and measurable. Different processes and metrics have been proposed to measure the quality of e-commerce systems. By measuring the performance of E-commerce system processes it is possible to implementation different business policies and tactics [25]. Web site design strategies and models propose different metrics to support e-commerce system success and assess the quality of e-commerce systems [26]. Based on this theoretical background, our first intention is to examine end user's quality perception of e-commerce systems based on existing web metrics.

The problem of identifying the factors that determine end-user perceived quality in software systems is not new [27]. This is not the case with other on-line software systems. Designing a successful B2C (Business to Consumer) system requires a bullet-proof underling business process workflow, or in other words fulfilment of specific functional requirements. The latter, and quality in general, is often underestimated especially at the first stages of the system design/development.

Quality is important and can be examined from two different perspectives: from the developer's and the end-user's point of view. The developer-centred perspective explains and predicts consumer's acceptance of e-commerce systems by examining the technical specifications of a system. These technical specifications include both technological infrastructure and services [28]. Developers may use web metrics to measure the quality of the services provided to the end user. End user, especially in B2C systems, sets the quality attributes that influence shopping decisions [29]. Undoubtedly, to ensure the production of highquality e-commerce systems, it is important to be able to assess the quality of B2C systems from the point of the user as well. Quality is by default linked with the end-user's perception of quality. So, the question arises: how can one evaluate B2C systems using metrics and define the extent to which they meet endusers' requirements? To this end, it is necessary to provide a framework for assessing B2C system quality, a framework which combines web metrics of different types based on a formal standard. There are several reasons for using web metrics for such a cause. A metric is measurements of some property of a piece of software or its specifications, a subjective factor since a value can be assigned to it. In this work we refer to metrics applied to an e-commerce system as seen from the end-user point of view; for example, number of colours used, or number of clicks needed to reach the description of a product. Since the interface of the application at hand is based on World Wide Web technology, these metrics are called "web metrics". But how can it be measured?

2.2. Measuring with Metrics

Web metrics are not subjective; they are generally easily understandable by both developers and users; and most importantly as demonstrated in this work, they can be mapped to quality characteristics and sub-characteristics of formal quality standards like ISO25010. They are a means to be as objective as possible in a subjective matter such as quality. Although the use of individual or even sets of metrics may not always give the correct image of an e-commerce system, their

use within a framework may yield better results [30]. Thus, using objective measures of software under a framework, a result that is considered to be reached subjectively, is achieved. This is the goal of this work. This is an area that has not been well covered. Few approaches partially cover these requirements [30] [31] [32]. In this context, some interesting research questions arise.

"How existing web metrics can be related with B2C e-commerce systems?" Online shopping behaviour can be presented as a function of the interaction between the users and the software system per se. Quality may be modelled using three complementary facets that, when put together, provide a complete description of the system. Based on these three facets a categorization of existing web metrics is produced. This is the first step of relating existing web metrics with end-user's shopping behaviour.

"Which web metrics can be used in a specific quality evaluation scenario?" Meta-metrics represent different aspects of the measurement procedure like automation, measurement issues and reliability of provided measures. Meta metrics introduce the facet of measurement process at the evaluation framework. The selection of the appropriate evaluation process on each evaluation case ensures the reliability of the evaluation results.

"How web metrics can be related with end user's perception of quality?" The use of the external quality characteristics of ISO25010 provides the baseline on which an e-commerce system may be built, considering end-users' requirements.

A quality framework is proposed that includes three aspects (three dimensions) of quality evaluation process: facets, meta-metrics and external quality characteristics. These aspects are vertically related by providing a 3D-representation of e-commerce systems quality (**Figure 1**). Each metric is represented in this multi-dimensional model.



Facets are user-system interaction activities. They denote which metrics should be used in which part of the system (the "were"). Metrics are action-depended,



meaning that there is usually a one-to-one mapping between them and an interaction activity. By using facets, metrics are clustered according to their connection with end-user actions. Thus, facets categorize metrics focused on end user actions in the e-commerce system. There are three facets: Presentation, Navigation and Purchasing. Navigation is the facet that describes the various mechanisms provided to the end user for accessing information and services of the e-commerce system via alternative routes. Presentation is the facet that describes how a product or service is presented to the user. Purchasing refers to the facilities provided for the commercial transaction per se.

Meta-metrics denote which metrics should be used for evaluating the e-commerce application based on specific performance characteristics of the metrics themselves. These characteristics are divided into five categories (the actual meta-metrics) which measure the accuracy, the automation ability, bias, ease of use and units of measurement. So, the meta-metrics categorization provides an evaluation of metrics. The goal of this evaluation is not to criticize the actual usefulness of the metrics (this is subjective) or to directly compare them but to aid the practitioner in selecting an appropriate set of metrics suitable for a particular case. Although many web metrics can be of some value during a specific evaluation process, many may not fit entirely into a specific evaluation method.

External quality characteristics are the link between metrics and Software Quality dimensions as they are formally perceived by the software engineering community. They denote end-user's perception of these web metrics by providing the "how": a quality mapping of metrics to quality. For the shake of formality, four specific external quality characteristics of ISO25010 were used: Functionality, Usability, Reliability and Efficiency. ISO25010 is a general-purpose topdown approach based on its hierarchical structure of quality characteristics and sub-characteristics. In ISO25010 the quality characteristic of the system causes or facilitates or supports the use of the software system. This top-down approach is often referred to as domain decomposition, which consists of the decomposition of the e-commerce systems into its functional areas and subsystems.

The very nature of the metrics, the nature of the artefact they measure, contains valuable information which is not captured in the three dimensions described above. For this purpose, a tree-like taxonomy was incorporated in the framework (**Figure 2**).

In this taxonomy, first level nodes correspond to metrics related to Content, Structure of the e-commerce application as well as Visualisation and Process related metrics. The taxonomy has two levels. Further decomposition is made in the 2nd level (leaf nodes). Depending on its nature, an assumption is made that a metric belongs only to one leaf of the tree. Some metrics are of mixed nature, but this one-to-one relationship was kept. There is no direct connection between these nodes and the Facets. Content related web metrics measure attributes related to text, hypertext, or multimedia (audio, video, animation) properties. Similarly, structure-related metrics measure attributes related to the structure of



Figure 2. The taxonomy of e-commerce web metrics.

either a web page or the entire web site. Visualization metrics concern the appearance and Process metrics measure process-specific attributes. Further decomposition of the taxonomy is possible, but this would reduce the flexibility of the model. In the following sections details concerning the categorisation of metrics in dimensions are analysed.

3. The Facet Taxonomy

The quality of web applications can be measured from two perspectives: quality perceived by the developers, and quality as experienced by the end-users. E-commerce systems provide a full range of attributes that compose the conformance of requirements, both stated and implied. Depending on the nature of the e-commerce system these quality attributes can be measured in different ways using the appropriate metrics. Metrics are better suited (they give better results, that is a better representation of the quality of the system) when used to evaluate specific components. Some metrics are universal in the sense that they can be applied effectively in all components. By clustering metrics, not only bias produced by "non-applicable" metrics is reduced but evaluators also save effort as well since unnecessary measurements are minimized. The term "non-applicable" does not actually mean non-applicability; for the shake of simplicity, those metrics that yield low level results when being applied to some facet are excluded. One could assign weights to the importance of one metric in each facet (in the case of universal metrics), although this would be quite subjective.

This is not an extensive description of all existing metrics or a full presentation/analysis of their use but rather a facilitation of their use. For facilitating the presentation, a 3-letters code for each metric (e.g., EMB—Emphasized Body Text) is used followed by the metric name, a brief definition and its references. The metrics presented hereinafter have been selected from well-known and recent works that have been proposed for online sites and could be applied to e-commerce systems as well. They are presented alphabetically according to their codes.

3.1. The Presentation Facet

The overall presentation of an e-commerce system is composed of hyper information which is measured based on the potential information of a web object. A web object can be either textual information, graphics, images, or a multimedia artefact. So, the Presentation facet contains metrics that measure the quality of content e.g., how the product is presented to the end user. Content metrics help developers to make content understandable and navigable. This includes not only making the language clear and simple, but also providing understandable mechanisms for navigating within and between pages. Providing navigation tools and orientation information in pages will maximise accessibility and usability. **Table 1** presents the most significant B2C metrics for presentation according to the bibliography (based on previous work mainly of [11] [32] [33] [34] [35]).

3.2. The Navigation Facet

The navigability of an e-commerce system is a critical factor for its success. Navigation is an important design element, allowing users to acquire more of the information they are seeking and making that information easier to find. In **Table 2**, B2C metrics for navigation according to the bibliography, are presented.

Navigation issues support e-commerce systems quality by considering the quality of components such as indexes, navigation bars, site maps and quick links. The availability of these components facilitates access of information and services and enables users to locate efficiently the information they need, while avoiding usability bottlenecks. Additionally, navigation concerns the facilities for accessing information and the connectivity of e-commerce system applications.

3.3. The Purchasing Facet

Purchasing refers to those specific features of the e-commerce system that strongly support its commercial character. The purchasing process includes the following basic steps: location of the product to buy (via catalogue or search engine services), purchase of the product (addition to the shopping cart, order process). A reference is made to the search features and to the features that support directly or indirectly the purchase process per se. Some of these features are Table 1. B2C metrics for the presentation facet.

Acronym	Presentation Metrics							
Actoliyin	Full Title	Description						
ALT	ALternative Text	Percentage of alternative text for the total number of images on a page						
AUF	AUdio Files	Number of unique audio files						
AVF	AVerage Font	Average Font Size						
BOC	BOdy Colour	Number of different body colours						
BOT	BOdy Text	Percentage of words that are body vs. display text (<i>i.e.</i> , headers)						
COC	COlor Count	Number of total colours employed						
DIC	DIfferent Image Count	Number of non-repeated images in the site						
EBT	Emphasised Body Text	Portion of body text that is emphasized (e.g., bold, capitalized)						
FOC	FOnt Count	Count Total fonts employed (<i>i.e.</i> , face + size + bold + italic)						
GRA	GRAphic	Percentage of page bytes that are used for graphics						
GRC	GRaphics Count	Total graphics on a page (not including graphics specified in scripts, applets and objects)						
IAT	Images Alternative Text	Number of images with alternative text						
IMR	IMage Redundancy	$100 \times (1-\text{DIC/IPP})$						
IPP	Images per Page	Number of images per page						
LAC	Length of Audio clips	Average length of audio clips in the site						
LVC	Length of Video Clips	Average length of video clips in the site						
MAF	MAximum Font	Maximum Font Size						
MEC	MEdia Count	Number of not reused media files						
MIF	MInimum Font	Minimum Font Size						
NOC	NOde Count	Number of html files in the site						
PAS	PAge Size	Total bytes of the page as well as elements graphics and style sheets						
REM	REused Media	Number of reused media files						
TCC	Text Cluster Count	Number of text areas highlighted with colour, bordered regions, rules or lists						
TDO	3D Objects	Number of files including 3D objects						
TEP	TExt Positioning	Number of changes in text position from left						
TMA	Total Media Allocation	Total space allocated for all the media files not reused						
TNA	Total Node Allocation	Total space allocated for all the html files						
TNC	Total Node Complexity	Average number of different types of media per page						
TPC	Total Page Complexity	Average number of different types of media per page						
TRM	Total Reused Media allocation	Total space allocated for all the reused media files						
WOC	WOrd Count:	Total number of words on a page						

also related to the Navigability of the system but they are categorized differently because of their great contribution to the purchasing process. **Table 3** presents the most significant metrics for the search process, according to the

Table 2. H	B2C metrics	for the	navigation	facet.
------------	-------------	---------	------------	--------

•	Navigation Metrics						
Acronym	Full Title	Description					
BLC	Broken Link Count	Number of broken links both internal and external to the site, not including dynamically generated pages and links					
BRL	BRoken Links	$BRL = 100 \times (INL + EBL)/LIC$					
COD	COnnectivity Density	COD = CON/NOC					
CON	CONnectivity	Number of internal links. Dynamically generated links are not included					
CYC	CYclomatic Complexity	(CON - NOC) + 2					
DBL	Different Broken Links	$DBL = 100 \times DBL/LIC$					
DBL	Different Broken Links	Percentage of Different Broken links					
EBL	External Broken Links	Number of external broken links in the site					
GLC	Graphic Link Count	Number of graphic links on a page					
HOS	HOrizontal Scrolling	Capability of Horizontal scrolling					
IBL	Internal Broken Links	Number of internal broken links in the site					
LIC	LInk Count	Number of total links on a page					
LII	LInk Image	Number of images used as a link					
NFV	Non Frame Version	Existence of none frame version					
NUC	Nodes Under Construction	Number of pages under construction					
ORP	ORphan Pages	Number of orphan pages					
PAC	PAge Count	Number of static pages					
PLC	Page Link Count	Count the number of links on a page					
QAP	Quick Access Page	The quotient between the Links Count and Page Count.					
TLC	Text Link Count	Count the number of text links on a page					
TOV	Text Only Version	Support for Text only version					
UPT	UPloading Time	Count the uploading of a page					
VES	VErtical Scrolling	Capability of vertical scrolling					

state of the art.

Search metrics measure the end-users easiness to locate the information needed inside the e-commerce system data corpus. If the end user cannot find any information, he/she will probably not use the system anymore. Search should be adjusted to any query that the end user poses and should only present results with high relevance per each search session. A search session represents a single attempt by an end user to find some specific piece of information. A session is defined as a group of search requests coming from a single IP address with no more than ten minutes break between them. **Table 3.** B2C metrics for the search process of the purchasing facet.

A	Purchasing Metrics						
Асгопуш	Full Title	Description					
AUT	AUThority	The reputation of the organization that produced the web page					
AVA	AVAilability	Number of broken links contained by the Web page					
СОН	COHesiveness	The degree to which the content of the page is focused on one topic					
CUR	CURrency	How recently a web page has been updated					
INR	Information-to-Noise Ratio	The proportion of useful information contained in a web page of a given size					
NST	New Search Terms	Number of new search terms on a session from a single IP					
РОР	POPularity	How many other web pages have cited this particular web page					
PRE	PREcision	The proportion of relevance between query and results					
REL	RELevance	The proportion of hits from each search engine that is followed by the end user					
RST	Repeated Search Terms	Number of repeated search terms on a session from a single IP					
SPS	Searches Per Session	The average duration of search on a session from a single IP					
WPQ	Words Per Query	The average number of words per query					

The end user navigates using alternative features that facilitate the purchasing process. The existence or not of these features defines binary metrics of e-commerce systems quality. These features support the interaction with the end user through the purchasing process. For example, features like indexes, FAQs and different language versions support end user's interaction by ensuring the reliability of the purchasing process. Additionally different web components (applets, agents) using the appropriate input data (*i.e.*, card number, name) help the end user to complete a purchase. In **Table 4**, the B2C metrics for interaction tasks are presented, according to the state of the art.

4. Meta-Categorisation

4.1. Meta-Metrics Definition

The framework uses five different meta-metrics that cover different aspects of the measurement procedure. The letters in parenthesis following the meta-metric name are used to facilitate and shorten future reference to the corresponding meta-metrics.

• Measurement scale (MS). The values assigned to a metric could be of various scales. Such scales are nominal, ordinal, interval, ratio and absolute. As expected, metrics on nominal or ordinal scale could not be used as easily as metrics on ratio or absolute scale.

A	Interaction Metrics—Purchasing Facet						
Acronym	Full Title	Description					
ALI	Alphabetical Index	Existence of alphabetical index					
DNM	Depth of Navigation Map	Maximum depth of navigation map					
FAQ	FAQ feature	Existence of FAQ features					
НОР	HOme Pages	Number of major entry points to the web applications					
IND	INput Data	Input data for each operation					
LAV	LAnguage Versions	Number of different languages					
NWC	Number of Web Components	Number of applets, agents					
OTI	OTher Indexes	Existence of other indexes					
OUD	OUtput Data	Output data for each operation					
SIM	SIte Map	Existence of navigation map					
SUI	SUbject Index	Existence of subject index					
WEP	WEb Pages	Number of web pages in an application					

Table 4. B2C metrics for the interaction features of the purchasing facet.

- Measurement's independence (MI). The ability of a metric to always offer the same result (measurement) for the same measured unit is important. Metrics that may have various interpretations for different users are not ideal for use.
- Automation (AU). The effort required to automate a metric varies. Automation refers to the ability to implement software that automatically assigns values to metrics. Since software quality is subjective, it is very difficult to measure some metrics this way; a human peer is necessary in this case. For example, the number of background colours in a page can be easily measured by software (by analysing the underlying code of the page) but the reputation of the organization that produced the web page (AUT, **Table 3**) can only be evaluated by a human expert.
- Simplicity (SI). This meta-metric examines how a metric is defined in relation to the simplicity of the metric's definition, how easily this definition can be understood and facilitate actions in the development plan.
- Accuracy (AC). This examines if the metric measures what is supposed to be measured and how the metric is related to the abstract software characteristics or factors to be measured.

The actual meta-metrics values are presented in conjunction with ISO25010 external quality characteristics in the next section.

4.2. Mapping to ISO25010

There are a lot of definitions for software quality. Quality is generally defined as "a set of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs". These broad definitions can be applied to B2C software as well for they are software also and highly user-interactive.

ISO25010 is a quality standard for software product evaluation which provides quality characteristics and guidelines for their use. This standard aims at defining a quality model for software and a set of guidelines for measuring the characteristics associated with it. The quality model proposed by the standard is subdivided into two parts: the quality model for internal and external quality characteristics and the quality model for quality in use. A quality characteristic is a property of the software product that enables the user to describe and appraise some quality aspect of a product. Internal quality characteristics provide developer's view of quality and external quality characteristics concern the end user's perception of quality. A characteristic can be further detailed into (or described by) multiple quality sub-characteristics. **Figure 3** presents the hierarchy of ISO25010 internal characteristics [7].

ISO25010 may be used as basis for e-commerce quality evaluation but further analysis and mapping of its characteristics is required. The main question posed is how the standard can be related to a set of existing and already used web metrics to assess the quality of B2C e-commerce systems. In this work, the following external quality characteristics of ISO25010 to evaluate e-commerce systems are used: Functionality, Usability, Efficiency and Reliability. These characteristics concern the end user's perception of quality; an extra dimension in B2C e-commerce systems. In the following paragraphs, a description of these quality characteristics and their definition of their e-commerce character is made.

Functionality (F) refers to a set of functions and specified properties that satisfy stated or implied needs. The goal of Functionality is to provide integrative and interactive functions to ensure end-user convenience. Especially in e-commerce systems functionality refers to all functions and services that





e-commerce provides to the end user at each one of the three facets of Presentation, Navigation and Purchasing.

Usability (U) is defined as a set of attributes that bear on the effort needed for the use of a product or service, based on the individual assessment of such use by a stated or implied set of users. Usability is an important quality characteristic as all functions of an e-commerce system are usually developed in a way that seeks to facilitate the end user by simplifying end user's actions; this fact can however negatively affect the system in certain cases. In e-commerce systems usability can be defined as the usefulness of the B2C functions during the interaction of the end user with the system.

Efficiency (E) is a complex concept that entails both conceptual challenges as well as implementation difficulties. Efficiency is defined as the capability of the system to provide appropriate performance, relative to the number of resources used, under stated conditions. It refers to a state where system functions are both usable and successful, *i.e.*, they achieve their aim, the reason for their existence. One of the main criteria of efficiency of an e-commerce system is the quality of sub characteristics relating to time and resource behaviour.

Reliability (R) is the quality characteristic that refers to a set of attributes that bear on the capability of software to maintain its performance level under stated conditions for a stated period. Reliability is comprised of three quality sub-characteristics: maturity, fault tolerance, and recoverability. Reliability refers to error-free and unconfused user experiences during navigation but also to support in bottleneck situations. Characteristics like 'Undo' functions and error recovery for broken links, data entry errors and orphan pages are the most popular methods for increasing reliability.

The mapping of metrics to facets and ISO25010 quality characteristics fuels the next step of this research: the mining of hidden relations between metrics or groups of metrics through association rules mining.

5. Data Mining for Association Rules

5.1. Parameterization

For examining the connection between web metrics and quality characteristics the "+" sign is used to denote metrics that can be used to provide measures for each quality characteristics.

For examining measurement scale (MS) two symbols are used, the "+" and "-". The "-" characterizes metrics that offer results on absolute, ration and interval scale, while "+" characterizes metrics on nominal and ordinal scale. According to the measurements' independence, (MI) the "+" sign is used to denote metrics that are always measured in the same way and "-" for metrics that their data collection may vary according to each case. For the automation (AU) easiness, the "+" is used to denote metrics that are automated easily, "=" for metrics that require significant effort to automate and the "-" sign, for metrics that

cannot be automated. For the value of simplicity (SI) three symbols are used: "+" for very well-defined metrics, "=" for fairly defined metrics and "-" for metrics that are difficult to be understood, interpreted, and related to external software characteristics. Finally, the symbols "+" and "-" are also used for accuracy (AC).

5.2. Results Presentation and Analysis

Evaluation results for navigation and especially for connectivity metrics are presented in **Table 5**. These metrics are well defined; they are measurement independent but the measurement of some of them are not easily automated. None

Matula	Meta metrics					Quality Characteristics			
Metric	MS	MI	AU	SI	AC	F	U	Е	R
BLC	_	+	=	+	+				+
BRL	_	+	+	+	+				+
COD	_	+	+	+	+				+
CON	_	+	+	+	+				+
CYC	-	+	+	+	+	+			+
DBL	_	+	=	+	+				+
DBL	_	+	+	+	+				+
EBL	_	+	+	+	+				+
GLC	_	+	+	+	+	+	+		
HOS	+	-	+	+	+		+		
IBL	+	+	+	=	+	+			+
LIC	_	+	+	+	+		+		
LII	_	+	+	+	+		+		
NFV	+	-	=	=	+	+		+	
NUC	-	+	+	+	+		+		+
ORP	_	-	=	=	+				+
PAC	-	+	+	+	+		+		
PLC	-	+	+	+	+	+	+		
QAP	+	-	=	=	+				+
TLC	_	+	+	+	+		+		
TOV	+	-	=	=	+	+	+		
UPT	+	-	=	=	+			+	+
VES	+	-	+	+	+		+		

Table 5. Evaluation results for the navigation metrics.

of the metrics (23 in all) is mapped to all four quality characteristics or at least to three of them. Most of the metrics are mapped to Reliability and Usability (56.5% and 43.4% respectively).

Table 6 presents the evaluation results for the Presentation facet. From the results a conclusion is drawn: most of these metrics can be automated and can

Metric	Meta metrics					Quality Characteristics			
	MS	MI	AU	SI	AC	F	U	Е	R
ALT	_	+	+	+	+	+	+		
AUF	-	+	+	+	-	+			+
AVF	-	+	+	+	+		+		
ВОС	_	+	+	+	+		+		
BOT	_	-	+	=	-	+	+		
COC	_	+	+	=	+		+		
DIC	-	+	+	+	+	+	+		
EBT	-	-	+	=	+		+		
FOC	-	-	+	=	+		+		
GRA	-	+	+	+	-	+		+	
GRC	-	+	+	=	+	+			
IAT	-	+	+	+	+	+	+		
IMR	-	+	+	+	+	+		+	
IPP	_	+	+	+	+		+		
LAC	-	+	+	+	+				+
LVC	-	+	+	+	+				+
MAF	-	+	+	+	+		+		
MEC	-	+	+	=	+				+
MIF	-	+	+	+	+		+		
NOC	-	+	+	+	+	+			
PAS	-	+	+	=	+				+
REM	-	+	+	=	+	+			+
TCC	+	+	+	=	-		+		
TDO	-	+	+	=	+	+			
TEP	+	+	+	=	+		+		
TMA	-	+	+	+	+				+
TNA	-	+	+	+	+			+	+
TNC	_	+	+	=	+			+	+
TPC	-	+	+	=	+			+	+
TRM	-	+	+	=	+			+	+
WOC	-	+	+	+	+		+		

Table 6. Evaluation results for the presentation metrics.

present accurate results of measurement. As expected, this facet's metrics are mapped to the Functionality and Usability characteristics of ISO25010. Some of them can also be used to evaluate the reliability of a system. Out of 31 metrics attributed to this facet, none of them is mapped to all four quality characteristics, 4 (12.9%) are mapped to three quality characteristics, 9 (29%) are mapped to two, leaving 18 (58.1%) mapped to only one characteristic.

Finally, **Table 7** presents the evaluation results for Interaction in two groups: (a) 12 web metrics for search features and 12 web metrics for navigation features. Most of these metrics are binary and cannot be easily automated so end user participation in the evaluation process is needed.

Motric	Meta metrics					Quality Characteristics			
Metric	MS	MI	AU	SI	AC	F	U	Е	R
AUT	+	-	-	=	-			+	
AVA	-	+	=	+	+				+
СОН	+	-	-	=	-		+		
CUR	+	-	+	+	+		+		
INR	+	-	-	-	-		+	+	
NST	+	+	=	=	+	+			
POP	-	-	+	+	+			+	
PRE	_	_	+	+	+			+	
REL	+	+	+	=	+			+	
RST	+	+	=	=	+	+	+		
SPS	+	+	=	=	+			+	+
WPQ	+	+	=	=	+		+	+	
ALI	+	-	=	=	+	+	+		
DNM	-	+	=	+	+		+		
FAQ	+	+	+	=	+		+		
HOP	+	-	=	=	+	+	+		
IND	+	-	=	=	-	+		+	
LAV	+	+	+	+	+	+	+		
NWC	-	+	+	+	+		+		
OTI	+	-	+	=	+	+			
OUD	+	-	+	=	+			+	
SIM	+	-	=	=	+	+			
SUI	+	-	=	=	+		+		
WEP	-	+	+	+	+	+		+	

Table 7. Evaluation results for the purchasing facet: searching and interaction metrics.

Out of 24 web metrics of the two groups, 9 (37.5%) are mapped to 2 external quality characteristics and 15 (62.5%) are mapped to one quality characteristic. Most metrics are mapped to Usability (12 metrics), Efficiency (11 metrics) and Functionality (9 metrics). This distribution denotes the difficulty underlying the purchasing process since its quality depends heavily on satisfying the rules of three characteristics with an almost equal distribution. Thus, developers of the functions of this specific facet should try to reach a quality equilibrium for these three characteristics. This is rather difficult since the satisfaction of one quality characteristic hampers the satisfaction of a sub-characteristic of another characteristic. For example, the inclusion of many functions serves Functionality (the system is more complete) but may hamper Usability (novice users are faced with an overcrowded user interface). This difficulty is also implied by the low automation values of these metrics.

Having in mind the metric categorisation into facets presented in **Tables 1-4** and the structure of the taxonomy (**Figure 2**), the two tables are combined in one, twhich maps the metrics into the leaves of 2nd level of the taxonomy and to Facets. The mapping is a one-to-one relation meaning that a metric belongs only to one leaf of the taxonomy tree of **Figure 2**. Some metrics have an ambiguous nature that is, it is difficult to decide the taxonomy leaf they belong to. For the shake of uniformity and simplicity, the one-to-one relationship is kept by assigning these metrics to the closest match possible. The result, **Table 8**, is another useful categorisation for selecting the most appropriate metrics for targeted evaluation.

5.3. Associations Rules

In order to find more relations between the metrics and the meta-metrics and/or quality characteristics, a data mining tool is used for discovering association rules that are not so obvious to find. The analysis used Weka [36] to analyse the metrics per facet and then the whole set. The data were modified for tool compatibility: for the meta-metrics, the "+" signs were replaced by "1", the "-" by "-1" and the "=" by 0. For the quality characteristics Boolean values were used: a "yes", if there exists a relation between a metric and a quality characteristic and a "no" otherwise. There were a lot of rules produced by the tool. In the following, only those who are useful and have a large confidence factor (they are valid for the majority (>70%) of metrics in the facet) are used. The rules are applied to the specific e-commerce related metrics presented in this paper and are not necessarily applicable to general purpose web metrics.

In the presentation Facet, two rules were discovered:

• Association Rule 1 (confidence factor: 100%):

 $MS = -1 \rightarrow AU = 1$ $MI = 1 \rightarrow AU = 1$ $AC = 1 \rightarrow AU = 1$

Terror		Metrics Facet							
Taxor	юшу								
1st level	2nd level	Presentation	Navigation	Purchasing					
Structure	Web site	NOC, TNO	BLC, BRL, COD, CON, CYC, DBL, EBL, IBL, QAP	НОР					
	Web page		GLC, LIC, PLC, TLC	AVA					
Content	Multimedia	AUF, DIC, GRA, GRC, IAT, IMR, IPP, LAC, LVC, MEC, RES, TDO, TMA, TNC, TPC	LII	NWC					
	Text	ALT, BOT, WOC		COH, INR, LAV					
	Hypertext		NUC, OPR, PAC	CUR, POP, WEP					
Visualization	-	AVF, BOC, COC, EBT, FOC, MAF, MIF, TCC, TEP	HOS, NFV, TOV, VES	ALI, DNM					
Processes	-		UPT	AUT, NST, PRE, REL, RST, SPS, WPQ, FAQ, IND, OTI, OUD, SIM, SUI					

Table 8. Mapping of metrics to the taxonomy of Figure 2 and to the three facets.

A rule that is somewhat self-evident: if a metric is accurate or has absolute/interval values or is always measured in the same way, then it is also easily automated. Most metrics in this facet are easily understandable so a connection between SI and AU is self-evident also.

- Association Rule 2 (confidence factor: 70%):
 - $U = yes \rightarrow R = no$
 - $R = no \rightarrow U = yes$

A metric mapped to U or R is not mapped to the other. This means that most metrics for this facet cannot be used to evaluate both Usability and Reliability characteristics.

In the Navigation facet one new rule was discovered and one was re-evaluated:

- Association Rule 3 (confidence factor: 80%):
 - $R = no \rightarrow U = yes$
 - $U = yes \rightarrow R = no$

This actually affirms Assoc. Rule 2 for the Navigation facet.

• Association Rule 4 (confidence factor: 88%):

 $E = no and R = no \rightarrow U = yes$

Metrics that are not mapped to E and R are mapped to U. This means that there are no metrics for measuring these three characteristics at the same time.

In the Purchasing Facet one rule was discovered:

- Association Rule 5 (confidence factor: 93%):
 - $E = no \rightarrow R = no$

Metrics not mapped to E are not mapped to R either. This means that in the

Purchasing facet there are no metrics that can be used to measure both Efficiency and Reliability. An Association Rule 4 is not that strong for this facet (it has confidence factor of 30%) was also discovered.

Finally, by putting all the metrics in one set, some rules with a global effect were discovered:

- Association Rule 6 (confidence factor: 97%):
 - $U = yes \rightarrow R = no$

Metrics mapped to U are not mapped to R. This means that most of the metrics that measure Usability do not measure Reliability as well.

- Association Rule 7 (confidence factor: 96%):
 - $R = yes \rightarrow U = no$

Metrics mapped to R are not mapped to U. This means that most of the metrics that measure Reliability don not measure Usability as well.

- Association Rule 8 (confidence factor: 95%):
 - $U = yes \rightarrow E = no$

Metrics mapped to U are not mapped to E. This means that most of the metrics that measure Usability don not measure Efficiency as well.

• Association Rule 9 (confidence factor: 100%):

 $F = no and E = no and R = no \rightarrow U = yes$

Metrics not measuring F, E and R are mapped to U. So there is no metrics that can measure all four quality characteristics.

6. An Ontology of E-Commerce Metrics

The tables of Section 5 can be used by a human peer or an automatic mechanism to answer simple questions involving few parameters. When encoded in a decision support mechanism the relations and data of these tables are hard to change, extended or shared. Most importantly, although data exist, it is not easy to answer more complex questions such as: "which metrics are appropriate for evaluating the efficiency and reliability of the purchasing process of an e-commerce site and are measurement independent?" or "which metrics can be used by an automatic procedure to evaluate the multimedia used in the navigation mechanism of an e-commerce site in terms of usability and effectiveness?". A different representation of the framework and the data involved is required, a representation that enables the reuse of domain knowledge and separates this knowledge from the operational knowledge (the decision support mechanisms). Such a representation is ontologically principled. By making use of the framework and the taxonomy, the classes, the sub-classes and the relationships of an e-commerce metrics ontology were built (Figure 4).

Classes and sub-classes are marked with a "C". The actual metrics are subclasses of the leaves (2nd level) of the taxonomy. Class or sub-class attributes include among others name, value, description, reference (citation) and special factors described in detail later. Sub-classes inherit all the attributes of a class. Besides the "is SubClass" relation there are three other relations that bind the



Figure 4. ISO25010 external and internal quality characteristics.

framework together: "is MeasuredBy" which is a many-to-many relation between a metric and the five meta-metrics of the framework, "isMappedto" which is also a many-to-many relation between a metric and the quality sub-characteristics of ISO25010 and "isUsedin" which is a one-to-one-relation between a metric and a facet.

By filling-in the values of the metrics described by the tables of Section 5, the ontology becomes a knowledge base. This ontology can be used by tools or humans (with the appropriate reasoning mechanisms) to suggest good combinations of metrics for targeted evaluation of e-commerce applications.

The framework and subsequently the ontology reason on how, where and which metrics should be used in different evaluation scenarios. The framework does not provide a firm ranking of metrics (e.g. "which are the best metrics?") A ranking of this type would be subjective; different users (*i.e.*, quality experts) would probably choose different metrics. Meta-metrics score, facet and quality characteristics mapping tell only one side of the story. Since a consensus on the significance (how good is a metric as an evaluation mean) of metrics presented is

subjective (*i.e.*, user dependent) the inclusion of a significance factor (SF) in the ontology was foreseen. This factor denotes how important is a metric and is set by the user, taking values ranging from [0,1]. The factor is set by default to 1 for all metrics in the ontology (*i.e.*, all metrics are equally important). This makes the ontology flexible by attacking the problem of subjectivity in the evaluation of the significance of metrics. One could also assign significance weights to facets or meta-metrics and derive a more parameterised version. Thus, different users may operate on different instances of the ontology, by increasing or decreasing the significance of metrics (or other classes/sub-classes) depending on their perception of quality. Using a customised decision mechanism, users can operate on their own version, at least until new research shades light on this subject.

Another important feature of the ontology is the possibility of defining meta metrics, metrics that combines two or more metrics to give a more compact view of quality. Ideally the proper combination of all metrics in one "super" meta-metric would give a clear indication of the quality of the system. Instead of having one metric to rule them all, simpler metrics, more realistic and unbiased can be constructed. Construction through combination is difficult and subjective. Which metrics should (and can) be combined and how? The ontology provides, along with the SF, one more tool for doing this, leaving the subjective issues again to the user: the metric normalisation factor (MNF). The MNF is used to convert the value of a metric (VM) to a value in the interval [0,1]. This factor is different for every metric since metric values use different units of measurement (from percent to sec or Boolean). The MNF is used to provide a unified measure for all metrics. The conversion of a value to the predefined interval is subjective and has to do primarily with the definition of a best- and worst-case value for this metric. For example, the GRC metric defines the number of graphics in a page. A user considers that a page in an e-commerce site should have at least 1 graphic (e.g. the product to be purchased) and at most 10 graphics (more would deem the page difficult to download). Based on this, a score of MNF = 1/10 is derived. So a page with 5 graphics would have a VM of 0.5. Values greater than 1 are again normalised to 1. This is a rather simplistic example, but it gives the general idea behind the use of this factor. MNF can either be set by the user or be defined by a survey with a rather large set of users. A meta metric MM can then be calculated by the following formula:

$$\mathbf{MM} = \sum_{i} \mathbf{MNF}_{i} * \mathbf{SF}_{i} * \mathbf{VM}_{i}$$

where *i* metrics (i > 1, selected by the user) are combined in a sum, with MNF_{*i*} being the metric normalisation factor of metric *i*, SF_{*i*} the significance factor and VM_{*i*} the corresponding value (**Figure 5**). The VM, SF and MNF factors are attributes of the Metric Class.

Depending on the nature of the metrics involved, the above-mentioned formula may include more factors that reduce the bias or give better results. In any case this procedure is defined by the user and realised by a mechanism that uses



Figure 5. A procedure for calculating a meta-metric.

the ontology.

The ontology is also extendable since new classes and relationships can be easily added, or the taxonomy can be easily rearranged. The association rules presented in Section 5 or other rules may be build-in to a decision mechanism to facilitate answers to complex questions. The rules were not used as relationships inside the ontology to retain a high level of flexibility.

The ontology was developed using the Protégé editor and is available in OWL (Web Ontology Language) making its use efficient by customised query engines or decision support mechanisms.

7. Discussion and Conclusions

Quality evaluation of B2C e-commerce systems can take a numerical form by using metrics. B2C systems, being web based may be evaluated in terms of quality by web metrics. However not all web metrics are suitable for such an evaluation. Starting from this point, the first goal of this research was to choose e-commerce-specific web metrics and categorize them according to both B2C-related and general attributes. The definition of these attributes was based on a literature review, the quality evaluation of several e-commerce systems and on development experience.

The resulting framework is based on three dimensions, each one contributing to goal of metric categorisation from a different perspective: either internal or external to the metric itself, user-oriented or evaluation expert-oriented. The measurement scale by using simple formalization contributes to the evaluation of e-commerce metrics by demonstrating that there might be two general views in quality evaluation, even for metrics: process perceived quality and user perceived quality. To conceptualize metric quality into three dimensions increases our ability to explain their relationship in a better way. In process perceived quality aspect the evaluator defines the resources (evaluation tools, human resources) of the evaluation process in order to select the appropriate metrics.

The results of our analysis using this framework is not a conclusion on how e-commerce systems can be measured qualitatively by metrics, but it rather provides an extendable tool useful for evaluation experts and developers alike. This is a step towards more effective measurements of e-commerce systems quality. The use of some of web metrics for e-commerce systems becomes more difficult because an e-commerce system is a general platform for several web applications.

Although the method proposed offers a well-defined evaluation framework, the evaluator plays an important role. The evaluator can use the default values of each quality characteristic but can also change the evaluation results to place emphasis on specific quality characteristics. Extreme modifications of the proposed evaluation results may significantly lead to meaningless results. The authors propose to an inexperienced evaluator to use the model as presented herein. Another limitation of the model is that the set of web-metrics that it defines may change over time as e-commerce technology is a rapidly growing area. This, however, does not affect the evaluation framework since an experienced evaluator can change or add web metrics and the values for the measurement scale(s) or easily expand/change the ontology.

This paper employed a quantitative research method to develop and validate a framework of e-commerce systems' quality; future qualitative studies on the topic will extend the reliability and validity of the findings of this study, possibly map metrics to quality sub-characteristics (ideally keeping the framework simple) or by simply adding new quality dimensions (in the condition that they keep the model tight and targeted on software quality).

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Campisi, T., Russo, A., Tesoriere, G., Bouhouras, E. and Basbas, S. (2021) COVID-19's Effects over E-Commerce: A Preliminary Statistical Assessment for Some European Countries. In: Gervasi, O., *et al.*, Eds., *Computational Science and Its Applications—ICCSA* 2021. *Lecture Notes in Computer Science*, Vol. 12954, Springer, Cham, 370-385. <u>https://doi.org/10.1007/978-3-030-86979-3_27</u>
- [2] Dionysiou, G., Fouskas, K. and Karamitros, D. (2021) The Impact of Covid-19 in E-Commerce. Effects on Consumer Purchase Behavior. In: Kavoura, A., Havlovic, S.J. and Totskaya, N., Eds., *Strategic Innovative Marketing and Tourism in the COVID*-19 *Era*, Springer, Cham, 199-210. https://doi.org/10.1007/978-3-030-66154-0_22

- [3] Suguna, M., Shah, B., Raj, S.K., et al. (2021) A Study on the Influential Factors of the Last Mile Delivery Projects during Covid-19 Era. Operations Management Research. <u>https://doi.org/10.1007/s12063-021-00214-y</u>
- [4] Winarsih, Indriastuti, M. and Fuad, K. (2021) Impact of Covid-19 on Digital Transformation and Sustainability in Small and Medium Enterprises (SMEs): A Conceptual Framework. In: Barolli, L., Poniszewska-Maranda, A. and Enokido, T., Eds., *Complex, Intelligent and Software Intensive Systems. CISIS* 2020. Advances in Intelligent Systems and Computing, Vol. 1194, Springer, Cham, 471-476. https://doi.org/10.1007/978-3-030-50454-0_48
- [5] Jílková, P. and Králová, P. (2021) Digital Consumer Behaviour and eCommerce Trends during the COVID-19 Crisis. *International Advances in Economic Research*, 27, 83-85. <u>https://doi.org/10.1007/s11294-021-09817-4</u>
- [6] Stefani, A. and Xenos, M. (2009) Meta-Metric Evaluation of E-Commerce-Related Metrics. *Electronic Notes in Theoretical Computer Science*, 233, 59-72. <u>https://doi.org/10.1016/j.entcs.2009.02.061</u>
- [7] International Standardization Organisation (2011) ISO 25010: Software Engineering—Software Product Quality Requirements and Evaluation (SQuaRE)—Measurement Reference Model and Guide. ISO, Switzerland.
- [8] Bures, M. (2015) Metrics for Automated Testability of Web Applications. Proceedings of the 16th International Conference on Computer Systems and Technologies, Dublin Ireland, 25-26 June 2015, 83-89. <u>https://doi.org/10.1145/2812428.2812458</u>
- [9] Martinez-Ortiz, A.L., Lizcano, D. and Ortega, M. (2019) Software Metrics Artifacts Making Web Quality Measurable. *Proceedings of the 14th International Workshop* on Automation of Software Test (AST'19), Montreal, 27 May 2019, 1-6. <u>https://doi.org/10.1109/AST.2019.000-2</u>
- [10] Mendes, E., Mosley, N. and Counsell, S. (2000) Comparison of Web Size Measures for Predicting Web Design and Authoring Effort. *IEEE Software*, **149**, 86-92. <u>https://doi.org/10.1049/ip-sen:20020337</u>
- [11] Mendes, E., Watson, I., Triggs, C., Mosley, N. and Counsell, S. (2002) A Comparison of Development Effort Estimation Techniques for Web Hypermedia Applications. 8th IEEE International Software Metrics Symposium (Metrics 2002), Ottawa, 4-7 June 2002, 131-140. https://doi.org/10.1109/METRIC.2002.1011332
- [12] Mendes, E., Counsell, S. and Mosley, N. (2003) Do Adaptation Rules Improve Web Cost Estimation? *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia*, Nottingham, 26-30 August 2003, 174-183. https://doi.org/10.1145/900051.900091
- [13] Nielsen, J. (2000) Designing Web Usability: The Practice of Simplicity. New Riders Publishing, San Francisco.
- Offut, J. (2002) Quality Attributes of Web Software Applications. *IEEE Software*, 19, 25-32. <u>https://doi.org/10.1109/52.991329</u>
- [15] Xie, X., Mao, J., Liu, Y., de Rijke, M., Chen, H., Zhang, M. and Ma, S. (2020) Preference-Based Evaluation Metrics for Web Image Search. *Proceedings of the* 43*rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, China, 25-30 June 2020, 369-378. https://doi.org/10.1145/3397271.3401146
- [16] Crichton, K., Christin, N. and Cranor, L.F. (2021) How Do Home Computer Users Browse the Web? ACM Transactions on the Web, 16, Article No. 3. <u>https://doi.org/10.1145/3473343</u>

- [17] Chen, Y., Zhou, K., Liu, Y., Zhang, M. and Ma, S. (2017) Meta-Evaluation of Online and Offline Web Search Evaluation Metrics. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Tokyo, 7-11 August 2017, 15-24. <u>https://doi.org/10.1145/3077136.3080804</u>
- [18] Orehovački, T. (2019) Objective and Subjective Metrics Meant for Evaluating Quality of Social Web Applications. *Proceedings of the* 2019 8th International Conference on Software and Information Engineering, Cairo, 9-12 April 2019, 25-29. <u>https://doi.org/10.1145/3328833.3328852</u>
- [19] Abrahão, S., Olsina, L. and Oscar, P. (2003) Towards the Quality Evaluation of Functional Aspects of Operative Web Applications. In: Olivé, A., Yoshikawa, M. and Yu, E.S.K., Eds., Advanced Conceptual Modeling Techniques. ER 2002. Lecture Notes in Computer Science, Vol. 2784, Springer, Berlin, Heidelberg, 325-338. https://doi.org/10.1007/978-3-540-45275-1_29
- [20] Tamm, Y.M., Damdinov, R. and Vasilev, A. (2021) Quality Metrics in Recommender Systems: Do We Calculate Metrics Consistently? 15th ACM Conference on Recommender Systems, Amsterdam, 27 September-1 October 2021, 708-713. https://doi.org/10.1145/3460231.3478848
- [21] Sloss, B.T., Nukala, S. and Rau, V. (2019) Metrics That Matter. *Communications of the ACM*, 62, 88. <u>https://doi.org/10.1145/3303874</u>
- [22] Song, S., Bu, J., Shen, J., Artmeier, A., Yu, Z. and Zhou, Q. (2018) Reliability Aware Web Accessibility Experience Metric. *Proceedings of the* 15th International Web for All Conference, Lyon, 23-25 April 2018, Article No. 24. <u>https://doi.org/10.1145/3192714.3192836</u>
- [23] Kirsh, I. and Joy, M. (2020) Splitting the Web Analytics Atom: From Page Metrics and KPIs to Sub-Page Metrics and KPIs. *Proceedings of the 10th International Conference on Web Intelligence, Mining and Semantics, Biarritz, 30 June 2020-3* July 2020, 33-43. <u>https://doi.org/10.1145/3405962.3405984</u>
- [24] Malhotra, R. and Sharma, A. (2015) A Web Metric Collection and Reporting System. Proceedings of the 3rd International Symposium on Women in Computing and Informatics, Kochi, 10-13 August 2015, 661-667. https://doi.org/10.1145/2791405.2791561
- [25] Beyer, B., Murphy, N.R., Rensin, D.K., Kawahara, K. and Thorne, S. (2018) The Site Reliability Workbook: Practical Ways to Implement SRE. O'Reilly Media, Sebastopol.
- [26] Stefani, A., Xenos, M. and Stavrinoudis, D. (2003) Modeling E-Commerce Systems' Quality with Belief Networks. *Proceedings of the IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems*, Lugano, 27-29 July 2003, 13-18.
- [27] Chen, L., Gillenson, M. and Sherrell, D. (2004) Consumer Acceptance of Virtual Stores: A Theoretical Model and Critical Success Factors for Virtual Stores. ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 35, 8-31. https://doi.org/10.1145/1007965.1007968
- [28] Fogli, D. and Guida, G. (2018) Evaluating Quality in Use of Corporate Web Sites: An Empirical Investigation. ACM Transactions on the Web, 12, Article No. 15. <u>https://doi.org/10.1145/3184646</u>
- [29] Diaz, E., Arenas, J.J., Moquillaza, A. and Paz, F. (2019) A Systematic Literature Review About Quantitative Metrics to Evaluate the Usability of E-Commerce Web Sites. In: Karwowski, W. and Ahram, T., Eds., *Intelligent Human Systems Integration* 2019. Advances in Intelligent Systems and Computing, Vol. 903, Springer,

Cham, 332-338. https://doi.org/10.1007/978-3-030-11051-2_51

- [30] Abbasi, M.Q., Weng, J., Wang, Y., Wang, I., Rafique, I., Wang, X. and Lew, P. (2012) Modeling and Evaluating User Interface Aesthetics: Employing ISO 25010 Quality Standard. *Proceedings of the 8th International Conference on Quality of Information and Communications Technology*, Lisbon, 3-6 September 2012, 303-306. https://doi.org/10.1109/QUATIC.2012.39
- [31] Fernandez, A., Insfran, E. and Abrahão, S. (2011) Usability Evaluation Methods for the Web: A Systematic Mapping Study. *Information and Software Technology*, 53, 789-817. <u>https://doi.org/10.1016/j.infsof.2011.02.007</u>
- [32] Lachner, F., Fincke, F. and Butz, A. (2017) UX Metrics: Deriving Country-Specific Usage Patterns of a Website Plug-In from Web Analytics. In: Bernhaupt, R., Dalvi, G., Joshi, A., et al., Eds., Human-Computer Interaction—INTERACT 2017. Lecture Notes in Computer Science, Vol. 10515, Springer, Cham, 142-159. https://doi.org/10.1007/978-3-319-67687-6_11
- [33] Ivory, M. and Megraw, R. (2005) Evolution of Web Site Design Patterns. ACM Transactions on Information Systems, 23, 463-497. <u>https://doi.org/10.1145/1095872.1095876</u>
- [34] Malak, G., Badri, L., Badri, M. and Sahraoui, H. (2004) Towards a Multidimensional Model for Web-Based Applications Quality Assessment. In: Bauknecht, K., Bichler, M. and Pröll, B., Eds., *E-Commerce and Web Technologies. EC-Web* 2004. *Lecture Notes in Computer Science*, Vol. 3182, Springer, Berlin, Heidelberg, 316-327. <u>https://doi.org/10.1007/978-3-540-30077-9_32</u>
- [35] Fetterly, M. Manasse, M., Najork, J. and Wiener, J. (2004) A Large-Scale Study of the Evolution of Web Pages. *Software. Practice and Experience*, 34, 213-221. <u>https://doi.org/10.1002/spe.577</u>
- [36] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H. (2009) The WEKA Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter*, 11, 10-18. https://doi.org/10.1145/1656274.1656278