Decision support system for blockchain (DLT) platform selection based on ITU Recommendations: A systematic literature review approach

Sylvain Kubler[a,*], Matthieu Renard[a], Sankalp Ghatpande[b], Jean-Philippe Georges[a], Yves Le Traon[b]

[a] *Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France*
[b] *Interdisciplinary Centre for Security, Reliability & Trust (SnT), University of Luxembourg, Luxembourg*

Abstract

Blockchain technologies, also known as Distributed Ledger Technologies (DLT), are increasingly being explored in many applications, especially in the presence of (potential) dis-/mis-/un-trust among organizations and individuals. Today, there exists a plethora of DLT platforms on the market, which makes it challenging for system designers to decide what platform they should adopt and implement. Although a few DLT comparison frameworks have been proposed in the literature, they often fail in covering all performance and functional aspects, adding that they too rarely build upon standardized criteria and recommendations. Given this state of affairs, the present paper considers a recent and exhaustive set of assessment criteria recommended by the ITU (International Telecommunication Union). Those criteria (about fifty) are nonetheless mostly defined in a textual form, which may pose interpretation problems during the implementation process. To avoid this, a systematic literature review regarding each ITU criterion is conducted with a twofold objective: (i) to understand to what extent a given criterion is considered/evaluated by the literature; (ii) to come up with 'formal' metric definition (i.e., on a mathematical or experimental ground) based, whenever possible, on the current literature. Following this formalization stage, a decision support tool called CREDO-DLT, which stands for *"multiCRiteria-basEd ranking Of Distributed Ledger Technology platforms"*, is developed using AHP and TOPSIS, which is publicly made available to help decision-maker to select the most suitable DLT platform alternative (i.e., that best suits their needs and requirements). A use case scenario in the context of energy communities is proposed to show the practicality of CREDO-DLT.

Keywords: Blockchain, Distributed Ledger Technology, Decision Support System, Multicriteria Decision Making, Analytic Hierarchy Process, TOPSIS

1. Introduction

Blockchain is expected to revolutionize computing in many sectors, particularly where centralization is undesired and trust is an issue. After the recognition gained by Blockchain 1.0 (cryptocurrency), Blockchain 2.0 (financial applications), there is a high demand for Blockchain 3.0 (industrial applications other than finance) (Daim et al., 2020; Ho et al., 2021; Budak and Çoban, 2021; Bhatt et al., 2021). Within this context, the number of startups pitching ideas continues to grow and distributed ledger models continue to evolve (Maesa and Mori, 2020; Pisa, 2018). This growth comes along with a wide range of blockchain/Distributed Ledger Technologies (DLT)[1] platforms, which may share common features and functionalities but also integrate specific ones. Recently, an empirical study conducted by Deloitte in 2021 about DLT projects in the GitHub open-source environment has revealed that more than 85.000 projects are available today, with around 9.000 new projects every year. Although only 8% of those projects are maintained in the long run, this nonetheless leads to a substantial number of platform alternatives, making it difficult for DLT practitioners to know what platform(s) they should adopt/select for their applications (Nanayakkara et al., 2021; Sharma et al., 2021).

Over the past few years, a couple of DLT comparison frameworks and decision support tools have

*Corresponding author
Email addresses: `s.kubler@univ-lorraine.fr` (Sylvain Kubler), `matthieurenard6@gmail.com` (Matthieu Renard), `sankalp.ghatpande@uni.lu` (Sankalp Ghatpande), `jean-philippe.georges@univ-lorraine.fr` (Jean-Philippe Georges), `yves.letraon@uni.lu` (Yves Le Traon)

[1] Blockchain and DLT is used interchangeably in this paper.

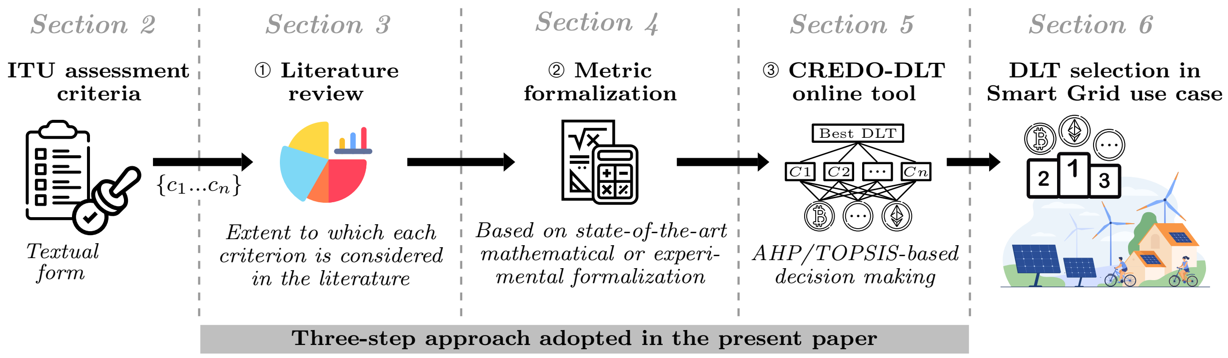


Figure 1: Three-step methodology applied in the present paper to move from the ITU recommendations to a practical decision support tool to guide blockchain practitionners in the selection process of a DLT platform

been proposed to help practitioners to deal with this decision-making problem (Ar et al., 2020). Among other frameworks, let us mention Nanayakkara et al. (2021); Labazova (2019); Gräbe et al. (2020) who introduced distinct approaches to evaluate a set of DLT platforms, although they do not build upon, or position themselves with reference to standardized criteria and recommendations, nor allows decision-makers to specify their requirements and/or preferences regarding one or more of the criteria. Six et al. (2020) have addressed such limitations by proposing and releasing an online decision support tool (called Blade[2]) that helps decision makers to select the most desirable DLT platform alternative. However, as will be further analyzed and discussed in section 2, this framework does not cover all performance and functional aspects of a DLT platform. Overall, there is a lack of normative, universally applicable decision-support framework for DLT platform assessment, which has a direct impact on the risk management of selecting and deploying a given DLT solution. This gap in research has been discussed, among other studies, in (Drljevic et al., 2020) and (Böckel et al., 2021). This motivates us to overcome this gap in research by investigating and proposing a decision support model/tool complying with a relevant DLT assessment criteria standard.

After having studied existing standards (further discussed in Section 2.2), we chose the recommendations of the ITU-T Focus Group on Application of Distributed Ledger Technology (FG DLT), which provide an exhaustive list of criteria (about fifty) covering most of the requirements for a DLT solution. These criteria are nonetheless defined in a textual form, which may pose interpretation problems when implementing them. Two research questions are thus formulated and addressed in this paper:

1. How assessment criteria textually defined can be turned into consensual formal mathematical definitions?

2. How requirements and preferences of DLT practitioners can be efficiently integrated into the decision-making process?

To address these two research questions, a three-step approach is adopted in this paper, as emphasized in Figure 1. First, a systematic literature review regarding each criterion is conducted with a twofold objective: (i) to understand the extent to which a given criterion is considered/evaluated by the literature (see stage denoted by ① in Figure 1); (ii) to come up with a 'formal' metric definition based on the current state-of-the-art (see stage denoted by ②). Following this formalization stage, a decision support tool called CREDO-DLT, standing for *"multiCRiteria-basEd ranking Of Distributed Ledger Technology platforms"*, is developed using Analytic Hierarchy Process (AHP) and TOPSIS (*cf.*, stage ③). The originality of this paper lies in the fact that this is the first study that proposes a decision-support tool built upon a "standardized" assessment criteria taxonomy for DLT platforms, whose metric formalization is obtained through a rigorous systemic literature review. CREDO-DLT, which is publicly available, is to the best of our knowledge the first decision support tool in that respect.

As emphasized in Figure 1, section 2 discusses past and ongoing standardization activities related to DLT/blockchain, and then details the assessment criteria defined by the ITU. Sections 3 to 5 respectively detail stages ① to ③ of Figure 1. The practicality of CREDO-DLT is showcased in section 6 considering a smart grid scenario; the discussion and conclusion sections follow.

[2] https://recommender.blade-blockchain.eu

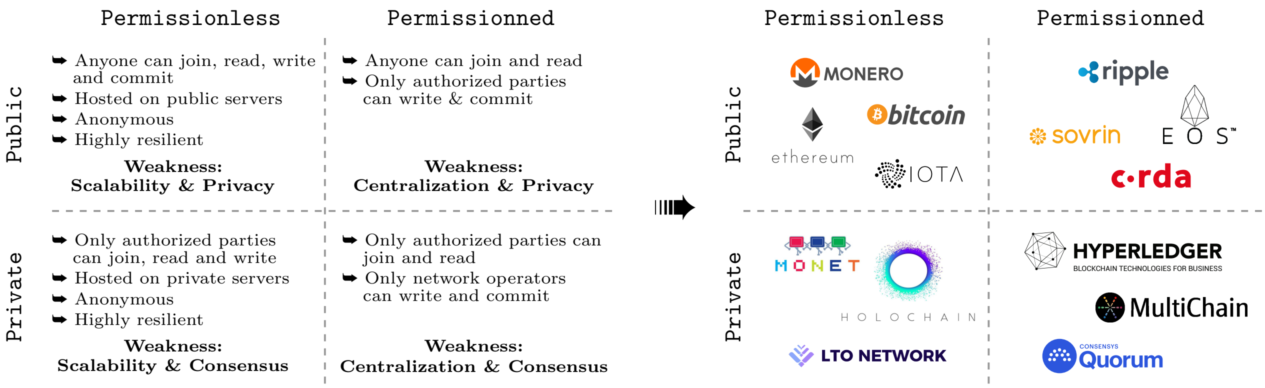


Figure 2: Blockchain (DLT) landscape

2. DLT: Definition & Standardization landscape

Section 2.1 briefly discusses the different types of DLT platforms that shape today's cryptocurrency world. Section 2.2 discusses past and ongoing standardization efforts for blockchain/DLT. Based on this discussion, the assessment criteria for DLT recommended by the ITU are introduced in section 2.3, along with an analysis of the extent to which existing DLT comparison frameworks cover those criteria/recommendations.

2.1. What does a DLT platform consist of?

A DLT platform is a complex system that builds upon different protocols, components and interacting subsystems. Abstractly speaking, a DLT platform can be seen as a four-layer model (Xie et al., 2019), namely:

1. *Network Layer:* P2P overlay network protocols play a crucial role at this layer, as they are used to efficiently handle distributed object storing, searching, and sharing among the blockchain network participants (Wang et al., 2019);

2. *Consensus Layer:* Consensus mechanisms are configured at this layer to decide how new blocks are added to the blockchain (i.e., to prevent frauds, duplicated entries, *etc.*), the most well-known mechanisms being Proof-of-Work (PoW), Proof-of-Stake (PoA), and Practical Byzantine Fault Tolerance (PBFT) (Bouraga, 2021; Xiao et al., 2020);

3. *Contract Layer:* some DLT platforms allow executing a set of logical instructions in the form of scripts (or "smart contracts");

4. *Social Layer:* it refers to economic and social considerations that must be addressed (e.g., how the DLT-related cryptocurrency can be efficiently integrated into the existing fiat economy).

Along with this four-layer model, decision-makers need to consider the class of blockchain (DLT) they would like to implement, namely (i) Public/Permissionless; (ii) Public/Permissioned; (iii) Private/Permissionless; or (iv) Private/Permissioned. Each class of blockchain serves specific use-cases and comes with its own advantages and disadvantages, as summarized in Figure 2. All this shows how complex it is for system designers/engineers to decide what class they should go for, and then, what specific platform they should select depending on their needs and requirements (Figure 2 giving a brief overview of what platforms could be selected in each of these classes). To lower the complexity in the selection process, but also ease and foster the adoption of DLT platforms in all sectors, several regulation and standardization initiatives currently co-exist, as reviewed in the next section.

2.2. Standardization initiatives

Standardization of blockchain technology is an essential step towards a common concept, interoperability, scaling, auditing and possible further technology regulations. Over the past few years, Several industry alliances and standards developing organizations (SDOs) have undertaken initiatives in that respect. For example, ISO (International Organization for Standardization) is working out the ISO/TC307 standard based on seven Working Groups (WG). CENELEC (European Committee for Electrotechnical Standardization) is collaborating with ISO to adapt the ISO standards to meet the European legislative requirements. The ITU has also created a Focus Group on Application of DLT (FG-DLT) to work out recommendations for the evaluation

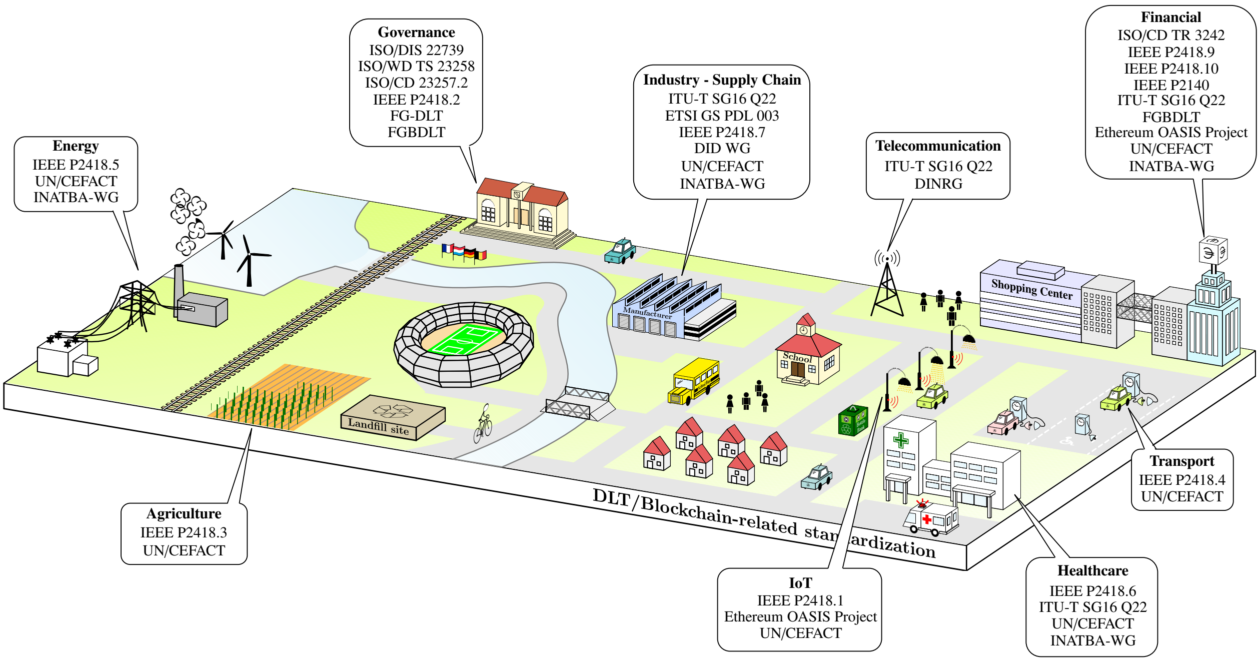


Figure 3: Overview of which standards covers what sector

Table 1: Overview of standard initiatives (whether under development or completed)

Standard	Issuing Organization	Accessibility	Latest Revision	Layer Focus (w.r.t the 4-layer model)			
				Network	Consensus	Contract	Social
ISO/CD TR 3242	International	Fee-based	Under development	✔	✔	✔	✔
IEEE P2418	International	Fee-based	Under development	✔	✔	✔	✔
Ethereum OASIS Open Project	International	Free	Under development	✔	✔	✔	✔
UN/CEFACT White Paper 2	International	Free	07/01/2019	✔	✔	✔	✔
UN/CEFACT Use Case paper	International	Free	03/28/2019	✔	✔	✔	✔
CEN/CENELEC-FGBDLT	European	Free	09/20/2018	✔	✔	✔	✔
ETSI GS PDL 003	European	Free	12/01/2020	✔	✔	✔	✔
ISO/DIS 22739	International	Fee-based	07/13/2020	✔	✔	✔	
ISO/WD TS 23258	International	Fee-based	Under development	✔	✔	✔	
ITU-T SG17 Q14	International	Free	Under development	✔	✔	✔	
ITU-T SG16 Q22 F.751.0	International	Free	08/13/2020	✔	✔	✔	
ITU-T SG16 Q22 F.751.1	International	Free	08/13/2020	✔	✔	✔	
ITU-T SG16 Q22 F.751.2	International	Free	08/13/2020	✔	✔	✔	
FG-DLT	International	Free	08/01/2019	✔	✔	✔	
IEEE P2140.1	International	Fee-based	11/04/2020	✔	✔	✔	
IEEE P2140.4	International	Fee-based	Under development	✔		✔	✔
ISO/CD 23257	International	Fee-based	Under development	✔		✔	
ISO/CD TR 23576	International	Fee-based	12/10/2020	✔			✔
IEEE P2140.2	International	Fee-based	Under development	✔			✔
IEEE P2140.3	International	Fee-based	Under development	✔			✔
IEEE P2140.5	International	Fee-based	07/17/2020	✔			✔
W3C-DID WG	International	Free	06/16/2021	✔			✔
IETF-DINRG	International	Free	Under development	✔	✔		
ISO/AWI TS 23259	International	Fee-based	Under development			✔	✔
ISO/CD TR 23244	International	Fee-based	05/07/2020	✔			
ETSI GS PDL 004	European	Free	02/01/2021			✔	
ISO/NP TS 23635	International	Fee-based	Under development				✔
INATBA-WG	International	Free	Under development				✔
				24	**16**	**19**	**16**

of DLT platforms. In 2019, this group was split into two sub-groups: (i) ITU-T Study Groups Q22/16 that focuses on standardization in application sectors; (ii) ITU-T SG Q14/17 that essentially focuses on security and privacy regulations. The IETF (Internet Engineering Task Force) also focuses on these aspects (security and privacy) through the Decentralized Internet Infrastructure Research Group (DINRG). The IEEE Standards Association, W3C (World Wide Web Consortium), ETSI (European Telecommunications Standards Institute), UN/CEFACT (United Nations Center for Trade Facilitation and Electronic Business), and INATBA (International Association of Trusted Blockchain Applications) have taken a slightly different track by developing standards for specific application sectors such as agriculture, healthcare, finance, supply chain, or still energy. An overview of what standard covers what sector is depicted in Figure 3. Furthermore, Table 1 provides a summary of the ongoing standard initiatives along the following criteria:

a) *Issuing Organization:* discerns between European and International SDOs;

b) *Accessibility:* whether the standard is accessible by the general public for free or if it needs to be purchased;

c) *Latest Revision:* date of the latest revision (indication about how up to date the standard is); some of them still being under development;

d) *Layer Focus:* emphasizes what layer(s) – *on the basis of the 4-layer model introduced in section 2.1* – the standard is covering/addressing;

It can be noted that half of the standards are accessible for free, and many are still under development. Second, it can be seen that only a few standards (7 out of 28) address the four layers described in section 2.1, although more than half address the network, consensus and (smart) contract layers. Although this landscape proves that standardization is an essential step for successful adoption of DLT solutions in all sectors, most of the standards published so far are merely informative rather than normative. This finding has been very well stressed and analyzed by König et al. (2020) in their recent article entitled *"Comparing Blockchain Standards and Recommendations"*. Along with the need to move towards more normative standards, there is also a need to design decision support tools based on standardized criteria. This what our paper is achieving considering the ITU-FG SG16 standard, which provides an exhaustive list of assessment criteria for DLT platform assessment. The next section briefly introduces those criteria, while discussing the extent to which existing DLT comparison frameworks/tools cover those criteria.

2.3. ITU assessment criteria & State-of-the-art DLT comparison frameworks

The ITU-FG SG16 (F.751.1: Assessment criteria for distributed ledger technologies) defines around fifty assessment criteria, which are split into five categories:

1. *Core Functions:* criteria for evaluating the extent to which a DLT platform fulfills the expected blockchain functionalities;

2. *Application functions:* criteria for evaluating the extent to which a DLT platform provides end-users with the possibility to efficiently interact with the DLT platform;

3. *Operation functions:* criteria for evaluating the extent to which a DLT platform allows for monitoring (and controlling) the status of nodes and of the underlying network;

4. *Ecosystem:* criteria for evaluating the extent to which a DLT platform is attractive (openness, community support…);

5. *Perfomance:* criteria for evaluating the intrinsic performance of the DLT platform (e.g., maximum throughput achievable, testing tool compatibility…).

Table 2 details the list of criteria defined by the ITU, and although they are textual defined, they form a good basis to start turning them into more formal metrics, which is the objective of section 3. But before doing so, let us review and analyze to what extent existing DLT comparison frameworks cover the ITU criteria.

A first evaluation framework was proposed by Labazova (2019) using a "Design Science Research" approach, which comprises six steps: problem identification, objective definition, design and development, demonstration, evaluation, and communication. In total, 21 criteria have been defined based on a systematic literature review, split into five categories (Blockchain Innovation, Blockchain Design, Inter-Organizational Integration, Implementation Environment, and Interconnections). A second evaluation framework was proposed by Polge et al. (2020), but it only focuses on permissioned blockchains and considers 5 criteria. Gräbe et al. (2020) went a step further in the decision-making process by proposing an approach that aggregates the different scores obtained by a set of platform alternatives

Table 2: Summary table of articles selected after filter two

Criterion				Definition
Core		c1	Account creation	Ability to create user accounts
		c2	Transaction processing	Ability to process – *(i) asset transfer and (ii) non-asset transfer* – transactions
	Query	c3.a	Balance query	Ability for end-users to acquire their account balance
		c3.b	Conditional query	Ability for end-users to search for historical information
	Consensus	c4.a	Data consistency	Ability to keep distributed ledger consistent and correct
		c4.b	BFT/CFT	Ability to continue working in presence of malicious actors or system failures
	Private key	c5.a	Software wallet	Ability to automatically generate private keys and include wallet backup mechanisms
		c5.b	Hardware wallet	Ability to store the generated keys in physical medium
	Smart contract	c6.a	Participants' status	Ability to monitor the status of nodes participating in the execution of a smart contract (e.g., computational resource used/consumed, *etc.*)
		c6.b	Lifecycle Management	Availability of lifecycle management functions (Create, Deploy, Update, Delete)
		c6.c	Reliability/Security	Support end-users in writing high-quality smart contracts (i.e., low bug rate)
		c6.d	Data Access control	Ability to support smart contract mechanisms (incl., code compilation/execution. . .)
	Crypto-graphy	c7.a	Encryption declaration	Ability to specify whether the encryption is derived from an open source solution or through regulatory compliance
		c7.b	Pluggable encryption	Ability to use a pluggable set of modular encryption algorithms
		c7.c	Encryption efficiency	Ability to provide strong encryption mechanisms with acceptable efficiency
		c7.d	Encryption strenght	Extent to which the DLT platform declares the security level of the used cryptographic schemas
		c8	Decentralization	Ability to provide operators with support for decentralized network management
Application	User authenti-cation	c9.a	Account verification	Ability to provide users with a password or two-step verification mechanism
		c9.b	Login management	Ability to update the status after user login
		c9.c	User classification	Ability to assign permissions to users and manage their access rights
		c9.d	Authorization	Ability to authorize third parties to access and modify their private data
	System stability	c10.a	Node management	Ability to grant operations when some nodes join, leave or upgrade
		c10.b	Cross-chain	Ability to grant normal operations when co-operating with other DLT/cloud systems
		c10.c	Network latency	Ability to remain stable despite network latency
		c10.d	Memory utilization	Ability to remain stable despite memory exceptions
		c10.e	CPU utilization	Ability to remain stable despite CPU exceptions
		c10.f	Concurrency	Ability to remain stable with bursts of concurrent transactions
	Economics design	c11.a	Incentive schemes	Ability to provide participants with incentives (financial or non-financial)
		c11.b	Token disclosure	Ability to provide end-users with a tokenomics report (fungible or non-fungible)
		c11.c	Token lifecycle	Ability to support (standardized) token issuance, transfer, withdrawal and clearance
	Information privacy	c12.a	Secure transmission	Ability to transfer information over a secure channel
		c12.b	Restricted data access	Ability to support access control and security protection mechanisms for confidential or personal data storage and exchange
		c12.c	Privacy protection	Ability to support/integrate privacy protection algorithm(s)
	Application functions	c13.a	UI for query	Should provide UIs to perform queries, visualize results, and show ledger's status
		c13.b	UI for smart contract	Should provide UIs to visualize the deployment and invocation of smart contracts
		c13.c	Multi-language SDK	Should provide at least one SDK
	Transaction origin	c14.a	Node	Ability to provide a mechanism to identify the origin of a transaction
		c14.b	Account	Ability to provide mechanisms to segregate the account signing a transaction from the account dispatching the transaction to the network
Operation	Network management	c15.a	Node status	Ability to monitor overall node status (e.g., number of online nodes, client version. . .)
		c15.b	Multi type nodes	Ability to classify nodes (full or lightweight)
		c15.c	Node configuration	Ability to support hot modification of node's setting parameters (e.g., block size)
		c15.d	Network fairness	Ability to support a mechanism to balance the sharing of network usage among nodes
	Risk management	c16.a	Recovery mechanisms	Ability to resist attacks (distributed denial of service, Sybil. . .) and recover efficiently
		c16.b	Trouble shooting	Ability to execute rapid trouble shooting and automatically send failure notifications
		c16.c	Single point of failure	Ability to be independent of any centralized system
Ecoystem		c18	Platform maturity	Extent to which a DLT platform is mature (year of creation, number of projects. . .)
		c19	Open source	Extent to which a DLT platform is open sourced and announces the licence used
		c20	Maintenance	Extent to which the platform is maintained (e.g., GitHub ticket management)
		c21	Professional support	Extent to which professionals can support the platform deployment and development
		c22	Running cost	Extent to which a platform is costly (e.g., audit fee, transaction fees. . .)
		c23	Avoid vendor lock-in	Ability to support standardized APIs (e.g., for service discovery, access. . .)
Perf.		c24	Transactions	Refer to the Throughput performance, i.e. how many transactions (Tx) per unit of time can be performed by the DLT platform
		c25	Testing tools	Extent to which the experimental settings (network topology, block size. . .) have been documented when publishing performance results

with regard to the considered criteria (21 in total), although the approach is quite simplistic (summing of the ratings). Nanayakkara et al. (2021) introduce a more advanced method using the Simple Multi Attribute Rating Technique (SMART) over 13 criteria. Although the approach is interesting, the authors have not released any tool/software that would allow end-users to use (benefit from) it. Such a framework has been proposed by Six et al. (2020) (tool called “BLADE”), which is to the best of our knowledge the only publicly available tool that has been designed based on a standard (ISO25010). One small criticism that could be made is that this standard has not been developed specifically for blockchain/DLT, but for software quality, and thus do not cover all the expected assessment criteria of DLT, as will be evidenced in the next paragraph. Having said that, the methodology underlying BLADE is well suited for capturing the end-user’s needs and requirements.

The five above-introduced DLT evaluation/comparison frameworks are further analyzed in Table 3, which highlights the extent to which those frameworks cover the ITU criteria. It can be observed that the frameworks of (Gräbe et al., 2020) and (Six et al., 2020) are the ones that best cover the ITU criteria, respectively with 24 and 21 criteria out of 53 (see last row of Table 3), against 12, 16 and 19 for the three others. As a second observation, some criteria/categories are well covered such as c4 (Consensus), c6 (Smart Contract), c10 (System stability), c11 (Economics design), or still the Ecosystem and Performance categories, while others are rarely considered/addressed (see e.g. c1, c2, c3, c5, c14, c15, c25). The objective of the present paper is to propose a decision-support tool that does cover all the ITU criteria, which is the subject of sections 3 to 5.

3. Literature review-based metric collection

This section is dedicated to the review of the literature with the objective to identify studies that cover one or more of the ITU criteria. This corresponds to step ① of the three-step approach presented in Figure 1. Section 3.1 presents the methodology applied to collect and select relevant state-of-art articles. Section 3.2 discusses the outcomes and findings of this analysis.

3.1. Research methodology

Scientific papers have been searched and collected from several scientific databases, including Elsevier, IEEE, Springer, Multidisciplinary Digital Publishing Institute, and arXiv. Based on common systematic literature review practices (Brereton et al., 2007; Kitchenham et al., 2009), a search has been carried out for

Table 3: How state-of-the-art DLT evaluation/comparison frameworks stand with respect to the ITU criteria/recommendations

	[N21]	[L19]	[P20]	[G20]	[S20]	[Y21]	[B21]
c1							
c2						■	
c3.a							
c3.b							
c4.a	■	■	■	■		■	■
c4.b	■	■	■	■	■	■	■
c5.a							
c5.b							
c6.a			■	■	■		
c6.b			■	■	■		
c6.c			■	■	■		
c6.d			■	■	■		■
c7.a					■		
c7.b					■		
c7.c				■	■		
c7.d				■	■		
c8				■		■	
c9.a					■		
c9.b					■		
c9.c	■			■	■	■	
c9.d					■		
c10.a		■	■	■		■	■
c10.b				■			■
c10.c		■		■			
c10.d		■		■	■		
c10.e		■		■	■		
c10.f				■			
c11.a			■		■	■	
c11.b			■	■	■	■	
c11.c		■	■		■		■
c12.a		■	■	■			
c12.b			■				
c12.c			■	■			
c13.a	■						
c13.b							
c13.c	■				■		
c14.a							
c14.b							
c15.a							
c15.b							
c15.c		■					
c15.d							
c16.a	■	■					
c16.b	■	■					
c16.c	■	■				■	
c18	■	■	■	■	■		■
c19		■	■				■
c20	■	■		■			
c21	■	■	■				■
c22	■			■			■
c23							
c24	■		■	■	■		
c25							
Total	**13**	**16**	**17**	**24**	**21**	**9**	**10**

[N21] Nanayakkara et al. (2021) [L19] Labazova (2019)
[P20] Polge et al. (2020) [G20] Gräbe et al. (2020)
[S20] Six et al. (2020) [Y21] Yang et al. (2021)
[B21] Büyüközkan and Tüfekçi (2021)

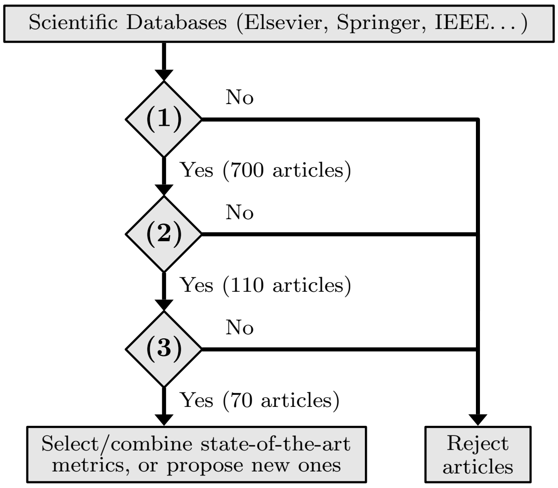


Figure 4: Methodology applied for the presented literature review

each ITU criterion (i.e., c1 to c25, *cf.* Table 2) using the following search terms: `"ci.j" + blockchain evaluation`, where `"ci.j"` refers to criterion *i.j* (e.g., `c4.a = "Data consistency"`). Based on those search queries, the paper selection process depicted in Figure 4 has been applied with the following filters:

- **(1)** *Has ci,j been addressed in one or more papers?* (some criteria might not have been addressed or discussed in the literature);
- **(2)** *Does the paper propose a "formal" definition of ci,j?* (some articles might discuss one or more criteria, but without formalizing how to measure/assess them);
- **(3)** *Is the (formal) definition aligned with the ITU definition?* (some articles might propose definitions that are different – *drastically in some cases* – from the ITU definition/recommendation).

As highlighted in Figure 4, more than 700 scientific articles were collected (i.e., when considering the 53 ITU criteria). Some queries resulted in thousand of scientific papers (e.g., c4.a: "Data consistency", c12.c: "Privacy protection"), while others were almost never covered, or even mentioned, such as c6.b: "Lifecycle management of smart contract", c7.b: "Pluggable encryption algorithm", or still c13.b: "User interface for smart contract". After applying filters (2) and (3) respectively 110 and 70 papers were identified (*cf.*, Figure 4). The next section discusses the results of these two steps in order to provide an indication of the extent to which a given criterion is discussed/considered in the literature, and the extent to which the state-of-the-art metric definitions are aligned with the ITU ones.

3.2. Literature review outcomes & findings

After applying filter (2) of the selection process, 110 papers were identified. Note that the median publication year of those papers is 2019, which shows that evaluation is becoming increasingly important for scholars working in the blockchain area. Figure 5 gives an overview of the proportion of papers (out of the 110) discuss/consider each of the ITU criteria. The two most considered ones are c4.a (Data consistency) and c4.b (BFT/CFT) with respectively 17 and 21 articles, which is not surprising as the consensus mechanism is the cornerstone of any DLT platform. On the other hand, seven criteria have not been discussed in the reviewed literature, among which c1 (Account creation), c13.b (UI for smart contract), c14.a/b (Transaction origin), c15.b/c (Network management), or still c25 (Testing tools). Figure 6 provides a more high-level view of how each ITU category is covered. It can be observed that Application and Core function categories are the two most covered ($\approx$ 40% each), which can be explained by the fact that those categories (i) cover the largest number of criteria; (ii) deal most with 'scientific' (performance) criteria.

After applying filter (3) (*cf.*, Figure 4), 70 articles were identified, which can be classified into three distinct metric classes:

- *Mathematical metric:* papers provide a mathematical definition (equation) of the metric;
- *Experimental metric:* papers provide an experimental methodology to quantify the metric;
- *Textual metric:* papers only provide a textual definition of what the metric is referring to.

Table 4 summarizes which article(s) – *among the* 70 – cover what criteria, and whether the proposed definitions are mathematically, experimentally or textually formulated. A complementary graphic representation is given in Figure 7, where it can be observed that criteria are often experimentally-evaluated rather than mathematically. This is not surprising considering the complexity of a DLT platform (from a design viewpoint), as was previously discussed in section 2.1. Another observation is that a significant proportion of scientific articles only provide textual definitions of criteria, which could lead to different (mis)interpretations when implementing them. This is one of key motivation underlying this research work. The next section further analyzes the 70 identified articles.

Table 4: Classification of the reviewed article(s) according to the definition class: Mathematical, Experimental, Textual

	Mathematical	Experimental	Textual
c1			
c2	(Park et al., 2019), (Fan et al., 2020)		
c3.a		(Han et al., 2020)	
c3.b		(Han et al., 2020)	
c4.a	(Gräbe et al., 2020), (Gopalan et al., 2020)	(Hao et al., 2018), (Gervais et al., 2016), (Baliga et al., 2018), (Gowat, 2020), (Srivastav et al., 2020), (Chaudhry and Yousaf, 2018), (Bhatt et al., 2021)	(Cong and Zi, 2020), (Qing et al., 2020),(Nanayakkara et al., 2021)
c4.b	(Gopalan et al., 2020), (Goffard, 2019), (Gräbe et al., 2020)	(Bhatt et al., 2021),(Dinh et al., 2018), (Gervais et al., 2016), (Zhang and Preneel, 2017), (Six et al., 2020)	(Smetanin et al., 2020), (Cong and Zi, 2020), (Maranhão et al., 2019), (Nanayakkara et al., 2021)
c5.a			(Eskandari et al., 2018), (Maranhão et al., 2019), (Mackay, 2019)
c5.b			(Mackay, 2019), (Suratkar et al., 2020)
c6.a			
c6.b			
c6.c		(Parizi et al., 2018), (Kirillov et al., 2019), (Prechtel et al., 2019), (Honig et al., 2019), (Kalra et al., 2018)	(Luu et al., 2016)
c6.d		(Kirillov et al., 2019), (Wöhrer and Zdun, 2018)	
c7.a		(Six et al., 2020)	
c7.b			
c7.c		(Mathew and Jacob, 2010), (Balasch et al., 2013), (Kobayashi et al., 2010)	
c7.d	(Gräbe et al., 2020), (Gupta and Shankarananda, 2015)		
c8	(Gräbe et al., 2020), (Polge et al., 2021)	(Bhatt et al., 2021)	
c9.a			
c9.b			
c9.c	(Gräbe et al., 2020)	(Six et al., 2020)	(Labazova, 2019), (Nanayakkara et al., 2021)
c9.d	(Gräbe et al., 2020)		
c10.a	(Gopalan et al., 2020)	(Dinh et al., 2018), (Dinh et al., 2018), (Dong et al., 2019), (Bhatt et al., 2021), (Belotti et al., 2019), (Polge et al., 2020)	(Labazova, 2019), (Maranhão et al., 2019)
c10.b			
c10.c	(Gräbe et al., 2020), (Tucker, 2004)	(Wan et al., 2019), (Zhong and Cole, 2018), (Miyamae et al., 2018)	(Maranhão et al., 2019)
c10.d	(Suankaewmanee et al., 2018), (Gräbe et al., 2020), (Tucker, 2004)	(Selimi et al., 2018), (Kabbinale et al., 2020), (Sagirlar et al., 2018), (Lohachab et al., 2021), (Six et al., 2020), (Dinh et al., 2018)	(Smetanin et al., 2020), (Maranhão et al., 2019)
c10.e	(Gräbe et al., 2020), (Tucker, 2004)	(Selimi et al., 2018), (Kabbinale et al., 2020), (Sagirlar et al., 2018), (Huang et al., 2019), (Dinh et al., 2018), (Six et al., 2020), (Lohachab et al., 2021)	(Smetanin et al., 2020), (Maranhão et al., 2019), (Qing et al., 2020), (Smetanin et al., 2020)
c10.f		(Zhong and Cole, 2018), (Saad et al., 2021), (Baqer et al., 2016)	
c11.a		(Bhatt et al., 2021)	(Maranhão et al., 2019), (Cong and Zi, 2020)
c11.b	(Gräbe et al., 2020)	(Polge et al., 2020), (Six et al., 2020), (Bhatt et al., 2021)	(Oliveira et al., 2018)
c11.c		(Borkowski et al., 2019), (Frauenthaler et al., 2020)	(Labazova, 2019)
c12.a	(Gräbe et al., 2020)		
c12.b		(Maranhão et al., 2019)	
c12.c	(Gräbe et al., 2020)	(Polge et al., 2020)	(Labazova, 2019)
c13.a		(Nanayakkara et al., 2021)	
c13.b			
c13.c		(Gowat, 2020), (Six et al., 2020)	(Nanayakkara et al., 2021)
c14.a			
c14.b			
c15.a			
c15.b			
c15.c			
c15.d	(Jain et al., 1984), (Gochhayat et al., 2020)		
c16.a		(Dinh et al., 2018)	(Nanayakkara et al., 2021)
c16.b			(Nanayakkara et al., 2021)
c16.c		(Modrak et al., 2014), (Bhatt et al., 2021)	

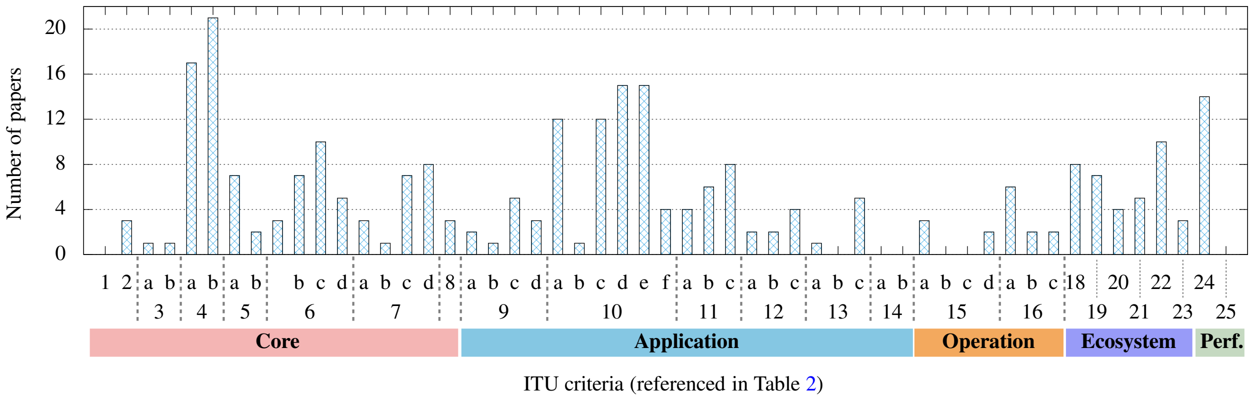


Figure 5: Overview of the extent to which the reviewed papers cover/address the IUT criteria

Table 4: Classification of the reviewed article(s) according to the definition class: Mathematical, Experimental, Textual

	Mathematical	Experimental	Textual
c18	(Gräbe et al., 2020), (Polge et al., 2020), (Tang et al., 2019)	(Shaw, 2018), (Gowat, 2020), (Six et al., 2020)	(Nanayakkara et al., 2021)
c19	(Polge et al., 2020), (Tang et al., 2019)	(Shaw, 2018)	
c20	(Gräbe et al., 2020)	(Shaw, 2018)	(Nanayakkara et al., 2021)
c21	(Gräbe et al., 2020)	(Shaw, 2018), (Six et al., 2020)	(Nanayakkara et al., 2021)
c22	(Gräbe et al., 2020)	(Kshetri, 2018), (Bai and Sarkis, 2020), (Szczerbowski, 2018), (Uesugi et al., 2020), (Bhatt et al., 2021), (Srivastav et al., 2020), (Delgado-Mohatar et al., 2019)	(Nanayakkara et al., 2021)
c23		(Nanayakkara et al., 2021)	(Tolk et al., 2007)
c24		(Baliga et al., 2018), (Dinh et al., 2018), (Dong et al., 2019), (Fan et al., 2020), (Kabbinale et al., 2020), (Li et al., 2017)	(Bhatt et al., 2021), (Gräbe et al., 2020), (StudyGroup16/22, 2020), (Maranhão et al., 2019), (Polge et al., 2020), (Qing et al., 2020), (Srivastav et al., 2020), (Tang et al., 2019)

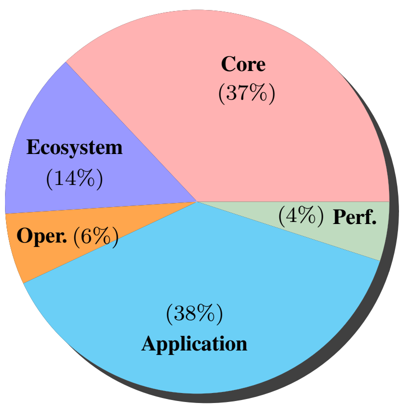


Figure 6: Percentage of articles per ITU (criteria) categories

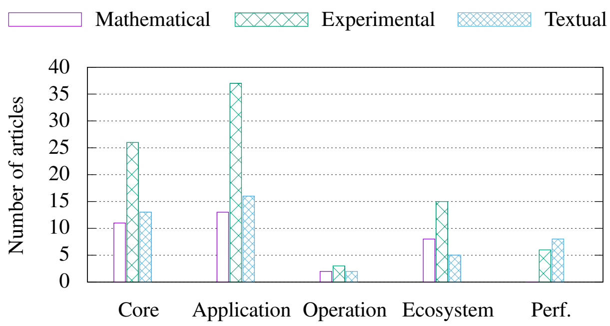


Figure 7: Article distribution per ITU category and metric definition class (Mathematical, Experimental, Textual)

4. Metric formalization

The 70 articles classified in Table 4 are further analyzed to identify or adapt, whenever possible, the definition/formalization proposed in those papers. Where this is not possible (e.g., because not aligned with the ITU definition), we propose our own metric definition. Table 5 shows the outcome of this analysis, which not only provides the mathematical formalization for each metric, but also the origin of that formalization (i.e., whether it comes from a scientific paper, directly from

Table 5: How state-of-the-art DLT evaluation/comparison frameworks stand with respect to the ITU criteria/recommendations

	Origin	Metric formalization	
c1	ITU-b	$m_1 = \begin{cases} 9 & \text{if account can be created manually AND automatically (e.g., Smart contract, API, ...)} \\ 5 & \text{if account can be created manually OR automatically} \\ 1 & \text{otherwise} \end{cases}$	
c2	ITU-b	$m_2 = \begin{cases} 9 & \text{if support asset transfer AND non-asset transfer} \\ 5 & \text{if support asset transfer OR non-asset transfer} \\ 1 & \text{otherwise} \end{cases}$	
c3.a	ITU-b	$m_{3a} = \begin{cases} 9 & \text{if user can get her/his account balance} \\ 1 & \text{otherwise} \end{cases}$	
c3.b	ITU-b	$m_{3b} = \begin{cases} 9 & \text{if user can search for historical information} \\ 1 & \text{otherwise} \end{cases}$	
c4.a	(Qing et al., 2020)	$x_{14a} = generation\_speed$	$generation\_speed$: block generation speed
c4.b	(Gopalan et al., 2020)	$x_{4b} = N/\sum T_{cons}(N_{\text{mal}})$	N: number of measurements; T_{cons}: time needed to append a Our block to all peers' chain; N_{mal}: number of malicious nodes
c5.a	(Eskandari et al., 2018)	$m_{5a} = \frac{nb\_swf}{nb\_wf}$	nb_wf: number of software wallet features based on the five-scale model defined in (Eskandari et al., 2018); nb_swf: number of software wallet features supported by the DLT platform
c5.b	(Suratkar et al., 2020)	$m_{5a} = (\frac{p}{max(p)} + (1 - \frac{c}{max(c)}))/2$	For a given technology, p : number or plateform support, c : the cost
c6.a	ITU-b	$m_{6a} = \begin{cases} 9 & \text{if can monitor status of participants} \\ 1 & \text{otherwise} \end{cases}$	
c6.b	ITU-b	$m_{5a} = \frac{nb\_sf}{nb\_f}$	nb_f: number of fuactivatnctions based on ITU's definition: {1-Create; 2-Deploy; 3-Activate; 4-Suspend; 5-Destroy}; nb_sf: number of fonctions supported by the DLT
c6.c	ITU-b	$m_{6c} = \begin{cases} 9 & \text{if a software/module to verify smart contract vulnerability and business logic is available} \\ 5 & \text{if a software/module for syntactic and/or semantic error detection and correction is available} \\ 1 & \text{otherwise} \end{cases}$	
c6.d	ITU-b	$m_{6c} = \begin{cases} 9 & \text{if possibility to add access control to a given contract} \\ 1 & \text{otherwise} \end{cases}$	
c7.a	ITU-b	$m_{7a} = \begin{cases} 9 & \text{if encryption declaration} \\ 1 & \text{otherwise} \end{cases}$	
c7.b	ITU-b	$m_{7b} = \begin{cases} 9 & \text{if a pluggable modular encryption is available AND can be switched online} \\ 5 & \text{if a pluggable modular encryption is available AND can only be switched offline} \\ 1 & \text{otherwise} \end{cases}$	
c7.c	(Kuznetsov et al., 2021)	$m_{7c} = nb\_hash$	nb_hash: Number of hashes per second (KHash/s)
c7.d	(Gupta and Shankarananda, 2015)	$m_{7d} = k(p,b) \approx \sqrt{2 * 2^b * ln\left(\frac{1}{1-p}\right)}$	k: number of tests to be performed to obtain (under p probability) the same output value for two distinct inputs (b: number of bits to hash)
c8	(Gräbe et al., 2020)	$m_8 = \frac{1}{N}\sum_{i=1}^{N} C_i$	C_i refers to the clustering coefficient of a node n_i computed as follows: $C_i = \frac{2.L_i}{k_i(k_i-1)}$, k_i referring to the degree of node n_i and L_i to the number of edges between the k_i neighbors of n_i
c9.a	ITU-b	$m_{9a} = \begin{cases} 9 & \text{if authentication is made based on a 2-step verification process} \\ 5 & \text{if authentication is made based on a password only} \\ 1 & \text{otherwise} \end{cases}$	
c9.b	ITU-b	$m_{9b} = \begin{cases} 9 & \text{if the platform updates the user login-state after login} \\ 1 & \text{otherwise} \end{cases}$	
c9.c	ITU-b	$m_{9c} = \begin{cases} 9 & \text{if end-users have access to an user interface (UI) to manage roles and access rights} \\ 1 & \text{otherwise} \end{cases}$	
c9.d	ITU-b	$m_{9d} = \begin{cases} 9 & \text{if end-users can grant authority to others to access or modify their private data} \\ 1 & \text{otherwise} \end{cases}$	
c10.a	(Gopalan et al., 2020)	$m_{10a} = \begin{cases} 9 & \text{if the platform has at most one block difference between all nodes of the network} \\ 1 & \text{otherwise} \end{cases}$	
c10.b	ITU-b	$m_{10b} = \lvert C \rvert$	C: set of DLT platforms with it is allowed to do cross-swap operations
c10.c	(Miyamae et al., 2018)	$m_{10c} = TPS\ (net\_lat)$	net_lat: Network latency

	Origin	Metric formalization	
c10.d	(Smetanin et al., 2020)	$m_{10d} = 1/memory$	$memory$: The amount of RAM required for efficient transaction/block processing (in Gigabytes).
c10.e	(Smetanin et al., 2020)	$m_{10d} = 100 - CPU$	CPU: Hardware utilized for blockchain-related data processing.
c10.f	ITU-b	$m_{10f} = \frac{1}{T_r} \mid T_r = c \times T\,(n-1)$	T_r: Time to return in normal mode; c: chain capacity (Leduc et al., 2021), $n \times c$ ($n > 1$): number of transactions sent during T sec
c11.a	(Bhatt et al., 2021)	$m_{11a} = \begin{cases} 9 & \text{if a rewarding mechanism exist (financial or non-financial)} \\ 1 & \text{otherwise} \end{cases}$	
c11.b	(Gräbe et al., 2020)	$m_{11b} = s\_tokens/tokens$	$tokens$: set of token alternatives, classified as {1-payment; 2-utility; 3-security}; s_tokens: number of supported tokens
c11.c	(Frauentha- ler et al., 2020)	$m_{11c} = \frac{1}{Tt}$	Tt: Time needed for token transfer
c12.a	(Gräbe et al., 2020)	$m_{12a} = \begin{cases} 9 & \text{if a secure transmission protocol is implemented} \\ 1 & \text{otherwise} \end{cases}$	
c12.b	ITU-b	$m_{12b} = \begin{cases} 9 & \text{if a differential access control mechanism is made available} \\ 1 & \text{otherwise} \end{cases}$	
c12.c	(Polge et al., 2020)	$m_{12c} = pr_{\text{mec}}$	pr_{mec}: number of privacy-preserving mechanisms available for use
c13.a	ITU-b	$m_{13a} = \begin{cases} 9 & \text{if M2M (Machine-to-Machine) AND M2H (Machine-to-Human) interfaces available} \\ 5 & \text{if M2M (Machine-to-Machine) OR M2H (Machine-to-Human) interfaces available} \\ 1 & \text{otherwise} \end{cases}$	
c13.b	ITU-b	$m_{13b} = \begin{cases} 9 & \text{if users can vizualize a smart contract (e.g., Remix-like interfaces)} \\ 1 & \text{otherwise} \end{cases}$	
c13.c	ITU-b	$m_{13c} = nb_{\text{SDK}}$	nb_{SDK}: number of SDK available for the DLT platform
c14.a	ITU-b	$m_{14a} = \begin{cases} 9 & \text{if mechanism allows for identifying the origin of a transaction} \\ 1 & \text{otherwise} \end{cases}$	
c14.b	ITU-b	$m_{14b} = \begin{cases} 9 & \text{if a mechanism allows for segregating the account signing a given transaction} \\ 1 & \text{otherwise} \end{cases}$	
c15.a	ITU-b	$m_{15a} = \begin{cases} 9 & \text{if M2M (Machine-to-Machine) AND M2H (Machine-to-Human) interfaces available} \\ 5 & \text{if M2M (Machine-to-Machine) OR M2H (Machine-to-Human) interfaces available} \\ 1 & \text{otherwise} \end{cases}$	
c15.b	ITU-b	$m_{15b} = \begin{cases} 9 & \text{if the DLT platform supports multi type nodes (full, lightweight...)} \\ 1 & \text{otherwise} \end{cases}$	
c15.c	ITU-b	$m_{15c} = \begin{cases} 9 & \text{if hot modification (i.e., online) is possible} \\ 5 & \text{if cold modification (i.e., offline) is possible} \\ 1 & \text{otherwise} \end{cases}$	
c15.d	(Gochhayat et al., 2020)	$m_{15d} = \left(\sum_{i=1}^{N} p_i\right)^2 / \left(N * \sum_{i=1}^{N} p_i^2\right)$	p_i: Total number of blocks mined by node i; N: number of miners
c16.a	(Dinh et al., 2018)	$m_{16a} = 1/T_{rec}$	T_{rec}: Recovery time needed to return to the normal operating mode
c16.b	(Sayadi et al., 2019)	$m_{16b} = (TP + TN)/(TP + TN + FP + FN)$	TP: True Positive Detection, TN: True Negative Detection, FP: False Positive Detection, FN: False Negative Detection
c16.c	(Modrak et al., 2014)	$m_{16c} = 1 - \frac{\sum\left(deg_{max}(v) - deg(v_i)\right)}{(n-1) * \sum\left(deg_{max}(v) - 1\right)}$	n: number of nodes, $deg\,(v)$: number connections of v node
c18	(Polge et al., 2020)	$m_{18} = \left(\frac{c}{max(c)} + \frac{u}{max(u)} + \frac{f}{max(f)} + \frac{t}{max(t)}\right)/4$	u: number of contributors to GitHub, c: number of commits, f: number of Twitter followers , t: number of tweets
c19	(Schmitz et al., 2018)	$m_{19} = s\_class/licence\_class$	s_class: number of supported licence classes; $licence\_class$: set of classes {1-reproduction; 2-modify; 3-distribute; 4-permissive; 5-freedom of linking; 6-covered by law; 7-OSI approved}
c20	ITU-b	$m_{20} = c\_iss/iss$	iss: Number of Github issues; c_iss: number of closed issues
c21	(Gräbe et al., 2020)	$m_{21} = a\_users$	a_users: number of active users (at least 1 commit over the last 3 months)
c22	(Gräbe et al., 2020)	$m_{22} = 1/\,(HE + TC * NT + MS)$	HE: average electricity/equipment cost; TC: average transaction cost , NT: number of transactions, MS: average maintenance cost
c23	(Kolbe et al., 2019)	$m_{23} = \begin{cases} 9 & \text{if rely on standardized semantically interoperable information models (e.g., schema.org...)} \\ 5 & \text{if rely on standardized syntactic formats (e.g., JSON, XML...)} \\ 1 & \text{otherwise} \end{cases}$	
c24	ITU	$m24 = n_{Tx}/t_{proc}$	n_{Tx}: number of transactions; t_{proc}: time needed to process them
c25	ITU-b	$m_{25} = \begin{cases} 9 & \text{if Testing tools for performance evaluation} \\ 1 & \text{otherwise} \end{cases}$	

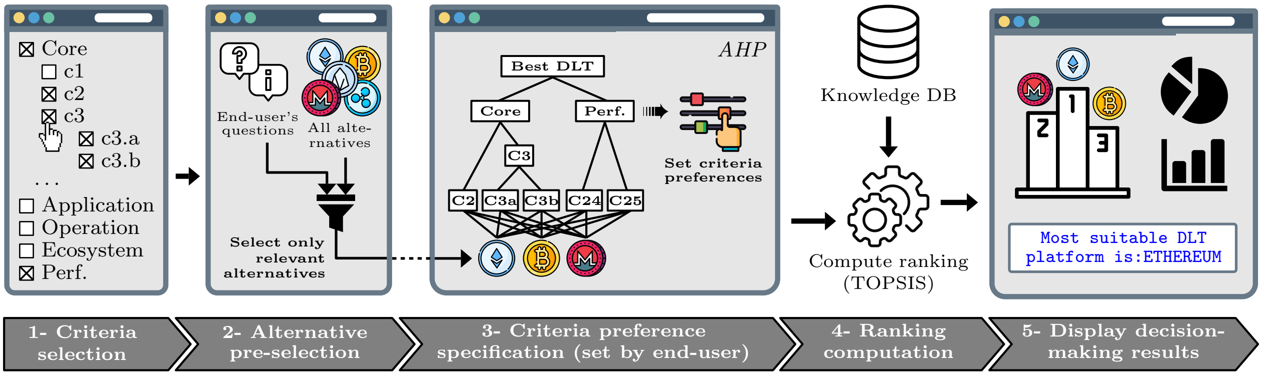


Figure 9: Illustration of the five-step approach underlying the CREDO-DLT decision support tool

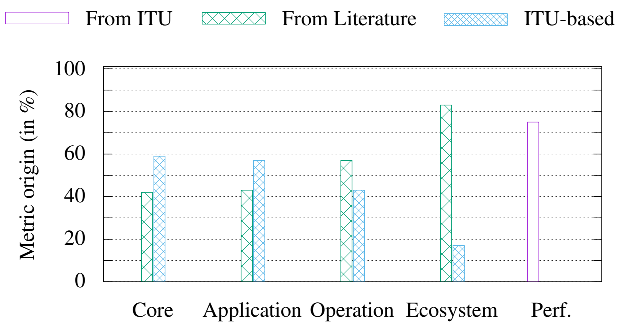


Figure 8: Overview of the origin of the metric formalizations (per category) that have been proposed in Table 3

the ITU report, or proposed based on the ITU definition). Figure 8 provides a complementary (graphical) view of the origin of the formalizations. It can be noted, in our study, that approximatively the same proportion of formalizations (around 50%) originates from the literature and is derived from the ITU for the Core, Application and Operation categories, while they have been directly extracted from the ITU regarding the Performance category.

As can be observed in Table 3, many criteria have been modeled in the form of system equations, which is mainly due to the fact that many ITU recommendations are condition-driven (i.e., *if-else*). Despite this, some of these system equations can be complemented, even extended, with other studies. For example, the system equation proposed for c3.a could eventually be extended to take into consideration the maximum number of queries that the DLT platform is able to handle (per unit of time), as proposed by Han et al. (2020). Regarding c6.c, several smart contract vulnerability detection software can be found in the literature, although they are usually platform-specific, as in (Parizi et al., 2018; Kirillov et al., 2019; Honig et al., 2019) whose solutions are only Ethereum's Solidity language-compliant. Good practices to write high-quality smart contracts could be followed, too, as the ones published by Kirillov et al. (2019) and Wöhrer and Zdun (2018).

5. CREDO-DLT: A multicriteria decision making tool for DLT platform selection

Based on the metric formalization detailed in the previous section, a decision support tool called CREDO-DLT (multiCRiteria-basEd ranking Of Distributed Ledger Technology platforms) has been developed and made publicly available online[3]. CREDO-DLT follows a five-step approach, as illustrated in Figure 9. Steps 1 and 2 are further presented in section 5.1, while steps 3-4-5 are detailed in section 5.2.

5.1. Steps 1-2: Criteria & Alternative (pre-)selection

CREDO-DLT provides decision-makers with the possibility to select only criteria they want to include in the decision-making process. Indeed, one might be interested in excluding one or more criteria, or even a whole category of criteria, as exemplified in Figure 9 (see step 1) where only the Core and Performance categories are being selected.

As a second step, CREDO-DLT provides decision-makers with a questionnaire in order to:

- guide them in selecting the right DLT platform class depending on their needs/constraints (i.e., should they adopt a Public/Permissionless, Public/Permissioned, Private/Permissionless, or Private/Permissioned DLT platform)? This allows

[3] http://www.credodlt.sylvainkubler.fr

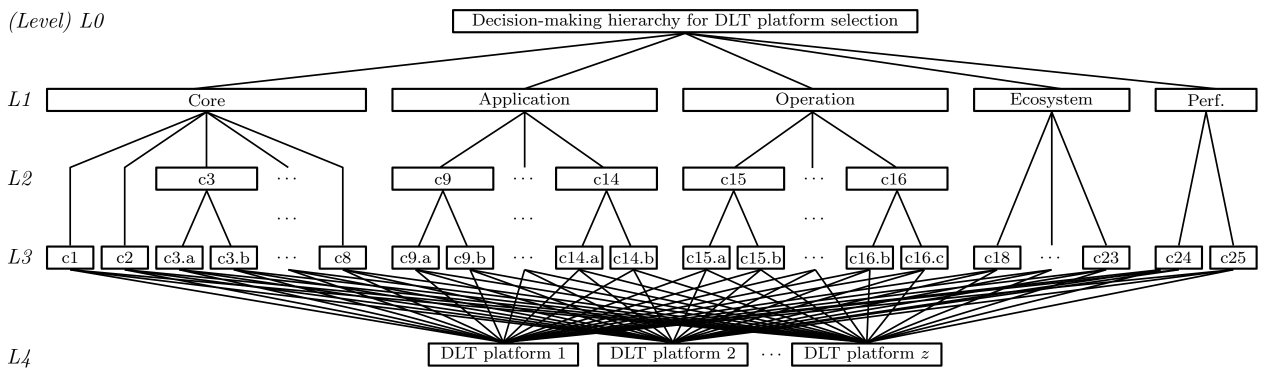


Figure 10: Percentage of articles per ITU (criteria) categories

CREDO-DLT for identifying the appropriate set of DLT platform alternatives to be considered and compared;

- take their specific needs into consideration (e.g., if a decision-maker requires a strong encryption mechanism). The set of questions to be asked to the decision-maker depends on the set of criteria selected in step 1.

5.2. Steps 3-4-5: Criteria preference specification, ranking computation & display

In step 3, CREDO-DLT uses AHP (Saaty, 1977) to structure the DLT platform selection decision-making problem, as depicted in Figure 10. Nonetheless, as discussed in the previous section, the final AHP tree will depend on the set of criteria selected by the decision-maker in step 1, as well as on the DLT platform class (alternatives) identified in step 2 (*cf.* AHP tree given in step 3 of Figure 9).

In compliance with the AHP method, CREDO-DLT provides decision-makers with the possibility to specify their preferences in terms of criteria importance (e.g., in Figure 9, the end-user specifies how much important the Core criteria category is compared to the Performance one). Preferences are specified using the 1- to 9-point Saaty's scale: $\{1, 3, 5, 7, 9\}$, as further detailed in Appendix A, which also details the complete mathematical formalization underlying AHP. Nonetheless, let us note that the pairwise comparisons performed at the alternative level relies on a "knowledge base" that has been established for that purpose. In other words, this knowledge base – *corresponding to "Knowledge DB in Figure 9* – allows CREDO-DLT to evaluate a given DLT platform alternative with regard to each criterion of level L3 in Figure 10. At the time of writing this article, seven DLT platform alternatives have been evaluated and included into the knowledge base, four public/permissionless platforms (bitcoin, Ethereum PoW, Ethereum PoZ, Monero) and three private/permissioned ones (Quorum, Hyperledger, MultiChain), with a long term objective of extending this base with other platforms such as IoTA, Corda, *etc.* To build this knowledge base, four different information sources have been used, namely:

- *Literature:* for some criteria, the performance scores of the alternatives with regard to those criteria have been collected based on state-of-the-art studies. Such cases have been highlighted using green cells in Table 6 (*cf.*, legend);

- *Source-code:* for some criteria, the performance scores have been assessed by analyzing the source code of the DLT platform (corresponding to cells highlighted in pink in Table 6);

- *Experiment:* for some criteria (see the ones highlighted in gray in Table 6), the performance scores must be evaluated under real-life conditions, as they depend on many applicative parameters and on the system's dynamic. One may state that this is in contradiction with the objective of the study, as CREDO-DLT aims at supporting decision-makers in the selection of a DLT platform before implementing it, but in fact blockchain simulators and experimental testbeds can be used to estimate such criteria beforehand. Among other simulators, let us mention BlockSim (Faria and Correia, 2019), Shadow (Miller and Jansen, 2015), Vibes (Stoykov et al., 2017), BlockPerf (Polge et al., 2021), or still Mininet (Kaur et al., 2014), Grid'5000 (Bolze et al., 2006) and Hyperledger Caliper (Sukhwani et al., 2018) for the experimental testbeds. Six et al. (2020), for example, used Grid'5000 in their study;

Table 6: Knowledge base used in CREDO-DLT to retrieve the score of a DLT platform alternative with regard to a given criterion, along with the type of information source used to obtain those scores (i.e., literature-based, source-code-based, experiment-based, API-based)

		Public/Permissionless				Private/Permissioned		
		bitcoin	Eth-PoW	Eth-PoA	Monero	Quorum	MultiChain	Hyperledger
Core	c1	5	9	9	5	1	1	1
	c2	5	9	9	5	5	5	9
	c3.a	9	9	9	9	9	9	9
	c3.b	9	9	9	9	9	9	9
	c4.a	exp	exp	exp	exp	exp	exp	exp
	c4.b	expF	expF	expF	expF	expF	expF	expF
	c5.a	0.5	0.33	0.33	0.33	0.66	0.33	0.33
	c5.b	1	1	1	1	1	1	1
	c6.a	1	9	9	1	9	9	9
	c6.b	1	5	5	1	5	5	6
	c6.c	1	9	9	1	9	5	9
	c6.d	1	9	9	1	9	9	9
	c7.a	9	9	9	9	9	1	1
	c7.b	1	1	1	1	1	1	5
	c7.c	exp	exp	exp	exp	exp	exp	exp
	c7.d	5.7E+38	5.7E+38	5.7E+38	5.7E+38	5.7E+38	5.7E+38	5.7E+38
	c8	exp	exp	exp	exp	exp	exp	exp
Application	c9.a	9	9	9	5	5	5	9
	c9.b	nf	nf	nf	nf	XX	nf	nf
	c9.c	1	1	1	1	9	9	9
	c9.d	1	1	1	1	1	1	9
	c10.a	expF	expF	expF	expF	expF	expF	expF
	c10.b	2	7	6	2	6	1	3
	c10.c	expF	expF	expF	expF	expF	expF	expF
	c10.d	exp	exp	exp	exp	exp	exp	exp
	c10.e	exp	exp	exp	exp	exp	exp	exp
	c10.f	expF	expF	expF	expF	expF	expF	expF
	c11.a	9	9	1	9	1	9	1
	c11.b	1	4	4	1	1	3	3
	c11.c	expF	expF	expF	expF	expF	expF	expF
	c12.a	1	1	1	9	9	9	9
	c12.b	1	9	9	1	9	9	9
	c12.c	1	1	1	4	5	5	6
	c13.a	9	9	9	9	9	9	9
	c13.b	1	9	9	1	9	9	9
	c13.c	4	2	2	1	3	1	3
	c14.a	1	1	1	1	1	1	9
	c14.b	1	1	9	1	9	9	9
Operation	c15.a	9	9	9	9	9	9	9
	c15.b	9	9	9	9	9	9	9
	c15.c	nf	nf	nf	nf	nf	nf	nf
	c15.d	exp	exp	exp	exp	exp	exp	exp
	c16.a	expF	expF	expF	expF	expF	expF	expF
	c16.b	expF	expF	expF	expF	expF	expF	expF
	c16.c	9	9	1	9	1	9	9
Ecosystem	c18	0.91	0.52	0.18	0.28	0.55	0.06	0.91
	c19	0.86	0.71	0.43	0.86	0.71	57	0.71
	c20	0.91	0.96	0.69	0.81	0.94	0.37	1
	c21	33	37	5	11	9	1	16
	c22	nf	nf	nf	nf	nf	nf	nf
	c23	5	9	5	5	5	5	5
Perf.	c24	exp	exp	exp	exp	exp	exp	exp
	c25	9	9	9	1	9	1	9

Literature-based | Source-code-based | experiment-based | API-based

nf: Non Found

exp: Scores obtained via experiments that are further conducted in Section 6

expF: Scores requiring experiments that are not included in this study/paper

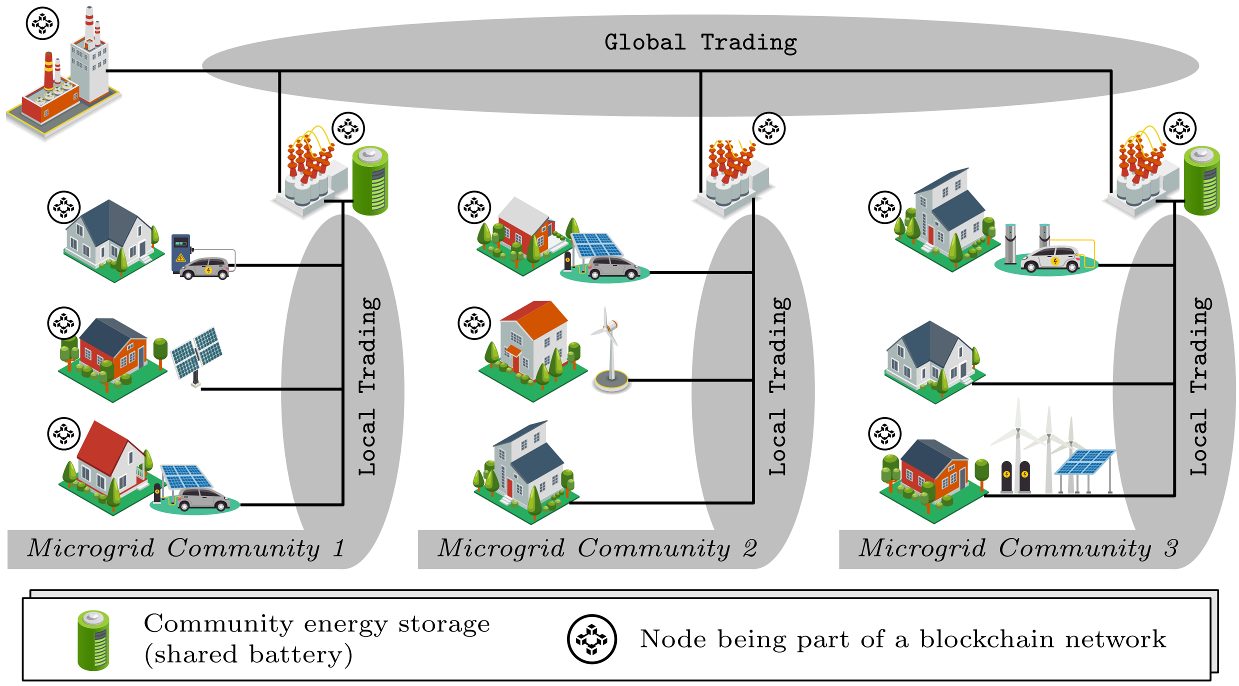


Figure 11: Scenario of decentralized energy communities making use of blockchain (DLT) technology to support automatic and peer-to-peer energy trading at both local and global levels (i.e., intra and inter-communities)

- *API:* for some criteria, the performance scores are computed based on information obtained via REST API calls, and can thus be updated on a regular basis (see cells highlighted in purple in Table 6);

Overall, Table 6 shows that it is not a straightforward process to evaluate and compare DLT platform alternatives, as some criteria scores are dependent on the targeted application (e.g., number of machines composing the blockchain network, communication capabilities, *etc.*), implementation details (source-code-related), or still the developer community (e.g., GitHub community). Furthermore, each criterion may be measured in different units (e.g., number of Tx/sec, using the Saaty's scale…), and, as a consequence, they all have to be normalized to obtain dimensionless classifications (i.e., a common numeric range/scale) to allow aggregation into a final score. To this end, a combination of max-normalization with linear-sum seems the most appropriate for AHP according to Vafaei et al. (2016). The mathematical formalization of this combination is detailed in Appendix A. Once normalized, all weights must be aggregated to obtain the final alternative score, based on which the DLT platform ranking is generated. To this end, the TOPSIS method combined with AHP, which has been proven a valuable/robust hybrid solution over the years (Chu et al., 2007; Kubler et al., 2016; James et al., 2021), is used, whose mathematical formalization is detailed in Appendix A.3. The final ranking is displayed to the decision-maker, as emphasized with step 5 in Figure 9.

6. DLT platform selection using CREDO-DLT: A decentralized energy communities' scenario

To showcase the practicality of CREDO-DLT, a use case scenario inspired by Wang et al. (2021) in the energy sector is presented, and more specifically about the use of blockchain to support decentralized energy communities and markets. An energy community refers to a group of interacting people who share both geographical location and energy needs (de Sõ José et al., 2021), as illustrated in Figure 11 (3 microgrid communities being considered in this scenario). Blockchain technology is seen as an enabler to support automatic and distributed energy exchanges at both the local level (energy traded among participants of a same community) and global one (energy traded between communities). Trading energy locally or globally brings different needs, whether from a security perspective (e.g., in terms of transaction encryption), a network and risk management, or still from an ecosystem perspective (i.e., maintenance, running cost…). This is where CREDO-DLT

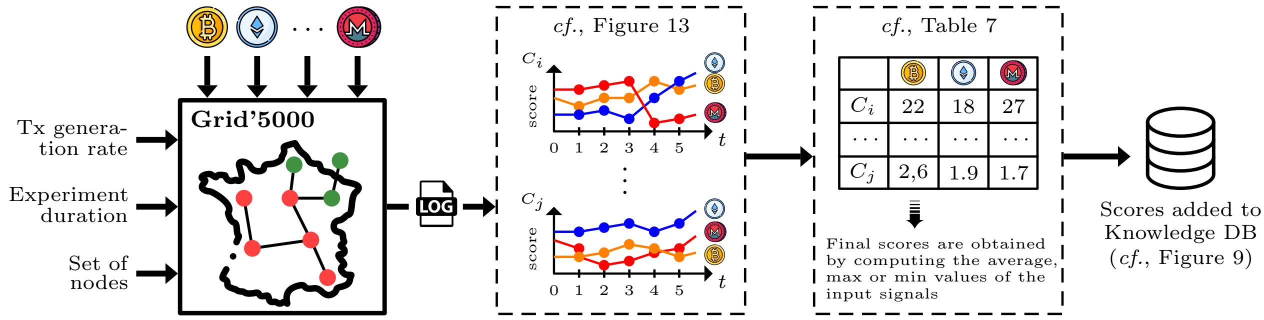


Figure 12: Experimental methodology applied to collect and compute the scores of the DLT platform alternatives w.r.t. ITU criteria (*cf.*, criteria denoted by "Exp" and "Exp(F)" in Table 6)

comes into play.

Section 6.1 presents the experimental approach defined and applied in this study to quantify the experiment-based scores expected by the knowledge base (*cf.*, criteria denoted by "Exp" in Table 6). Section 6.2 details how the system designer makes use of CREDO-DLT to specify her/his needs and preferences – *corresponding to steps 1, 2 and 3 in Figure 9* – with regard to the local and global energy trading scenarios. Section 6.3 details the computational stages related to AHP-TOPSIS (step 4), along with the results of the selection analyses (step 5).

6.1. *Knowledge base: Experiment-based scores*

As was previously discussed in section 5.2, some criteria scores can only be quantified under real-life application conditions. As in (Six et al., 2020), the Grid'5000 testbed is used for that purpose, whose overall approach underlying our experiments is summarized in Figure 12. First, the seven DLT platform alternatives reported in Table 6 have been implemented in Grid'5000[4], whose associated source codes are made freely available in an GitHub repository[5]. Second, inputs related to our applicative scenario are specified in Grid'5000, which include the transaction generation rate (i.e., [50;100] Tx/sec in our scenario), the duration of the experiment (10h in our scenario), the set of physical nodes/servers that compose the blockchain network (Luxembourg, Nancy and Lille in our experiments). After running the experiments, logs are analyzed and processed to compute the final scores of the concerned criteria, as illustrated with C_i and C_j in Figure 12, which are then added to the knowledge base. In the following, results obtained with the public/permissionless DLT platforms are presented. All nodes composing the Grid'5000 architecture have an Intel Xeon Gold 5220 (18 cores) processor, 96 GB of RAM, two SSD (480GB, 960GB), and a bandwidth of 2x25 Gbps.

Figure 13 gives insight into the experimental results obtained with Grid'5000 for 6 criteria out of the 14 requiring an experimental stage (*cf.*, criteria highlighted in gray in Table 6). The other criteria could not be evaluated because they required adaptations of the applicative scenario, such as introducing malicious nodes, network latency effects, *etc.*, which is out of scope of the present study. Based on these experimental data, the final score – *to be added to the knowledge base* – is obtained by computing the average values over the $10h$ experiment, resulting in the scores reported in Table 7. Let us note that it was not possible to represent the experimental data of c15.d in the form of a graph/curve.

6.2. *CREDO-DLT: Steps 1, 2*

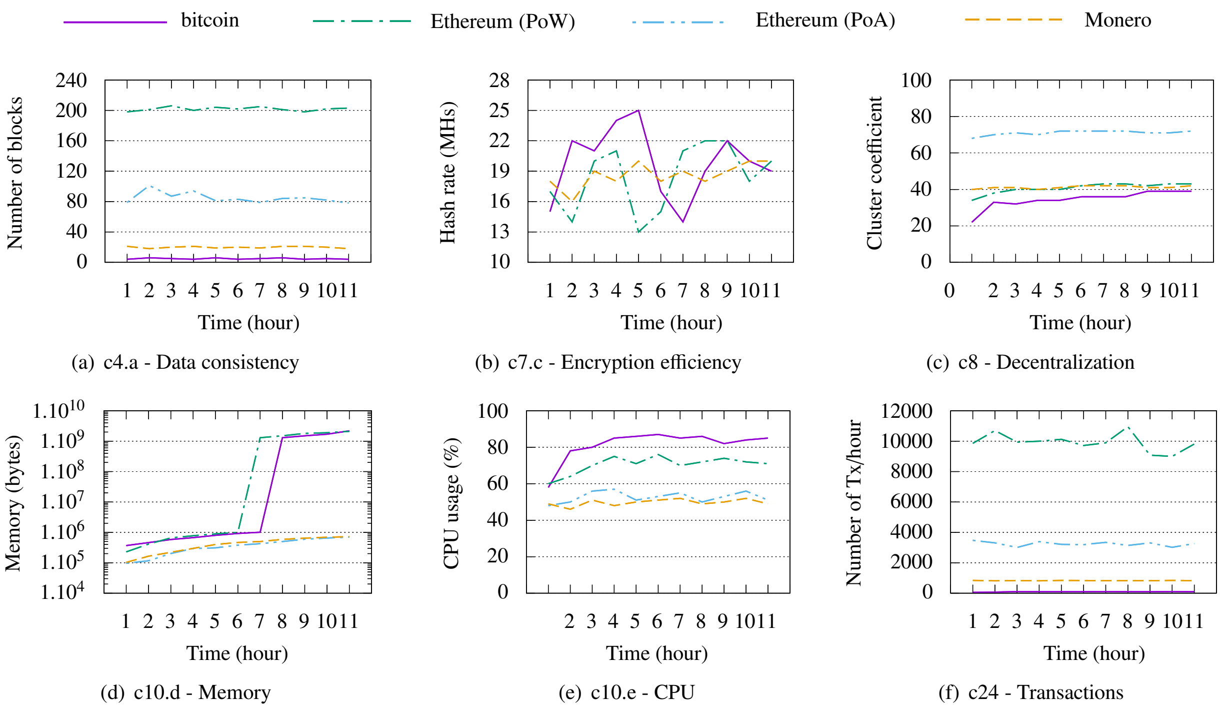
6.2.1. *Step 1: Criteria selection*

The decision-maker wants to identify what platforms are best suited to support the local and global energy trading markets. To do so, the decision-maker online connects to CREDO-DLT[3] (see screenshot denoted by ❶ in Figure 14) and selects the categories and criteria to be included in the decision-making process (see screenshot denoted by ❷). Since the decision-maker has different requirements and needs regarding the local and global energy trading markets, two distinct analyzes are performed. In the latter case (i.e., DLT platform selection for global energy trading), the decision-maker wants to include all categories of criteria, but to exclude some of the criteria belonging to that categories, as summarized in Table 8. In the former case (i.e., DLT platform selection for local energy trading), the decision-maker is not interested in – *and thus does not select* –

[4]https://www.grid5000.fr
[5]https://github.com/deadlyelder/grid5k-dlt

Table 7: Type of information sources considered to evaluate the DLT platform alternatives with regard to each criterion

	Public/Permissionless				Private/Permissioned		
	bitcoin	Eth-PoW	Eth-PoA	Monero	Quorum	MultiChain	Hyperledger
c4.a (Number of blocks per hour)	4.8	201.8	84.8	19.8	1072	335	44
c7.c (MHs)	19.5	17.5	N/A	18	N/A	N/A	N/A
c8 (Cluster coefficient)	0.35	0.41	0.71	0.41	0.87	0.88	0.88
c10.d (Gigabytes)	0.95	0.87	2.57	2.29	3.93	4.57	3.69
c10.e (%)	18.6	29.6	47.3	50.3	58.3	64.5	60.0
c15.d (%)	0.61	0.86	0.69	0.68	0.98	0.97	nf
m24 (Tx/hour)	94	9916	3244	827	1233	994	756



(a) c4.a - Data consistency
(b) c7.c - Encryption efficiency
(c) c8 - Decentralization
(d) c10.d - Memory
(e) c10.e - CPU
(f) c24 - Transactions

Figure 13: Experimental data/results obtained for several criteria under real-life conditions using the Grid'5000 testbed

the Operation and Performance categories, and decides to exclude the criteria listed in Table 8.

6.2.2. *Step 2: DLT platform class identification*

Once criteria have been selected, CREDO-DLT asks a couple of questions to the decision-maker to refine her/his requirements, which corresponds to step 2 of the five-step approach (*cf.*, Figure 10), and to the screenshot denoted by ❸ in Figure 14. One of the key questions aims at identifying what class of DLT platforms best suits the targeted application. In the first case (global trading), the decision-maker expresses the following need: *"Anyone can join, read, write and commit"*, while in the second case (local trading), the following need is expressed: *"Only authorized parties can join and read*. The "Public/Permissionless" and "Private/permissioned" classes are therefore respectively selected for the global and local DLT selection processes. At the time of writing, the following DLT platform alternatives are included in CREDO-DLT for these two classes, (the objective in the long run being to extend them with more platforms):

1. *Public/Permissionless:* {bitcoin, Ethereum-PoW, Ethereum-PoA, Monero};
2. *Private/Permissioned:* {Hyperledger, Multichain, Quorum}.

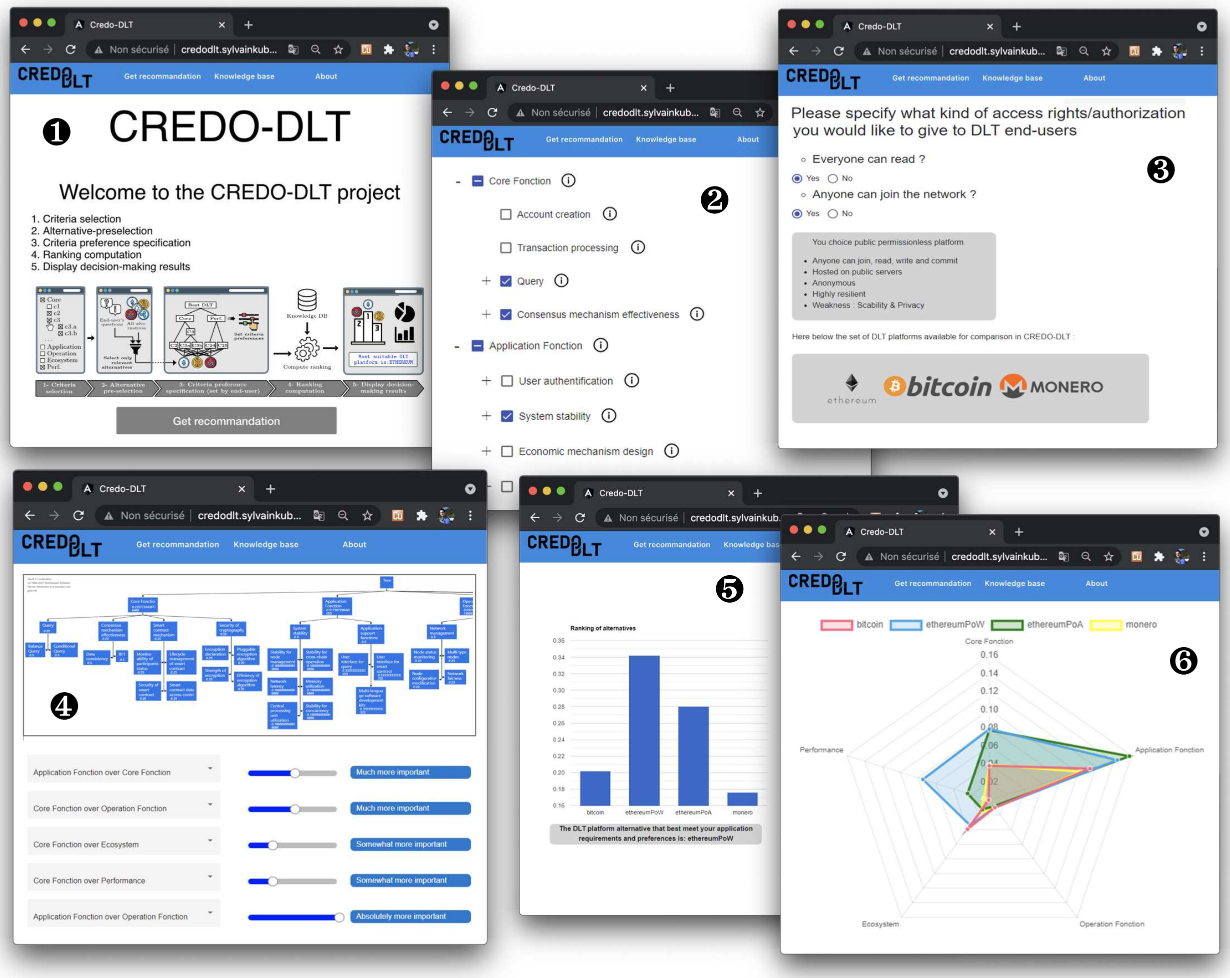


Figure 14: Screenshots of the five user (Web) interfaces of CREDO-DLT that correspond to the five-step approach introduced in Figure 9

Table 8: Set of categories and criteria selected by the decision-maker at step 1 (*cf.*, Figure 9)

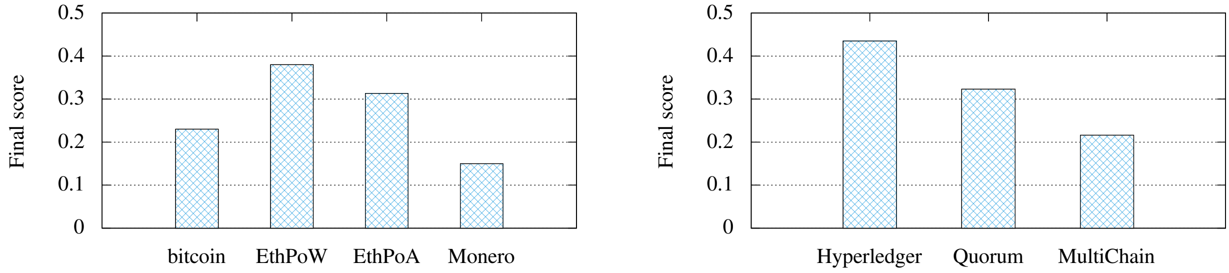
	Category	Selected	Non-selected
Global	Core	c3, c4, c6, c7	c1, c2, c5, c8
	Application	c10, c13	c9, c11, c12, c14
	Operation	c15, c16	-
	Ecosystem	c18, c20, c21	c19, c22, c23
	Performance	c24	c25
Local	Core	c3, c5, c6	c1, c2, c4, c7 c8
	Application	c9, c11, c12, c13	c10, c14
	Operation	-	c15, c16
	Ecosystem	c18, c19, c20, c21	c22, c23
	Performance	-	c24, c25

6.3. CREDO-DLT: Steps 3, 4, 5

6.3.1. Step 3: Criteria preference specification

The third step, which corresponds to screenshot denoted by ❹ in Figure 14, aims at taking into account the decision-maker preferences in terms of criteria importance. In this regard, the decision-maker can select a given level of the AHP structure, and then perform pairwise comparisons between criteria of the corresponding level. For illustrative purposes, let us consider that the decision-maker redefines the importance[6] of the criteria categories (i.e., criteria at L1 of Figure 10) by specifying the pairwise comparison weights (via sliders, as shown by screenshot ❹ in Figure 14), which results in the matrix and the eigenvector respectively given in (2) and (1) (cf., (A.6) for further details about the eigenvector computation). In the local energy trading scenario, all criteria, from all levels, are considered as equally important.

[6]By default, all criteria are defined as equally important.



(a) Final DLT platform scores obtained for Global energy trading (b) Final DLT platform scores obtained for Local energy trading

Figure 15: Score-based ranking of the compared DLT platform alternatives regarding the "Global" and "Local" energy trading scenarios

$$\begin{array}{c} \\ \text{Core} \\ \text{Appli} \\ \text{Oper} \\ \text{Ecos} \\ \text{Perf} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{Core} & \text{Appli} & \text{Oper} & \text{Ecos} & \text{Perf} \end{array} \\ \begin{bmatrix} 1 & \frac{1}{5} & 5 & 3 & 3 \\ 5 & 1 & 9 & 7 & 5 \\ \frac{1}{5} & \frac{1}{9} & 1 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{7} & 3 & 1 & 1 \\ \frac{1}{3} & \frac{1}{5} & 3 & 1 & 1 \end{bmatrix} \end{array} \quad (1)$$

$$\begin{array}{ccccc} W_{\text{Core}} & W_{\text{Appli}} & W_{\text{Oper}} & W_{\text{Ecos}} & W_{\text{Perf}} \end{array} \\ \begin{bmatrix} 0.23 & 0.52 & 0.04 & 0.10 & 0.11 \end{bmatrix} \quad (2)$$

6.3.2. Step 4 & 5: Ranking computation & Display

Once the decision-maker has specified her/his preferences regarding one or more levels of the AHP structure, pairwise comparisons between alternatives with regard to each criterion of level L3 is carried out based on the scores contained by the knowledge base (*cf.* Tables 6 and 7). For illustrative purposes, let us detail the pairwise comparisons between the four public/permissionless DLT platform alternatives {*Monero, bitcoin, Ethereum (PoW), Ethereum (PoA)*} with regard to criterion c18 (Platform maturity). First, the pairwise comparison matrix given in (3), denoted by N1, is defined based on the scores of c18 given in Table 6.

$$N1 = \begin{array}{c} \\ \text{bitc} \\ \text{Eth}_{\text{PoW}} \\ \text{Eth}_{\text{PoA}} \\ \text{Mone.} \end{array} \begin{array}{c} \begin{array}{cccc} \text{bitc} & \text{Eth}_{\text{PoW}} & \text{Eth}_{\text{PoA}} & \text{Mone} \end{array} \\ \begin{bmatrix} 1 & \frac{0.91}{0.52} & \frac{0.91}{0.18} & \frac{0.91}{0.28} \\ \frac{0.52}{0.91} & 1 & \frac{0.52}{0.18} & \frac{0.52}{0.28} \\ \frac{0.18}{0.91} & \frac{0.18}{0.52} & 1 & \frac{0.18}{0.28} \\ \frac{0.28}{0.91} & \frac{0.28}{0.52} & \frac{0.28}{0.18} & 1 \end{bmatrix} \end{array} \quad (3)$$

Second, the max-normalization matrix, denoted by N2 in (4), is computed using (A.4). Computation of element 2, 4 of N2 – *see element highlighted in bold in the matrix* – is detailed in (5).

$$N2 = \begin{array}{c} \\ \text{bitc} \\ \text{Eth}_{\text{PoW}} \\ \text{Eth}_{\text{PoA}} \\ \text{Mone} \end{array} \begin{array}{c} \begin{array}{cccc} \text{bitc} & \text{Eth}_{\text{PoW}} & \text{Eth}_{\text{PoA}} & \text{Mone} \end{array} \\ \begin{bmatrix} 0,20 & 0,36 & 1 & 0,65 \\ 0.11 & 0,20 & 0,56 & \mathbf{0,37} \\ 0.04 & 0.07 & 0.20 & 0.13 \\ 0.06 & 0.11 & 0.31 & 0.20 \end{bmatrix} \end{array} \quad (4)$$

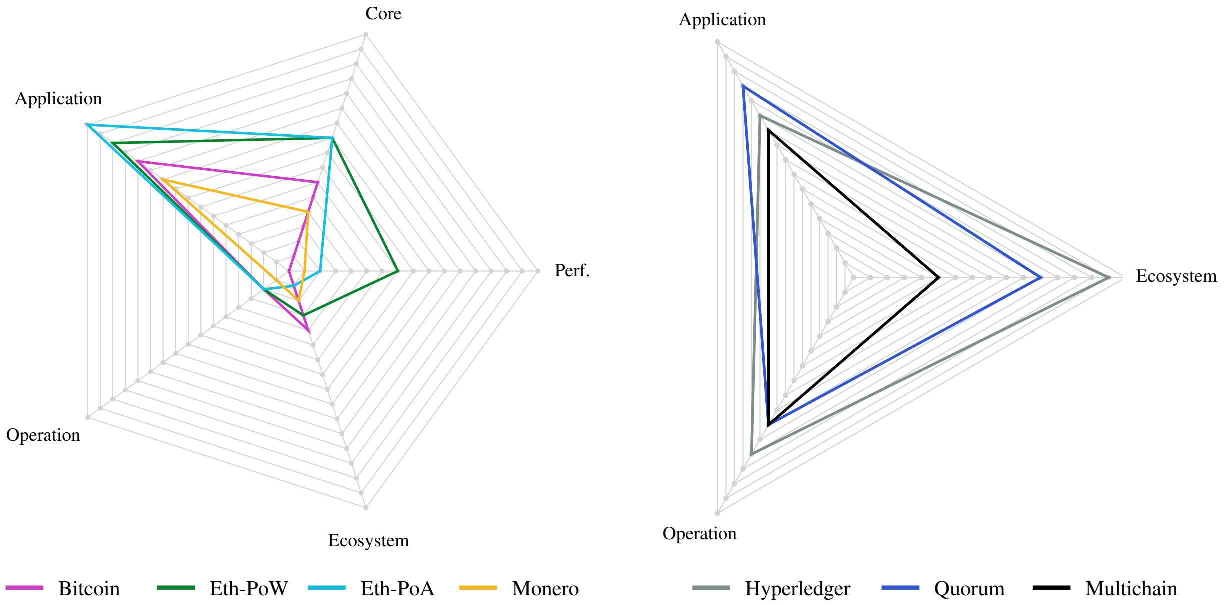
$$n'_{24} = \frac{0.52}{0.28} \times \frac{1}{max(N1)} = \frac{0,33}{2,34} \times \frac{0.18}{0.91} = \mathbf{0.37} \quad (5)$$

Third, the linear-sum matrix, denoted by N3 in (6), is computed using (A.5). Computation of element 2, 4 of N3 – *see element highlighted in bold in the matrix* – is detailed in (7). The eigenvector of N3 is then computed to obtain the final scores of each alternative with respect to each criterion of level L3.

$$N3 = \begin{array}{c} \\ \text{bitc} \\ \text{Eth}_{\text{PoW}} \\ \text{Eth}_{\text{PoA}} \\ \text{Mone} \end{array} \begin{array}{c} \begin{array}{cccc} \text{bitc} & \text{Eth}_{\text{PoW}} & \text{Eth}_{\text{PoA}} & \text{Mone} \end{array} \\ \begin{bmatrix} 0.48 & 0.48 & 0.48 & 0.48 \\ 0,27 & 0,27 & 0,27 & \mathbf{0,27} \\ 0.10 & 0.10 & 0.10 & 0.10 \\ 0.15 & 0.15 & 0.15 & 0.15 \end{bmatrix} \end{array} \quad (6)$$

$$n''_{24} = \frac{0.37}{0.65 + 0.37 + 0.13 + 0.20} = \mathbf{0,27} \quad (7)$$

As detailed in Appendix A.3, TOPSIS is finally applied to aggregate all scores of the AHP structure, based on which the final ranking of alternatives (DLT platforms) is established. For the sake of consistency, this computational step is not detailed in this article; only the final score obtained by each platform alternative regarding the two application scenarios (i.e., global and local energy trading) is presented in the form of histograms in Figure 15 (corresponding to screenshot ❺ in Figure 14). It can be noted that Ethereum PoW and Hyperledger are ranked as the top alternatives respectively regarding the global and local energy trading scenarios.



(a) Ranking of DLT platforms for Global energy trading market

(b) Ranking of DLT platforms for Local energy trading markets

Figure 16: In-depth analysis of how the compared DLT platforms perform with regard to the selected criteria categories

To more thoroughly analyze the results, and better understand how a DLT platform behaves regarding one or more of the criteria categories (i.e., level L1 of the AHP hierarchy), weight aggregation is computed up to L1 and represented in the form of a polar chart in Figure 16 (corresponding to screenshot denoted by ❻ in Figure 14). Overall, regarding the global energy trading scenario, Eth-PoA and Eth-PoW are the best alternatives with regard to the Application and Core categories (see Figure 16(a)); the latter outperforming all the DLT alternatives when it comes to the Performance category. It is also interesting to note that all platforms somehow fail to properly address the last two categories (i.e., Operation and Ecosystem). Considering the local energy trading scenario (see Figure 16(b)), the three Private/Permissioned DLT are compared only with regard to the Application, Ecosystem and Operation criteria categories due to the criteria selection made at step 1 (*cf.*, Table 8). Interestingly, the three platforms perform similarly regarding the Application and Operation categories, but not when it comes to the Ecosystem category, which can be partly explained by the fact that Hyperledger Fabric is highly supported by IBM, Intel, NEC, Linux Foundation and other major organizations.

7. Discussion

To the best of our knowledge, CREDO-DLT is the first decision support tool for DLT platform selection that builds upon the ITU recommendations, although several limitations and improvements still need to be considered.

First, some of the criteria definitions (recommendations) given in the ITU-T F.751.1 document are sometimes highly generic, which makes it difficult to come up with a proper (or unique) formalization of those metrics. This is the main reason why we adopted a systematic literature review approach for all criteria, as it allows us to formalize them in a consensual manner. However, even by doing so, some proposed formalizations can still be debatable and improved in future studies.

Second, as was discussed in Figure 5.2, some criteria require real-life conditions to be measured/quantified, as they are depend on applicative parameters and on the system's dynamic. One may fairly say that this is in contradiction with the objective of CREDO-DLT, namely to support decision-makers in selecting a suitable DLT platform before deploying/implementing it. Blockchain simulators and experimental testbeds (such as Grid'5000) makes possible pre-deployment analyses, but such testbeds are not straightforward to be used, requiring software development and integration stages. In future work, some of these stages coud be further automated and plugged with CREDO-DLT-like decision support tools.

Third, the questionnaire proposed to the decision-maker at step 2 of our approach, which aims both at selecting the right DLT platform class and taking into account specific decision-maker needs/requirements (*cf.*, Figure 9), should be more elaborated in the future by refining the set of questions. For example, when selecting c7.d, a relevant question could be to ask the decision-maker *"how strong the encryption should be?"*, or still, when selecting c13.c *"In which programming language(s) would you like to have a SDK?"*.

Fourth, one may question whether AHP is the most suitable MCDA technique to be applied, which is known to have some weaknesses when it comes to interdependence between criteria/alternatives, inconsistencies between judgments, and rank reversal. Despite this, AHP has many advantages, providing an efficient way to structure the problem in a hierarchically manner, which is suitable for decision-makers (the human brain being able to consider only a limited amount of information at any one time (Simpson, 1996)). Furthermore, AHP is combined with TOPSIS, which has been proven a valuable/robust hybrid solution over the years (Chu et al., 2007; Kubler et al., 2016; James et al., 2021). Finally, let us stress that the main contribution and originality of this paper does not lie much in the use of AHP-TOPSIS, but in the systematic literature approach defined to turn the textually defined ITU assessment criteria into formal (mathematical) ones.

Fifth, it should be noted that this article is not intended to discuss practical, social structure or intellectual property implications and challenges, as it is usually done in traditional systematic literature review. Indeed, the sole aim of our literature review is to identify whether one or more state-of-the-art metrics can be mapped to each ITU assessment criteria, and, if not possible, to propose a formal (mathematical) definition. Having said that, readers can refer to recent studies, as the one presented in (Yalcin and Daim, 2021; Zhang et al., 2021), in which the authors predict future development trends in the blockchain field.

8. Conclusion

Deciding what DLT platform to be used is never an easy task for organizations and developers due to the high number of platform alternatives, all having different characteristics, advantages and disadvantages. A few DLT comparison frameworks (reviewed in this article) have been proposed to help decision-makers in this process. However, they are too are rarely designed based on standardized criteria and/or recommendations. To the best of our knowledge, the only framework that builds upon a standard is the one proposed by Six et al. (2020), although the considered standard (ISO 25010) is not specifically designed for blockchain/DLT applications.

To overcome this lack of standard-based DLT comparison framework, this paper considers the assessment criteria published by the ITU-T Focus Group on Application of Distributed Ledger Technology (FG DLT). The problem is that such criteria are mostly textually defined, which may lead to different interpretations when implementing them, and this problem is facing many other DLT standards, as discussed by König et al. (2020). As a consequence, the present paper proposes a systematic literature approach with the aim of identifying whether one or more state-of-the-art metrics can be mapped to each ITU assessment criteria, and, if not possible, to propose a formal (mathematical) definition. These criteria are used as inputs of a decision support tool called CREDO-DLT[3], AHP and TOPSIS having been used as underlying techniques. At the time of writing, CREDO-DLT is, to the best of our knowledge, the only publicly accessible online tool with BLADE[1] that supports decision-makers in the specification of their requirements and preferences during the platform selection process.

This paper presents a use case scenario in the context of energy communities to showcase how CREDO-DLT could benefit system designers. An originality of this use case lies in the fact that an experimental platform (Grid'5000) has been used to experimentally evaluate some criteria, thus going further than most of today's DLT comparison frameworks that mostly conduct functional comparison analyses.

Acknowledgments

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-20-CE25-0001, as well as of grid5000 for the experiments. The authors also thank Valentin Absalon and Guilain Leduc for their respective inputs.

References

Ar, I. M., Erol, I., Peker, I., Ozdemir, A. I., Medeni, T. D., & Medeni, I. T. (2020). Evaluating the feasibility of blockchain in logistics operations: A decision framework. *Expert Systems with Applications, 158*, 113543.

Baqer, K., Huang, D. Y., McCoy, D., & Weaver, N. (2016, February). Stressing out: Bitcoin “stress testing”. In International Conference on Financial Cryptography and Data Security (pp. 3-18). Springer, Berlin, Heidelberg.

Bai, C., & Sarkis, J. (2020). A supply chain transparency and sustainability technology appraisal model for blockchain technology. *International Journal of Production Research, 58*(7), 2142-2162.

Baliga, A., Subhod, I., Kamat, P., & Chatterjee, S. (2018). Performance evaluation of the quorum blockchain platform. arXiv preprint arXiv:1809.03421.

Balasch, J., Ege, B., Eisenbarth, T., Gérard, B., Gong, Z., Güneysu, T., Heyse, S., Kerckhof, S., Koeune, F., Plos, T., Pöppelmann, T., Regazzoni, F., Standaert, F.-X., Van Assche, G., Van Keer, R., van Oldeneel tot Oldenzeel, L., & von Maurich, I. (2012, November). Compact implementation and performance evaluation of hash functions in attiny devices. In International Conference on Smart Card Research and Advanced Applications (pp. 158-172). Springer, Berlin, Heidelberg.

Bamakan, S. M. H., Motavali, A., & Bondarti, A. B. (2020). A survey of blockchain consensus algorithms performance evaluation criteria. *Expert Systems with Applications, 154*, 113385.

Bhatt, P. C., Kumar, V., Lu, T.-C., & Daim, T. (2019). Technology convergence assessment: Case of blockchain within the IR 4.0 platform. *Technology in Society, 67*, 101709.

Belotti, M., Božić, N., Pujolle, G., & Secci, S. (2019). A vademecum on blockchain technologies: When, which, and how. *IEEE Communications Surveys & Tutorials, 21*(4), 3796-3838.

Böckel, A., Nuzum, A.-K., & Weissbrod, I. (2021). Blockchain for the circular economy: analysis of the research-practice gap. *Sustainable Production and Consumption, 25*, 525–539.

Bolze, R., Cappello, F., Caron, E., Daydé, M., Desprez, F., Jeannot, E., Jégou, Y., Lanteri, S., Leduc, J., & Melab, N. (2006). Grid’5000: a large scale and highly reconfigurable experimental grid testbed. *The International Journal of High Performance Computing Applications, 20*(4), 481-494.

Borkowski, M., Sigwart, M., Frauenthaler, P., Hukkinen, T., & Schulte, S. (2019). DeXTT: Deterministic cross-blockchain token transfers. *IEEE Access, 7*, 111030-111042.

Bouraga, S. (2021). A taxonomy of blockchain consensus protocols: A survey and classification framework. *Expert Systems with Applications, 168*, 114384.

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software, 80*(4), 571-583.

Budak, A., & Çoban, V. (2021). Evaluation of the impact of blockchain technology on supply chain using cognitive maps. *Expert Systems with Applications, 184*, 115455.

Büyüközkan, G., & Tüfekçi, G. (2021). A decision-making framework for evaluating appropriate business blockchain platforms using multiple preference formats and VIKOR. Information Sciences, 571, 337-357.

Chaudhry, N., & Yousaf, M. M. (2018, December). Consensus algorithms in blockchain: comparative analysis, challenges and opportunities. In 12th IEEE International Conference on Open Source Systems and Technologies (pp. 54-63).

Cong, X., & Zi, L. (2020). Blockchain applications, challenges and evaluation: A survey. *Discrete Mathematics, Algorithms and Applications, 12*(04), 2030001.

Chu, M. T., Shyu, J., Tzeng, G. H., & Khosla, R. (2007). Comparison among three analytical methods for knowledge communities group-decision analysis. *Expert Systems with Applications, 33*(4), 1011-1024.

Daim, T. and Lai, K. K. and Yalcin, H. and Alsoubie, F. & Kumar, V. (2020). Forecasting technological positioning through technology knowledge redundancy: Patent citation analysis of IoT, cybersecurity, and Blockchain. *Technological Forecasting and Social Change, 161*, 120329.

de São José, D., Faria, P., & Vale, Z. (2021). Smart energy community: A systematic review with metanalysis. *Energy Strategy Reviews, 36*, 100678.

Delgado-Mohatar, O., Fierrez, J., Tolosana, R., & Vera-Rodriguez, R. (2019). Biometric template storage with blockchain: A first Look into cost and performance tradeoffs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops.

Dinh, T. T. A., Liu, R., Zhang, M., Chen, G., Ooi, B. C., & Wang, J. (2018). Untangling blockchain: A data processing view of blockchain systems. *IEEE transactions on knowledge and data engineering, 30*(7), 1366-1385.

Dong, Z., Zheng, E., Choon, Y., & Zomaya, A. Y. (2019, July). Dagbench: A performance evaluation framework for dag distributed ledgers. In IEEE 12th International Conference on Cloud Computing (pp. 264-271).

Drljevic, N., Aranda, D. A., Stantchev, V., & Weissbrod, I. (2020). Perspectives on risks and standards that affect the requirements engineering of blockchain technology. *Computer Standards & Interfaces, 69*, 103409.

Eskandari, S., Clark, J., Barrera, D., & Stobert, E. (2018). A first look at the usability of bitcoin key management. arXiv preprint arXiv:1802.04351.

Fan, C., Ghaemi, S., Khazaei, H., & Musilek, P. (2020). Performance evaluation of blockchain systems: A systematic survey. *IEEE Access, 8*, 126927-126950.

Faria, C., & Correia, M. (2019, July). BlockSim: blockchain simulator. In IEEE International Conference on Blockchain (Blockchain) (pp. 439-446).

Frauenthaler, P., Sigwart, M., Spanring, C., & Schulte, S. (2020). Advanced Cross-Chain Token Transfers. TU Wien.

Gervais, A., Karame, G. O., Wüst, K., Glykantzis, V., Ritzdorf, H., & Capkun, S. (2016, October). On the security and performance of proof of work blockchains. In Proceedings of the ACM SIGSAC conference on computer and communications security (pp. 3-16).

Gochhayat, S. P., Shetty, S., Mukkamala, R., Foytik, P., Kamhoua, G. A., & Njilla, L. (2020). Measuring Decentrality in Blockchain Based Systems. *IEEE Access, 8*, 178372-178390.

Goffard, P. O. (2019). Fraud risk assessment within blockchain transactions. *Advances in Applied Probability, 51*(2), 443-467.

Gopalan, A., Sankararaman, A., Walid, A., & Vishwanath, S. (2020). Stability and scalability of blockchain systems. *Proceedings of the ACM on Measurement and Analysis of Computing Systems, 4*(2), 1-35.

Gowat, R. C. (2020). Blockchain Traceability for the Counterfeit Detection and Avoidance Program (CDAP) Final Report. Accenture Federal Services Arlington United States.

Gräbe, F., Kannengiesser, N., Lins, S., & Sunyaev, A. (2020). Do not be fooled: Toward a holistic comparison of Distributed Ledger Technology designs. In Proceedings of the 53rd Hawaii International Conference on System Sciences (6297-6306).

Gupta, A., & Shankarananda, B. M. (2015, January). Fast interest recovery in content centric networking under lossy environment. In 12th Annual IEEE Consumer Communications and Networking Conference (pp. 802-807).

Han, R., Shapiro, G., Gramoli, V., & Xu, X. (2020). On the performance of distributed ledgers for internet of things. *Internet of Things, 10*, 100087.

Hao, Y., Li, Y., Dong, X., Fang, L., & Chen, P. (2018, June). Performance analysis of consensus algorithm in private blockchain. In IEEE Intelligent Vehicles Symposium (pp. 280-285).

Ho, G. T. S., Tang, Y. M., Tsang, K. Y., Tang, V., & Chau, K. Y.

(2021). A blockchain-based system to enhance aircraft parts traceability and trackability for inventory management. *Expert Systems with Applications, 179*, 115101.

Honig, J. J., Everts, M. H., & Huisman, M. (2019). Practical mutation testing for smart contracts. In Data Privacy Management, Cryptocurrencies and Blockchain Technology (pp. 289-303). Springer, Cham.

Huang, B., Jin, L., Lu, Z., Zhou, X., Wu, J., Tang, Q., & Hung, P. C. (2019). Bor: Toward high-performance permissioned blockchain in rdma-enabled network. *IEEE Transactions on Services Computing, 13*(2), 301-313.

Tzeng, G. H., & Huang, J. J. (2011). *Multiple attribute decision making: methods and applications*. CRC press.

Jain, R. K., Chiu, D. M. W., & Hawe, W. R. (1984). A quantitative measure of fairness and discrimination. Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA.

James, A. T., Vaidya, D., Sodawala, M., & Verma, S. (2021). Selection of bus chassis for large fleet operators in India: An AHP-TOPSIS approach. Expert Systems with Applications, 186, 115760.

Kabbinale, A. R., Dimogerontakis, E., Selimi, M., Ali, A., Navarro, L., Sathiaseelan, A., & Crowcroft, J. (2020). Blockchain for economically sustainable wireless mesh networks. *Concurrency and Computation: Practice and Experience, 32*(12), e5349.

Kalra, S., Goel, S., Dhawan, M., & Sharma, S. (2018, February). Zeus: Analyzing safety of smart contracts. In Ndss (pp. 1-12).

Kaur, K., Singh, J., & Ghumman, N. S. (2014, February). Mininet as software defined networking testing platform. In International Conference on Communication, Computing & Systems (pp. 139-42).

Kirillov, D., Iakushkin, O., Korkhov, V., & Petrunin, V. (2019, July). Evaluation of tools for analyzing smart contracts in distributed ledger technologies. In International Conference on Computational Science and Its Applications (pp. 522-536). Springer, Cham.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology, 51*(1), 7-15.

Kobayashi, K., Ikegami, J., Knežević, M., Guo, E. X., Matsuo, S., Huang, S., Nazhandali, L., Kocabaş, U., Fan, J., & Satoh, A. (2010, June). Prototyping platform for performance evaluation of SHA-3 candidates. In IEEE International Symposium on Hardware-Oriented Security and Trust (pp. 60-63).

Kolbe, N., Kubler, S., Robert, J., Le Traon, Y., & Zaslavsky, A. (2019). Linked vocabulary recommendation tools for internet of things: a survey. *ACM Computing Surveys (CSUR), 51*(6), 1-31.

König, L., Korobeinikova, Y., Tjoa, S., & Kieseberg, P. (2020). Comparing Blockchain Standards and Recommendations. *Future Internet, 12*(12), 222.

N. Kshetri, Kshetri, N. (2018). 1 Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management, 39*, 80-89.

Kubler, S., Robert, J., Derigent, W., Voisin, A., & Le Traon, Y. (2016). A state-of the-art survey & testbed of fuzzy AHP (FAHP) applications. *Expert Systems with Applications, 65*, 398-422.

Kuznetsov, A., Oleshko, I., Tymchenko, V., Lisitsky, K., Rodinko, M., & Kolhatin, A. (2021). Performance Analysis of Cryptographic Hash Functions Suitable for Use in Blockchain. *International Journal of Computer Network & Information Security, 13*(2).

Labazova, O. (2019). Towards a framework for evaluation of blockchain implementations. In International Conference on Information Systems (1-16).

Leduc, G., Kubler, S., & Georges, J. P. (2021). Innovative blockchain-based farming marketplace and smart contract performance evaluation. *Journal of Cleaner Production, 306*, 127055.

Li, Y., Chen, X., Wang, X., Xu, Y., & Chen, P. H. (2017). A review of studies on green building assessment methods by comparative analysis. *Energy and Buildings, 146*, 152-159.

Lohachab, A., Garg, S., Kang, B. H., & Amin, M. B. (2021). Performance evaluation of Hyperledger Fabric-enabled framework for pervasive peer-to-peer energy trading in smart Cyber–Physical Systems. *Future Generation Computer Systems, 118*, 392-416.

Luu, L., Chu, D. H., Olickel, H., Saxena, P., & Hobor, A. (2016, October). Making smart contracts smarter. In Proceedings of the ACM SIGSAC conference on computer and communications security (pp. 254-269).

Mackay, B. (2019). Evaluation of Security in Hardware and Software Cryptocurrency Wallets. Ph.D. thesis, Edinburgh Napier University.

Maesa, D. D. F., & Mori, P. (2020). Blockchain 3.0 applications survey. *Journal of Parallel and Distributed Computing, 138*, 99-114.

Maranhão, S., Seigneur, J. M., & Hu, R. (2019). Towards a standard to assess blockchain & other DLT platforms. ITU (1-7).

Mathew, S., & Jacob, K. P. (2010). Performance evaluation of popular hash functions. *World Academy of Science, Engineering and Technology, 61*, 449-452.

Miller, A., & Jansen, R. (2015). Shadow-bitcoin: Scalable simulation via direct execution of multi-threaded applications. In 8th Workshop on Cyber Security Experimentation and Test.

Miyamae, T., Honda, T., Tamura, M., & Kawaba, M. (2018). Performance improvement of the consortium blockchain for financial business applications. *Journal of Digital Banking, 2*(4), 369-378.

Modrak, V., Radu, S. M., & Modrak, J. (2014). Metrics in organizational centralization and decentralization. *Polish Journal of Management Studies, 10*.

Nanayakkara, S., Rodrigo, M. N. N., Perera, S., Weerasuriya, G. T., & Hijazi, A. A. (2021). A methodology for selection of a Blockchain platform to develop an enterprise system. *Journal of Industrial Information Integration, 23*, 100215.

Oliveira, L., Zavolokina, L., Bauer, I., & Schwabe, G. (2018). To token or not to token: Tools for understanding blockchain tokens. International Conference of Information Systems.

Parizi, R. M., Dehghantanha, A., Choo, K. K. R., & Singh, A. (2018). Empirical vulnerability analysis of automated smart contracts security testing on blockchains. arXiv preprint arXiv:1809.02702.

Park, S., Oh, S., & Kim, H. (2019, May). Performance analysis of DAG-based cryptocurrency. In IEEE International Conference on Communications workshops (pp. 1-6).

Pisa, M. (2018). Reassessing expectations for blockchain and development. *Innovations: Technology, Governance, Globalization, 12*(1-2), 80-88.

Polge, J., Ghatpande, S., Kubler, S., Robert, J., & Le Traon, Y. (2021). BlockPerf: A hybrid blockchain emulator/simulator framework. *IEEE Access, 9*, 107858-107872.

Polge, J., Robert, J., & Le Traon, Y. (2021). Permissioned blockchain frameworks in the industry: A comparison. *ICT Express, 7*(2), 229-233.

Prechtel, D., Gross, T., & Müller, T. (2019, June). Evaluating spread of 'gasless send'in Ethereum smart contracts. In 2019 10th IFIP International Conference on New Technologies, Mobility and Security (pp. 1-6).

Qing, S., Liu, X., & Zheng, H. (2020, October). An Assessment Framework for Distributed Ledger Technology in Financial Application. In Proceedings of the 2nd ACM International Symposium on Blockchain and Secure Critical Infrastructure (pp. 161-170).

Saad, M., Kim, J., Nyang, D., & Mohaisen, D. (2021). Contra-*: Mechanisms for countering spam attacks on blockchain's memory pools. *Journal of Network and Computer Applications, 179*, 102971.

T. L. Saaty, Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of mathematical psychology, 15*(3),

234-281.
Sagirlar, G., Carminati, B., Ferrari, E., Sheehan, J. D., & Ragnoli, E. (2018, July). Hybrid-iot: Hybrid blockchain architecture for internet of things-pow sub-blockchains. In IEEE International Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data (pp. 1007-1016).
Sayadi, S., Rejeb, S. B., & Choukair, Z. (2019, June). Anomaly detection model over blockchain electronic transactions. In 15th International Wireless Communications & Mobile Computing Conference (pp. 895-900).
Schmitz, P.-E., & Cacciaguerra Ranghieri, G. (2018). Joinup Licensing Assistant. White paper v1.01, European Commission - DG DIGIT Unit.D2 (Interoperability Unit).
Selimi, M., Kabbinale, A. R., Ali, A., Navarro, L., & Sathiaseelan, A. (2018, June). Towards blockchain-enabled wireless mesh networks. In Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems (pp. 13-18).
Sharma, M., Sehrawat, R., Daim, T., & Shaygan, A. (2021). Technology assessment: Enabling Blockchain in hospitality and tourism sectors. *Technological Forecasting and Social Change, 169*, 120810.
Shaw, M. R. (2018). A blockchain testbed for DoD applications. Naval Postgraduate School.
Simpson, L. (1996). Do decision makers know what they prefer?: MAVT and ELECTRE II. *Journal of the Operational Research Society, 47*(7), 919-929.
Six, N., Herbaut, N., & Salinesi, C. (2020). Which Blockchain to choose? A decision support tool to guide the choice of a Blockchain technology. arXiv preprint arXiv:2004.06080.
Smetanin, S., Ometov, A., Komarov, M., Masek, P., & Koucheryavy, Y. (2020). Blockchain evaluation approaches: State-of-the-art and future perspective. *Sensors, 20*(12), 3358.
Srivastav, R. K., Agrawal, D., & Shrivastava, A. (2020). A Survey on Vulnerabilities and Performance Evaluation Criteria in Blockchain Technology. *Advances in Distributed Computing and Artificial Intelligence Journal, 9*, 91-105.
L. Stoykov, K. Zhang, Stoykov, L., Zhang, K., & Jacobsen, H. A. (2017, December). VIBES: fast blockchain simulations for large-scale peer-to-peer networks. In Proceedings of the 18th ACM/IFIP/USENIX Middleware Conference: Posters and Demos (pp. 19-20).
StudyGroup16/22. (2020). F.751.1 Assessment criteria for distributed ledger technology platforms. International Telecommunication Union (ITU).
Suankaewmanee, K., Hoang, D. T., Niyato, D., Sawadsitang, S., Wang, P., & Han, Z. (2018, March). Performance analysis and application of mobile blockchain. In 2018 international conference on computing, networking and communications (pp. 642-646).
Sukhwani, H., Wang, N., Trivedi, K. S., & Rindos, A. (2018, November). Performance modeling of Hyperledger Fabric (permissioned blockchain network). In IEEE 17th International Symposium on Network Computing and Applications (pp. 1-8).
Suratkar, S., Shirole, M., & Bhirud, S. (2020, September). Cryptocurrency Wallet: A Review. In 4th IEEE International Conference on Computer, Communication and Signal Processing (pp. 1-7).
Szczerbowski, J. J. (2018). Transaction costs of blockchain smart contracts. *Law and forensic science, 16*, 2.
Tang, H., Shi, Y., & Dong, P. (2019). Public blockchain evaluation using entropy and TOPSIS. *Expert Systems with Applications, 117*, 204-210.
Tolk, A., Diallo, S. Y., & Turnitsa, C. D. (2007). Applying the levels of conceptual interoperability model in support of integratability, interoperability, and composability for system-of-systems engineering. *Journal of Systems, Cybernetics, and Informatics, 5*(5).
Tucker, A. B. (Ed.). (2004). *Computer science handbook*. CRC press.
Uesugi, T., Shijo, Y., & Murata, M. (2020). Short Paper: Design and Evaluation of Privacy-preserved Supply Chain System based on Public Blockchain. arXiv preprint arXiv:2004.07606.
Vafaei, N., Ribeiro, R. A., & Camarinha-Matos, L. M. (2016, April). Normalization techniques for multi-criteria decision making: Analytical Hierarchy Process case study. In doctoral conference on computing, electrical and industrial systems (pp. 261-269). Springer, Cham.
Wan, L., Eyers, D., & Zhang, H. (2019, July). Evaluating the impact of network latency on the safety of blockchain transactions. In IEEE International Conference on Blockchain (pp. 194-201).
Wang, X., Zha, X., Ni, W., Liu, R. P., Guo, Y. J., Niu, X., & Zheng, K. (2019). Survey on blockchain for Internet of Things. *Computer Communications, 136*, 10-29.
Wang, B., Zhao, S., Li, Y., Wu, C., Tan, J., Li, H., & Yukita, K. (2021). Design of a privacy-preserving decentralized energy trading scheme in blockchain network environment. *International Journal of Electrical Power & Energy Systems, 125*, 106465.
Wöhrer, M., & Zdun, U. (2018, July). Design patterns for smart contracts in the Ethereum ecosystem. In IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) (pp. 1513-1520).
Xiao, Y., Zhang, N., Lou, W., & Hou, Y. T. (2020). A survey of distributed consensus protocols for blockchain networks. *IEEE Communications Surveys & Tutorials, 22*(2), 1432-1465.
Xie, J., Tang, H., Huang, T., Yu, F. R., Xie, R., Liu, J., & Liu, Y. (2019). A survey of blockchain technology applied to smart cities: Research issues and challenges. *IEEE Communications Surveys & Tutorials, 21*(3), 2794-2830.
Yalcin, H., & Daim, T. (2021). Mining research and invention activity for innovation trends: case of blockchain technology. *Scientometrics, 126*(5), 3775-3806.
Yang, W., Garg, S., Huang, Z., & Kang, B. (2021). A decision model for blockchain applicability into knowledge-based conversation system. *Knowledge-Based Systems, 220*, 106791.
Zhang, H., Daim, T., & Zhang, Y. P. (2021). Integrating patent analysis into technology roadmapping: A latent dirichlet allocation based technology assessment and roadmapping in the field of Blockchain. *Technological Forecasting and Social Change, 167*, 120729.
Zhang, R., & Preneel, B. (2017, February). Publish or perish: A backward-compatible defense against selfish mining in bitcoin. In Cryptographers' Track at the RSA Conference (pp. 277-292). Springer, Cham.
Zhong, Q. T., & Cole, Z. (2018). Analyzing the Effects of Network Latency on Blockchain Performance and Security Using the Whiteblock Testing Platform. Blockchain Performance and Security.

Appendix A. AHP/TOPSIS-related computational stages applied in CREDO-DLT

In the following, section Appendix A.1 details the pairwise comparison process implemented at level L1, L2, L3 of the AHP structure (*cf.*, Figure 10), which is based on the end-user preferences in terms of criteria importance, while section Appendix A.2 details the process implemented at level alternative level (i.e. L4), which makes use of the performance scores stored in the knowledge base (*cf.*, section 5.2). Section Appendix A.3 finally details the weight aggregation process used to calculate the final ranking of the alternatives.

Appendix A.1. Preference-based pairwise comparisons

Let m be the number of criteria to be compared; e.g., at level L2 of the AHP structure in Figure 10, $m = |\{Core, Application, Operation, Ecosystem, Perf.\}| = 5$. The evaluation performed by the expert is made using the 1- to 9-point Saaty's scale: $\{1, 3, 5, 7, 9\}$, where $w_{ij} = 1$ means that C_i and C_j are of equal importance, while $w_{ij} = 9$ means that C_i is strongly favored over C_j. A pairwise comparison is denoted by P, as given in (A.1). The normalized eigenvector of $N1$ is computed using (A.6).

$$P = \begin{array}{c} \\ \mathrm{C}_1 \\ \vdots \\ \mathrm{C}_m \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{C}_1 & \dots & \mathrm{C}_m \end{array} \\ \begin{bmatrix} w_{11} & \dots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \dots & w_{mm} \end{bmatrix} \end{array} \tag{A.1}$$

$$W_{C_i} = \frac{\sum_{j=1}^{m} w_{ij}}{\sum_{k=1}^{m} \sum_{j=1}^{m} w_{kj}}, \quad w_{ji} = \begin{cases} 1 & i = j \\ \frac{1}{w_{ij}} & i \neq j \end{cases} \tag{A.2}$$
$$W_C = [W_{C_1}, .., W_{C_i}, .., W_{C_m}]$$

Appendix A.2. Score-based pairwise comparisons

Pairwise comparisons at the alternative level is based on the performance scores that have been quantified and stored in the knowledge base. For example, let us consider the pairwise comparison of the alternatives with regard to criterion c10.b[7], which is denoted by $N1$ in (A.3). If "Monero" (considered as alternative A_1) supports cross-swap operations with two other DLT platforms, while Ethereum (considered as alternative A_z) supports cross-swap operations with five other DLT platforms, then the corresponding pairwise comparison score, denoted by $\frac{\mathrm{I}_{C_{10.b}}(\mathrm{A}_1)}{\mathrm{I}_{C_{10.b}}(\mathrm{A}_z)}$ in (A.3) is equal to $\frac{2}{5}$.

$$N1 = \begin{array}{c} \\ \mathrm{A}_1 \\ \vdots \\ \mathrm{A}_z \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{A}_1 & \dots & \mathrm{A}_z \end{array} \\ \begin{bmatrix} 1 & \dots & \frac{\mathrm{I}_{C_{10.b}}(\mathrm{A}_1)}{\mathrm{I}_{C_{10.b}}(\mathrm{A}_z)} \\ \vdots & \ddots & \vdots \\ \frac{\mathrm{I}_{C_{10.b}}(\mathrm{A}_z)}{\mathrm{I}_{C_{10.b}}(\mathrm{A}_1)} & \dots & 1 \end{bmatrix} \end{array} \tag{A.3}$$

[7] c10.b defining the number of platforms with which a given DLT platform alternative can perform cross-swap operations

Given the fact that the performance scores at the alternative level rely on different numeric scales, a normalization stage is required to obtain dimensionless classifications (i.e., to transform data in different units into a common scale and comparable units) for aggregating all scores. According to Vafaei et al. (2016), the combination of the "max-normalization" with the "linear-sum" techniques seems the most suitable for AHP, which is adopted in this study. The max-normalization stage consists in applying (A.4), resulting in the pairwise comparison matrix denoted by N2 in that equation (C_p referring to criterion p).

$$N2_{i,j} = \frac{N1_{i,j}}{max\left(N1_{i,j}\right)} \tag{A.4}$$

$$N2 = \begin{array}{c} \\ \mathrm{A}_1 \\ \vdots \\ \mathrm{A}_z \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{A}_1 & \dots & \mathrm{A}_z \end{array} \\ \begin{bmatrix} \frac{1}{max(N1_{i,j})} & \dots & \frac{\mathrm{I}_{C_p}(\mathrm{A}_1)}{\mathrm{I}_{C_p}(\mathrm{A}_z) \cdot max(N1_{i,j})} \\ \vdots & \ddots & \vdots \\ \frac{\mathrm{I}_{C_p}(\mathrm{A}_z)}{\mathrm{I}_{C_p}(\mathrm{A}_1) \cdot max(N1_{i,j})} & \dots & 1 \end{bmatrix} \end{array}$$

The linear-sum technique is finally applied using (A.5), resulting in a pairwise comparison matrix denoted by N3. The eigenvector of N3, denoted by $W^A_{C_p}$, is then computed using (A.6) to obtain the final score of each alternative with regard to criterion C_p.

$$N3_{i,j} = \frac{N2_{i,j}}{\sum_{i=1}^{m} N2_{i,j}} \tag{A.5}$$

$$W^A_{C_p} = \frac{\sum_{j=1}^{z} N2_{i,j}}{\sum_{k=1}^{z} \sum_{j=1}^{z} N2_{k,j}} \quad N2_{ji} = \begin{cases} 1 & i = j \\ \frac{1}{N2_{ij}} & i \neq j \end{cases} \tag{A.6}$$
$$W^A_{C_p} = [W^{A_1}_{C_p}, .., W^{A_z}_{C_p}]$$

Appendix A.3. Alternative raking

TOPSIS, combined with AHP, has been proven a valuable/robust hybrid solution over the years for generating the final ranking of alternatives (Chu et al., 2007; Kubler et al., 2016; James et al., 2021). TOPSIS introduces for each alternative the closeness coefficient denoted by $R(\mathrm{A}_l)$, which implies computing for each criterion xh the positive ideal solution (PIS) denoted by d^+_{xh} and negative ideal solution (NIS) denoted by d^-_{xh}, as formalized in (A.8) and (A.9) respectively. The distances measuring the separation from PIS and NIS are then computed in (A.10) and (A.11), respectively denoted $D^+_{\mathrm{A}_l}$ and $D^-_{\mathrm{A}_l}$), where GW corresponds to the global weight of a given alternative A_k based on (A.7), where $W^{\mathrm{L}j}_{C_i}$, which is computed using (A.7), refers to the weight of the parent criterion of C_i at level L_j, and $\mathcal{Y}$ refers to the set of criteria at level L3 of the AHP structure (cf., Figure 10).

$$GW_{C_p}^{\mathrm{A}_l} = W_{C_p}^{\mathrm{A}_l} \cdot W_{C_p}^{\mathrm{L3}} \cdot W_{C_p}^{\mathrm{L2}} \cdot W_{C_p}^{\mathrm{L1}} \tag{A.7}$$

$$d_p^+ = \max_{l=1..z}\left(GW_{C_p}^{\mathrm{A}_l}\right) \tag{A.8}$$

$$d_p^- = \min_{l=1..z}\left(GW_{C_p}^{\mathrm{A}_l}\right) \tag{A.9}$$

$$D^+(\mathrm{A}_l) = \sqrt{\sum_p \left(GW_{C_p}^{\mathrm{A}_l} - d_p^+\right)^2} \quad l = 1, .., z \tag{A.10}$$

$$D^-(\mathrm{A}_l) = \sqrt{\sum_p \left(GW_{C_p}^{\mathrm{A}_l} - d_p^-\right)^2} \quad l = 1, .., z \tag{A.11}$$

A prior alternative has a longer distance to NIS and a shorter distance to PIS. Consequently, the closeness coefficient to the ideal solution for each alternative can be formulated as in (A.12), where $R(\mathrm{A}_l)$ denotes the final performance score of the DLT platform l. The larger the $R(\mathrm{A}_l)$ score, the more the DLT platform meets the decision-maker needs and requirements. The overall DLT platform ranking is finally generated using the $R(\mathrm{A}_l)$ performance scores.

$$R(\mathrm{A}_l) = \frac{D^-(\mathrm{A}_l)}{D^+(\mathrm{A}_l) + D^-(\mathrm{A}_l)} \quad l = 1, .., z \tag{A.12}$$