DREAM Principles and FAIR Metrics from the PORTAL-DOORS Project for the Semantic Web

1st Adam Craig
Brain Health Alliance
Ladera Ranch, California
acraig@bhavi.us

2nd Adarsh Ambati
Brain Health Alliance
Ladera Ranch, California
aambati@bhavi.us

3rd Shiladitya Dutta
Brain Health Alliance
Ladera Ranch, California
sdutta@bhavi.us

4th Pooja Kowshik
Brain Health Alliance
Ladera Ranch, California
pkowshik@bhavi.us

5th Sathvik Nori
Brain Health Alliance
Ladera Ranch, California
snori@bhavi.us

6th S. Koby Taswell
Brain Health Alliance
Ladera Ranch, California
ktaswell@bhavi.us

7th Qiyuan Wu
Brain Health Alliance
Ladera Ranch, California
qwu@bhavi.us

8th Carl Taswell*
Brain Health Alliance
Ladera Ranch, California
ctaswell@bhavi.us

***Abstract*—Articles published in *Scientific Data* by Wilkinson *et al.* argued for the adoption of the Findable, Accessible, Interoperable, and Reusable (FAIR) principles of data management without citing any of the prior work published by Taswell. However, these principles were first proposed and described by Taswell in 2006 as the foundation for work on the PORTAL-DOORS Project (PDP) and the Nexus-PORTAL-DOORS-Scribe (NPDS) cyberinfrastructure, and have been published in numerous conference presentations, journal articles, and patents. This work on PDP and NPDS has been continuously available since 2007 from a publicly accessible web site at www.portaldoors.org, and discussed in person at conferences with several key authors of the Wilkinson *et al.* papers. Paraphrasing without citing the PDP and NPDS principles while renaming them as the FAIR principles raises questions about both the 'FAIRness' and the fairness of the authors of the Wilkinson *et al.* papers. Promoting these principles with the use of the term 'metrics', which are not metrics by definition of the term metric as used in most fields of science, also raises questions about their commitment to maintaining consistency of usage for basic terminology across different fields of science as should be expected for terms in ontology mapping with knowledge engineering for the semantic web. Therefore, in the present report, we clarify the origin of their FAIR principles by identifying our PDP and NPDS principles that constitute the historical precedent for their FAIR principles. Moreover, as the comprehensively summarizing phrase for all of our PDP and NPDS principles, we rename them the DREAM principles with the acronym DREAM for *Discoverable Data with Reproducible Results for Equivalent Entities with Accessible Attributes and Manageable Metadata.* Finally, we define numerically valid quantitative FAIR metrics to monitor and measure the DREAM principles from the perspective of the most important principle, ie, the *Fair Acknowledgment of Information Records* and *Fair Attribution to Indexed Reports*, for maintaining fair standards of citation in scholarly research and publishing.**

***Index Terms*—Data stewardship, metadata management, NPDS cyberinfrastructure, DREAM principles, FAIR metrics.**

I. Introduction

A comment authored by Wilkinson *et al.* and published by Nature in *Scientific Data* stated in its abstract that "A diverse set of stakeholders – representing academia, industry, funding agencies, and scholarly publishers – have come together to design and jointly endorse a concise and measurable set of principles that we refer to as the FAIR Data Principles" [1]. While advocacy by more 'stakeholders' for making data findable, accessible, interoperable and reusable represented progress towards the goal of reproducible science, their use of the term 'design' gave the impression that these 'stakeholders' originated these principles. However, all of the FAIR principles by Wilkinson *et al.*, which were submitted 2015 and published 2016 in *Scientific Data* [1], simplified and paraphrased those previously introduced and expounded by Taswell in an article originally submitted in 2006 and published in *IEEE Transactions on Information Technology in Biomedicine* [2], [3] first online in 2007 and then print in 2008. The PORTAL-DOORS Project (PDP) and cyberinfrastructure system described in that article continues to serve as a distributed network of repositories with a message exchange format specification for managing data from and metadata about resources, but the PDP paper also defined and described the more general principles of data and metadata sharing for which that software engineering specification was designed. In this current report, we perform an item-by-item comparison and analysis of the principles as stated by Wilkinson *et al.* in 2015 and the corresponding principles as first articulated by Taswell in the original manuscript written in 2006. In addition, we compare the differences between their use and our use of the acronym 'FAIR', the term 'metric' and the phrase 'FAIR metrics' for the purpose of measuring fair standards of citation in scholarly research and publishing.

*Corresponding author (ctaswell@bhavi.us, ctaswell@ucsd.edu) also affiliated with UCSD Department of Electrical and Computer Engineering.

II. History of PDP and NPDS

As early as 1945, Vannevar Bush described the need for automated systems to help scientists sift through the ever-growing body of existing research [4]. Even before coinage of the phrase 'semantic web' and development of an ecosystem of documents engineered for processing by artificial intelligence with smart software systems empowered with logical

DOI: 10.1109/ECAI46879.2019.9042003

reasoning, the original concept of hypertext as proposed by Ted Nelson included semantic links between documents or parts of documents [5]. Taswell's work on the ManRay web-enabled ontology for nuclear medicine appeared in the *Journal of Nuclear Medicine* in May 2006 [6] and continued on PDP until submission of the manuscript in October 2006 [2]. This architectural design paper on PDP [2] cited the work done by Berners-Lee *et al.* on the semantic web in general [7], [8], but was unaware of his proposal on linked data principles that same year [9]. The original paper on PDP presented the separation of concerns between Problem-Oriented Registry of Tags and Labels (PORTAL) registries of unique identifiers with associated lexical metadata for resources and Domain-Ontology Oriented Resource System (DOORS) directories of locations with associated semantic descriptions for those same corresponding resources [2], [3]. This separation of concerns was inspired by analogy with that previously implemented between the Internet Registry Information Service and Domain Name System [2, Table I, page 194]. The literature review in sections II through VI of the original PDP paper [2] provided a more detailed review of prior work in the areas of data interoperability and metadata management circa 2006.

In contrast to the generally defined simple linked data principles of Berners-Lee from 2006 [9], the PDP principles of Taswell from 2006 [2] provided a much more technologically detailed, specific and comprehensive strategy for making resources with descriptions findable, accessible, interoperable, reusable and redistributable. The original architectural design paper for PDP published in *IEEE Transactions on Information Technology in Biomedicine* [2], [3] (also a USPTO patent [10]) described further a message exchange format for managing and sharing both lexical metadata with identifiers organized in PORTAL records, and semantic metadata with descriptions organized in DOORS records, along with location metadata for both offline physical resources and online virtual resources of any kind including data, metadata, and computing tools and services. Numerous conference presentations [11]–[21], a book chapter [22] and a journal article published in *Future Internet* [23] (also a USPTO patent [24]) detailed elaborations of the original design that included Nexus records, which could contain fields from both PORTAL and DOORS records for use cases where both lexical and semantic metadata are available. These principles for PDP expounded in the architectural design papers [2], [3], [10], [23], [24] described use of *metadata about metadata* and the hierarchically distributed mobile metadata (HDMM) architectural style for pervasive networks of data sharing [14], [23], [24], and provided a much more comprehensive systematic and detailed design and specification that encompassed all of the four basic considerations described in the linked data (LD) principles [9].

LD1: "Use URIs as names for things" [9]. Resources registered in Nexus-PORTAL-DOORS repositories can be identified by unique URIs called *resource labels* in [2] renamed *entity labels* in [23] with "a single required *canonical label* and multiple permitted *alias labels* for [each] resource".

LD2: "Use HTTP URIs so that people can look up those names" [9]. Resources registered in Nexus-PORTAL-DOORS repositories can be found by URLs called *resource locations* in [2] renamed *entity locations* in [23].

LD3: "When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)" [9]. Resources registered in Nexus-PORTAL-DOORS repositories can be described by lexical metadata called *other metadata* and semantic metadata called *resource descriptions* in [2] renamed *entity descriptions* in [23], defined as "an RDF mini-document, a collection of RDF triples that reference OWL ontologies, enabling semantic reasoning queries of descriptions" [2].

LD4: "Include links to other URIs. so that they can discover more things" [9]. Resources registered in Nexus-PORTAL-DOORS repositories can reference other different resources linked in the *resource descriptions* in [2] renamed *entity descriptions* in [23], as well as any other related resources called *resource cross-references* in [2] renamed *entity cross-references* in [23]. Resource records can also contain *supporting tags* with words or phrases [2], and *supporting labels* with URIs [23] for terms in controlled vocabularies [20], [21].

In the intervening years, authors affiliated with PDP have developed and tested software implementations of the design for the cyberinfrastructure system as well as ontologies for use with it [22], [25]–[29]. A paper presented at the SWAT4LS 2016 Conference introduced an improved read-only Nexus-PORTAL-DOORS web service with a RESTful API for accessing records and a separate read-write Scribe web service with a RESTful API for creating, updating, and deleting records, designed and implemented purposefully with a separation of concerns to minimize the chances of URL collisions and/or accidental corruption of records by different applications meant for use separately with read-only access versus read-write access [30]. The subsequent SWAT4LS 2017 Conference saw the demonstration of message exchange between a primary Nexus server implemented on the Microsoft web solution stack with a backend SQL Server database and a secondary, caching Nexus server implemented with a JavaScript and Node.js solution stack with records stored in a backend MongoDB database [31]. For our work with PDP, we now call the implementation of our Nexus-PORTAL-DOORS-Scribe cyberinfrastructure the NPDS system.

III. PDP, NPDS, and DREAM Principles

Both Taswell in 2006 [2] and Wilkinson *et al.* in 2015 [1] cited as their motivation the need for repositories of semantic descriptions and metadata accessible to both humans and automated agents. Compare the statement "The Domain Ontology Oriented Resource System (DOORS) and Problem Oriented Registry of Tags and Labels (PORTAL) are proposed as infrastructure systems for resource metadata within a paradigm that can serve as a bridge between the original web and the semantic web"[1] [2, abstract] with the statement "There is an urgent need to improve the infrastructure supporting the

[1]Quotes reproduce the original text including parentheses and italics.

DOI: 10.1109/ECAI46879.2019.9042003

reuse of scholarly data" and "Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals" [1, abstract].

Both also identified low barriers to entry as a key design consideration: "Minimizing requirements remains imperative during the transition from original web to semantic web" [2, Sec. VII-A, par. 1] and "By minimally defining each guiding principle, the barrier-to-entry for data producers, publishers and stewards who wish to make their data holdings FAIR is purposely maintained as low as possible" [1, Sec. III, par. 2]. The original paper by Taswell not only described in detail the message exchange specification of the PORTAL-DOORS System and BioPort, an example of a specific PORTAL registry service but also outlined the principles according to which they were designed: "Section VII provides a detailed exposition of the design principles and requirements necessary for both DOORS and PORTAL server functions and data records to operate as an effective infrastructure for registering resource labels and tags and publishing resource locations and descriptions intended for use by other semantic systems and applications. Similarly, Section VIII provides a description of the design principles and requirements for BioPORT as a registry for biomedical computing within the PORTAL-DOORS framework" [2, Sec. I, par. 3]. "Extensions of IRIS and analogs of DNS can also be developed for the semantic web and grid with a focus on *labeled resources* instead of *named domains*. Thus, basic principles and requirements for data records and server functions are proposed here for a new infrastructure technology as an extension and analog of the existing IRIS-DNS framework" [2, Sec. VII, par. 1], and also recommended the readily available web protocol stating that "PORTAL and DOORS could each be implemented as web services over `http`" [2, Sec. VII-E, par. 1].

Throughout [2], Taswell referred to the content of PORTAL and DOORS records as 'metadata' describing 'resources' in the sense of anything that can be identified via a URI or IRI [32], [33]. Since "these identifiers may specify either abstract or physical resources, neither of which are required to be accessible via the Internet" [2, Sec. VI, par. 1], this concept of resources includes data repositories and data sets, computing tools and computing services, and different types of media as has always been listed in the collection of *entity types* since the beginning of PDP in 2006. Wilkinson *et al.* use the terms 'data' and 'metadata' and the combined term '(meta)data' when applying a principle to both: "Throughout the Principles, we use the phrase '(meta)data' in cases where the Principle should be applied to both metadata and data" [1, Sec. III, par. 1]. However, in reference to non-data research objects, such as analytical workflows, they note that the "FAIR principles can equally be applied to these non-data assets, which need to be identified, described, discovered, and reused in much the same manner as data" [1]. From these passages, it follows that the scope of 'resources' in [2] and of 'data' in [1] are the same. With the exception of the particular use of the specific word 'license' by the Wilkinson *et al.* authors [1], which can be inferred obviously from the required metadata fields for *resource owner*, *record provenance* and *record distribution* of Taswell [2], the original paper [2] introduced, addressed and described each of the design principles simplified and paraphrased nine years later in [1]. These original NPDS principles first expounded by Taswell [2] identified and clarified all of the same principles that were later paraphrased by Wilkinson *et al.* [1], but Taswell [2] described a more versatile and flexible approach by specifying each principle as 'required' or 'permitted', while also avoiding the imposition of unnecessary restrictions and allowing for the possibility of derivations by problem-oriented and domain-oriented scientific communities who create their own customized extensions of NPDS.

Findable F1: "(meta)data are assigned a globally unique and persistent identifier" [1]. "the PORTAL operates as a resource label and tag registering system... and the DOORS operates as a resource location and description publishing system" [2, Sec. VII, par. 1]. Findable F2: "data are described with rich metadata (defined by R1 below)" [1]. The DOORS specification permits inclusion of "the *resource description* with an RDF mini-document, a collection of RDF triples that reference OWL ontologies, enabling semantic reasoning queries of descriptions" [2, Sec. VII-A, par. 1, itm. 6], while the PORTAL specification permits inclusion of "any *other metadata* permitted by the policies of the specific PORTAL registry type" [2, Sec. VII-B, par. 1, itm. 7]. Findable F3: "metadata clearly and explicitly include the identifier of the data it describes" [1]. Both PORTAL and DOORS each require "the *resource label* with a globally unique URI (or IRI) enabling nonsemantic string queries of labels" [2, Sec. VII-A, par. 1, itm. 1 and Sec. VII-B, par. 1, itm. 1]. Findable F4: "(meta)data are registered or indexed in a searchable resource" [1]. "The label and tags can be searched with a nonsemantic string query while the description can be searched with a semantic reasoning query" [2, Sec. VII-A, par. 3].

Accessible A1: "(meta)data are retrievable by their identifier using a standardized communications protocol" [1]. DOORS directories "Map label to location: Perform a lookup for a resource labeled uniquely by URI (or IRI) and return the associated URLs (or IDNs) required to be resolvable Internet locations for: a) the primary site and any mirror sites for the resource itself (a mapping via the associated URLs from the URI label to the resource itself); b) the URI (or IRI) namespace directory containing associated metadata maintained by the resource owner with descriptions in RDDL (a more indirect mapping via the associated URLs from the URI label to the metadata at the namespace directory linking to the resource); or c) the contact information maintained by the governing PORTAL registry if neither the resource itself nor its URI namespace is maintained online by the resource owner (the most indirect mapping via the associated URL from the URI label to the metadata at the registry enabling contact with the owner of the offline resource)." [2, Sec. VII-C, par. 1, itm. 1]. Accessible A1.1: "the protocol is open, free, and universally implementable" [1]. "PORTAL and DOORS could each be

 DOI: 10.1109/ECAI46879.2019.9042003

implemented as web services over http." [2, Sec. VII-E, par. 1]. Accessible A1.2: "the protocol allows for an authentication and authorization procedure, where necessary" [1]. DOORS directories "provide identification and authentication: Include the provenance and signature of each resource record returned in the response to the lookup or query request." [2, Sec. VII-C, par. 1, itm. 1]. Accessible A2: "metadata are accessible, even when the data are no longer available" [1]. In NPDS, resources are not required to be online entities in order to have a Nexus, PORTAL or DOORS record associated with them [2, Sec. VII-C, par. 1, itm. 1].

Interoperable I1: "(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation" [1]. NPDS records have been serialized in XML or JSON over the http or https protocol, but can also be represented in other serializations if so desired [2]. Interoperable I2: "(meta)data use vocabularies that follow FAIR principles" [1]. "As a bootstrap, the PORTAL-DOORS framework eschews debates about formal ontologies versus informal folksonomies and microformats" and instead "creates a hybrid with labels (URIs and IRIs) and tags (key word and phrase strings) for the original web, and with descriptions" [2, Sec. XII, par. 1]. For an example of a controlled vocabulary demonstrated with NPDS, see integration of the Medical Subject Headings for use with NPDS [20], [21]. Interoperable I3: "(meta)data include qualified references to other (meta)data" [1]. The PORTAL specification permits "*resource cross-references* with any globally unique identifiers permitted by the policies of the specific PORTAL registry type for identification of the resource in other systems unrelated to PORTAL-DOORS" [2, Sec. VII-A, par. 1, itm. 5].

Reusable R1: "meta(data) are richly described with a plurality of accurate and relevant attributes" [1]. The DOORS specification requires the *record provenance* [2, Sec. VII-A, par. 1, itm. 3], the *record distribution* [2, Sec. VII-A, par. 1, itm. 4], permits the optional *record signature* [2, Sec. VII-A, par. 1, itm. 7]. These requirements are elaborated more explicitly in [23]: "*Record metadata*: All metadata pertaining to the stored records about the entity and the process of registering and managing the records including timestamps for creating and updating the records, references to the governing registries and directories, as well as references to the registrant and agents for the records; note that the registrant and agent for the records may be different from the owner and contact for the entity" Reusable R1.1: "(meta)data are released with a clear and accessible data usage license" [1]. The PORTAL specification requires "the *resource owner* with contact information for the personnel who own and manage the resource" [2] and the DOORS specification requires the *resource location*, *record provenance* and *record distribution* which combined together serve the same purpose. Reusable R1.2: "(meta)data are associated with detailed provenance" [1]. The DOORS specification requires "the *record provenance* with identification of the: a) resource owner; b) authoritative master PORTAL registry; and c) authoritative primary DOORS server" [2, Sec. VII-A, par. 1, itm. 3]. Reusable R1.3: "(meta)data meet domain-relevant community standards" [1]. "Metadata items listed in Sections VII-A and VII-B are considered required or permitted with respect to the generic PORTAL registry type, not with respect to a semantic domain-specific PORTAL registry type (see Fig. 2). Thus, the schema imposed by the PORTAL root server (for the generic type) is least restrictive while a schema imposed by a PORTAL master server (for a specific type) may be more restrictive. An item considered permitted with respect to the generic PORTAL registry type may be considered required with respect to a specific PORTAL registry type if declared by its policies. Distinct registry types serving different semantic domains of inquiry may have very different policies regarding the manner in which unique labels and optional tags are created for each resource when registered" [2, Sec. VII-B, par. 2].

For a more extensive analysis with a detailed listing that compares the Wilkinson *et al.* FAIR principles with the Taswell NPDS principles, see Appendix Tables III, IV, V and VI. As a result of this item-by-item comparison and analysis, we cannot find any novel idea or concept in [1], [34] that can be described as fundamentally new and/or different from the content, principles, analysis and discussion in [2], [3], [10], [23], [24]. Moreover, the authors of [1], [34] did not cite any of the work previously published in [2], [3], [10], [23], [24] and online at www.portaldoors.org even though key authors of [1], [34] were aware of it by virtue of being informed about it by direct face-to-face communication in person by Taswell at conferences attended mutually by them and Taswell as early as 2009 at a W3C F2F Meeting [13]. In order to address these problems in the literature, we are renaming the NPDS principles previously published in [2], [10], [23], [24] as the DREAM principles with the acronym DREAM for *Discoverable Data with Reproducible Results for Equivalent Entities with Accessible Attributes and Manageable Metadata* as the comprehensively summarizing phrase to describe collectively all of the PDP and NPDS principles of Taswell [2], [10], [23], [24] that have been *unfairly* renamed the FAIR principles by Wilkinson *et al.* [1], [34].

IV. FAIR Metrics

Whereas the 'FAIR principles' have not been published *fairly* by the Wilkinson *et al.* [1] authors, the 'FAIR metrics' published by the Wilkinson *et al.* [34] authors should not be considered, called or termed 'metrics' in scientific research. In mathematics, statistics and the sciences, a 'metric' is usually defined as a quantitative numerical value corresponding to a measure of something, often within a specified interval of time or at a point in time, and/or within a specified region of space or at a point in space. Common lists of metrics used in theoretical, computational and experimental sciences generally do not include simple binary indicators of *true versus false* or *present versus absent*. These simple binary indicators as used by Wilkinson *et al.* [34] do not fulfill the criteria of non-negativity, identity, symmetry and sub-additivity that are usually required for mathematical and statistical metrics in scientific research. Therefore, the usage of the term 'metric'

DOI: 10.1109/ECAI46879.2019.9042003

Table I
NOTATION FOR FAIR METRICS

Symbol	Definition
C	set C of statements in a Control paper or in a Comparison Collection of papers
$F_1(T\|C)$	FAIR metric scaled on interval $[0, 1]$ for T compared with C (dependent on Q and S)
$F_2(T\|C)$	FAIR metric scaled on interval $[-1, +1]$ for T compared with C (dependent on Q, M and S)
$F_3(T\|C)$	FAIR metric scaled on interval $[-1, +1]$ for T compared with C (dependent on Q, P and S)
$F_4(T\|C)$	FAIR metric scaled on interval $[-1, +1]$ for T compared with C (dependent on Q, N and R)
$G(A)$	function G operates on set A
$G(A\|B)$	function G operates on set A in comparison with set B (on set A given set B)
$M(T\|C)$	number M of Misquoted (incorrectly cited) statements found in T compared with C
$N(T\|C)$	number N of Novel (uncited) statements found in T compared with C
$K(C)$	number K of Known statements found in C
$P(T\|C)$	number P of Plagiarized (uncited) statements found in T compared with C
$Q(T\|C)$	number Q of Quoted (correctly cited) statements found in T compared with C
$R(T\|C)$	number R of Reported statements found in T compared with C
$S(T\|C)$	number S of Similar statements found in T compared with C
T	set T of statements in a Test paper

Table II
FORMULAS FOR FAIR METRICS WITH CONDITION
$0 < S(T|C) \leq K(C) \leq R(T|C)$

Symbol		Formula
$F_1(T\|C)$	$=$	$Q(T\|C)/S(T\|C)$
$F_2(T\|C)$	$=$	$[Q(T\|C) - M(T\|C)]/S(T\|C)$
$F_3(T\|C)$	$=$	$[Q(T\|C) - P(T\|C)]/S(T\|C)$
$F_4(T\|C)$	$=$	$[Q(T\|C) - N(T\|C)]/R(T\|C)$
$S(T\|C)$	$=$	$M(T\|C) + Q(T\|C) + P(T\|C) \leq K(C)$
$R(T\|C)$	$=$	$M(T\|C) + Q(T\|C) + P(T\|C) + N(T\|C) \geq K(C)$

by the Wilkinson *et al.* [34] authors is not consistent with its usage in most fields of science. In order to address this problem in the literature with the misuse of the phrases 'FAIR principles' and 'FAIR metrics' by the Wilkinson *et al.* [1], [34] authors, we have proposed and published an alternative interpretation of the acronym 'FAIR' with our FAIR family of truly quantitative numerical metrics for maintaining *fair* standards in scholarly research and publishing [35], [36]. We defined and continue to use 'FAIR' as an acronym for the *Fair Acknowledgment of Information Records* and *Fair Attribution to Indexed Reports* [35], [36]. We continue the work on our FAIR metrics with a detailed mathematical exposition of the definitions, formulas, and simulated examples that can be found in [37] and that are also summarized here in Tables I and II. Our definitions and formulas allow for signed distance functions, also known as oriented distance functions. Therefore, the metrics in our FAIR family should be considered generalized metrics for the signed metrics that do not meet the non-negativity criteria of a classical unsigned metric. Currently, we are conducting experiments to validate our FAIR metrics on a collection of known plagiarized and plagiarizing papers. We plan to enhance the simple formulas for our FAIR metrics as summarized in Table II with additional weighting factors to account not only for commonality of author citations in the reference lists for published articles, but also for commonality of author attendance at conferences as inferred from authorship of articles published at conferences. We will use artificial intelligence with machine learning to infer values for these weighting factors derived from the problem-oriented collections of scientific literature appropriate for each community of authors publishing in a particular field of scholarly research.

V. CONCLUSION

Our PDP and NPDS principles that we now call the DREAM principles were originally published as the foundation for the PORTAL-DOORS Project and the Nexus-PORTAL-DOORS-Scribe cyberinfrastructure [2], [3], [10], [14], [23], [24]. Moreover, these DREAM principles of Taswell [2], [23] not only encompass the linked data principles of Berners-Lee [9] and the FAIR data principles of Wilkinson *et al.* [1], but were historically either contemporaneous with them [9] or preceded them [1]. Furthermore, the NPDS message exchange specification and web service APIs [28], [30], [31] that we have published for the DREAM principles address many of the software engineering challenges left unanswered by both [9] and [1]. As noted in a letter to IEEE Computer Magazine in 2010 by Taswell, "any discussion of provenance and reproducibility for computational science and engineering that does not also address citation and attribution leads to a contradiction in terms. It is not possible to maintain standards for scholarly peer-reviewed reproducible science without proper citation and attribution" [38]. This principle remains paramount when the professed goal has been "to improve the infrastructure supporting the reuse of scholarly data" as claimed by Wilkinson *et al.* [1], [34], but apparently not practiced by them with respect to citing *fairly* other authors such as Taswell [2], [23]. We believe that the Nexus-PORTAL-DOORS-Scribe cyberinfrastructure system, the DREAM principles and the FAIR metrics, as we have described and defined them in our published work, will serve as valuable software tools, web service applications and a distributed network system of data repositories for domain experts who wish to increase the impact of their scholarly research by making it discoverable on both the semantic web and the lexical web in a manner that also adheres to *fair* standards of citation in

DOI: 10.1109/ECAI46879.2019.9042003

scholarly research and publishing. Finally, we plan to require use of our FAIR metrics for the analysis of all manuscripts submitted for peer and editorial review prior to acceptance and approval for publication at our online open access *Brainiacs Journal of Brain Imaging and Computing Sciences* which will be available at www.BrainiacsJournal.org.

References

[1] M. D. Wilkinson, M. Dumontier *et al.*, "The FAIR guiding principles for scientific data management and stewardship," *Scientific Data*, vol. 3, p. 160018, Mar. 2016.

[2] C. Taswell, "DOORS to the semantic web and grid with a PORTAL for biomedical computing," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 2, pp. 191–204, Mar. 2008, in the Special Section on Bio-Grid.

[3] ——, "Corrections to "DOORS to the Semantic Web and Grid With a PORTAL for Biomedical Computing"," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 3, p. 411, Mar. 2008.

[4] V. Bush *et al.*, "As we may think," *The Atlantic Monthly*, vol. 176, no. 1, pp. 101–108, 1945.

[5] J. B. Smith and S. F. Weiss, "Hypertext," *Communications of the ACM*, vol. 31, no. 7, pp. 816–819, Jul. 1988.

[6] C. Taswell, B. Franc, and R. Hawkins, "The ManRay project: Initial development of a web-enabled ontology for nuclear medicine," in *Proceedings of the 53rd Annual Meeting of the Society of Nuclear Medicine, San Diego, CA*, vol. 47, no. suppl 1. Soc Nuclear Med, Jun. 2006, p. 1431.

[7] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, vol. 284, no. 5, pp. 34–43, May 2001.

[8] T. Berners-Lee, W. Hall, J. Hendler, N. Shadbolt, and D. J. Weitzner, "Computer science: Creating a science of the web," *Science*, vol. 313, no. 5788, pp. 769–771, Aug. 2006.

[9] T. Berners-Lee, "Linked data," 2006. [Online]. Available: https://www.w3.org/DesignIssues/LinkedData.html

[10] C. Taswell, "PORTALS and DOORS for the semantic web and grid," patentus 7,792,836, Sep., 2010.

[11] ——, "PORTAL-DOORS infrastructure system for translational biomedical informatics on the semantic web and grid," in *Proceedings of the American Medical Informatics Association Summit on Translational Bioinformatics, San Francisco, CA*, Mar. 2008, p. 43.

[12] ——, "Application of the PORTAL-DOORS system for use by clinical trials registries," in *Proceedings of the American Medical Informatics Association Summit on Translational Bioinformatics, San Francisco, CA*, Mar. 2009.

[13] ——, "Biomedical informatics for brain imaging and gene-brain-behavior relationships," in *W3C Semantic Web HCLSIG F2F Meeting 4/30-5/1*, 2009. [Online]. Available: http://esw.w3.org/topic/HCLSIG/Meetings/2009-04-30_F2F

[14] ——, "The hierarchically distributed mobile metadata (HDMM) style of architecture for pervasive metadata networks," in *Proceedings of ISPAN 2009 The 10th International Symposium on Pervasive Systems, Algorithms and Networks*. IEEE, Dec. 2009, pp. 315–320.

[15] ——, "Implementation of prototype biomedical registries for PORTAL-DOORS," in *Proceedings of the American Medical Informatics Association Summit on Translational Bioinformatics, San Francisco, CA*, Mar. 2009, pp. AMIA–0036–T2009.

[16] ——, "Knowledge engineering for PharmacoGenomic Molecular Imaging of the brain," in *2009 Fifth International Conference on Semantics, Knowledge and Grid*. Institute of Electrical and Electronics Engineers (IEEE), Sep. 2009, pp. 26–33.

[17] ——, "A new PDS PORTAL for clinical telegaming rehabilitation and intervention," in *Proceedings of 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM10)*, 2010.

[18] ——, "Concept validating methods for maintaining the integrity of problem oriented domains in the PORTAL-DOORS system," in *IDAMAP 2010: Intelligent Data Analysis in Biomedicine and Pharmacology*, S. Swift and K. Phillips, Eds., 2010.

[19] ——, "The ManRay project in biomedical informatics for nuclear medicine and pharmacogenomic molecular imaging," in *Proceedings 35th Annual Western Regional Meeting of the Society of Nuclear Medicine*, 2010.

[20] ——, "Use of the MeSH Thesaurus in the PORTAL-DOORS System," in *Proceedings AMIA 2010 Symposium on Clinical Research Informatics, San Francisco CA*, 2010, pp. AMIA–033–C2010.

[21] ——, "Use of NLM medical subject headings with the MeSH2010 thesaurus in the PORTAL-DOORS system," in *Proceedings of 8th HealthGrid 2010 Paris*, ser. Studies in Health Technology and Informatics, T. Solominides, I. Blanquer, V. Breton, T. Glatard, and Y. Legre, Eds., vol. 159. IOS Press, 2010, pp. 255–258.

[22] ——, "Alternative bootstrapping design for the PORTAL-DOORS cyberinfrastructure with self-referencing and self-describing features," in *Semantic Web*, G. Wu, Ed. Vukovar, Croatia: In-Teh, 2009, ch. 2, pp. 29–37. [Online]. Available: http://sciyo.com/books/show/title/semantic-web

[23] ——, "A distributed infrastructure for metadata about metadata: The HDMM architectural style and PORTAL-DOORS system," *Future Internet*, vol. 2, no. 2, pp. 156–189, 2010, in Special Issue on Metadata and Markup. [Online]. Available: http://www.mdpi.com/1999-5903/2/2/156/

[24] ——, "Management of multilevel metadata in the PORTAL-DOORS system with bootstrapping," patentus 8,886,628, Nov., 2014.

[25] D. Lockery, J. F. Peters, and C. Taswell, "CTGaming: A problem-oriented registry for clinical telegaming rehabilitation and intervention," *Journal of Emerging Technologies in Web Intelligence*, vol. 3, no. 1, pp. 28–37, Feb. 2011.

[26] M. Skarzynski, A. Craig, and C. Taswell, "SOLOMON: An ontology for Sensory-Onset, Language-Onset and Motor-Onset dementias," in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE. Institute of Electrical and Electronics Engineers (IEEE), Nov. 2015, pp. 969–972.

[27] S. H. Bae, A. G. Craig, and C. Taswell, "Expanding Nexus diristries of dementia literature with the NPDS concept-validating search engine agent," in *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Jul. 2017, pp. 3704–3707.

[28] A. G. Craig, S.-H. Bae, and C. Taswell, "Bridging the semantic and lexical webs: Concept-validating and hypothesis-exploring ontologies for the Nexus-PORTAL-DOORS System," *Journal of Systemics, Cybernetics and Informatics*, vol. 15, no. 5, pp. 8–13, 2017.

[29] S. Dutta and C. Taswell, "SPARQL-based search engine and agent for finding brain literature and converting references to NPDS metadata records," in *Proceedings of the 11th International Conference on Brain Informatics*, no. B277, 2018.

[30] A. Craig, S. H. Bae, T. Veeramacheneni, S. K. Taswell, and C. Taswell, "Web service APIs for Scribe registrars, Nexus diristries, PORTAL registries and DOORS directories in the NPD system," in *Proceedings of the 9th International SWAT4LS Conference*, 2016. [Online]. Available: http://ceur-ws.org/Vol-1795/paper4.pdf

[31] A. Craig, S.-H. Bae, and C. Taswell, "Message exchange between independent implementations of servers in the Nexus-PORTAL-DOORS System," in *Proceedings of the 10th International SWAT4LS Conference*, 2017. [Online]. Available: http://ceur-ws.org/Vol-2042/paper6.pdf

[32] T. Berners-Lee, R. Fielding, and L. Masinter, "STD66 RFC3986: Uniform resource identifier (URI): Generic syntax," Full Standard RFC [Online], IETF, RFC 3986, Jan. 2005. [Online]. Available: http://www.ietf.org/rfc/rfc3986.txt

[33] M. Dürst and M. Suignard, "RFC3987: Internationalized resource identifiers (IRIs)," IETF RFC 3987 Working Draft, IETF, RFC 3987, Jan. 2005. [Online]. Available: http://www.ietf.org/rfc/rfc3987.txt

[34] M. D. Wilkinson, S.-A. Sansone, E. Schultes, P. Doorn, L. O. B. da Silva Santos, and M. Dumontier, "A design framework and exemplar metrics for FAIRness," *Scientific Data*, vol. 5, p. 180118, Jun. 2018.

[35] A. Craig and C. Taswell, "The FAIR metrics of adherence to citation best practices," in *Proceedings of ASIS&T 81st Annual Meeting SIGMET Workshop*, 2018.

[36] ——, "Formulation of FAIR metrics for primary research articles," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, Dec. 2018.

[37] A. Craig, A. Ambati, S. Dutta, A. Mehrotra, S. K. Taswell, and C. Taswell, "Definitions, formulas, and simulated examples for plagiarism detection with FAIR metrics," in *Proceedings of ASIS&T 82nd Annual Meeting*, 2019, in press.

[38] C. Taswell, "Provenance and reproducibility," *Computer*, vol. 43, no. 12, p. 8, 2010.

 DOI: 10.1109/ECAI46879.2019.9042003

APPENDIX

The Appendix includes Tables III, IV, V and VI with a more detailed listing that compares the Wilkinson *et al.* FAIR principles with the Taswell NPDS principles. The Wilkinson *et al.* FAIR principles [1], [34] paraphrase a simplified subset of the Taswell NPDS principles [2], [3], [10] which also include and support use of *metadata about metadata* and the hierarchically distributed mobile metadata (HDMM) architectural style for pervasive networks of data sharing [14], [23], [24].

Table III
WILKINSON *et al.* FAIR PRINCIPLES, TAGGED F, PARAPHRASED FROM TASWELL NPDS PRINCIPLES WITHOUT CITATION

Wilkinson *et al.* FAIR quotes	Taswell NPDS quotes
"F1. (meta)data are assigned a globally unique and persistent identifier"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "the PORTAL operates as a resource label and tag registering system (i.e., IRIS extension) and the DOORS operates as a resource location and description publishing system (i.e., DNS analog)" [2, pag. 194, Sec. VII, par. 1]; "the *resource label* with a globally unique URI (or IRI) enabling nonsemantic string queries of labels" [2, pag. 195, Sec. VII-A, par. 1, itm. 1]; "the *resource tags*, if registered at the governing registry, including a tokenized name and/or phrases enabling nonsemantic string queries of tags" [2, pag. 195, Sec. VII-A, par. 2, itm. 5]; "These resources may be assigned unique labels with associated tags consisting of one or more trade or service marks, and with associated descriptions referencing ontologies for patent and trademark classes and the semantic definitions for entities within those classes" [2, pag. 199, Sec. IX, par. 2]. [23, Tab. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6]; "Hierarchical authorities (root, primary, secondary, forwarding, caching) and globally unique identifiers to prevent namespace conflicts when identifying resources while maintaining autonomy of local communities with control over local policies" [23, pag. 167, Sec. 5.1, par. 2].
"F2. data are described with rich metadata (defined by R1 below)"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "the *resource description* with an RDF mini-document, a collection of RDF triples that reference OWL ontologies, enabling semantic reasoning queries of descriptions" [2, pag. 195, Sec. VII-A, par. 2, itm. 6]; "The label and tags can be searched with a nonsemantic string query while the description can be searched with a semantic reasoning query" [2, pag. 195, Sec. VII-A, par. 2]; "any *other metadata* permitted by the policies of the specific PORTAL registry type" [2, pag. 196, Sec. VII-B, par. 1, itm. 7]; [23, Tab. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6]; "Pervasively distributed and shared infrastructure, content, and control of content including distributed and shared control over both the contribution and distribution of the content defined as the mobile metadata records" [23, pag. 163, Sec. 4, par. 4, itm. 1].
"F3. metadata clearly and explicitly include the identifier of the data it describes"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "the *resource label* with a globally unique URI (or IRI) enabling nonsemantic string queries of labels" [2, pag. 195, Sec. VII-A, par. 1, itm. 1]; "the *resource label* with a globally unique URI (or IRI) required by the generic PORTAL registry type for identification of the resource in PORTAL-DOORS" [2, pag. 196, Sec. VII-B, par. 1, itm. 1]; "the *resource owner* with contact information for the personnel who own and manage the resource" [2, pag. 196, Sec. VII-B, par. 1, itm. 2]; "the *owner signature* with the XML-Signature of the owner permitted by the generic PORTAL registry type" [2, pag. 196, Sec. VII-B, par. 2, itm. 6]. [23, Tab. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6] "A relative freedom of choice in the selection of identifiers with purposeful absence of any requirement to use the *same* root name or label for *all* identifiers, thus enabling essentially unrestricted choice of naming or labeling schemes for identification and thereby avoiding monopolistic control by any single organization" [23, pag. 164, Sec. 4, par. 4, itm. 5].
"F4. (meta)data are registered or indexed in a searchable resource"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "The label and tags can be searched with a nonsemantic string query while the description can be searched with a semantic reasoning query" [2, pag. 195, Sec. VII-A, par. 3]; "*Search nonsemantic strings in labels or tags:* Find resources by string query of character substrings in labels or tags and return the associated URIs and URLs recognizing that the search may yield nonunique results when performed across resources governed by registries of different registry types or of the same registry type without a policy imposing at least one unique tag." [2, pag. 196, Sec. VII-C, par. 1, itm. 3]; "*Search semantic statements in descriptions:* Find resources by semantic query with SPARQL [72] of semantic statements in descriptions and return the associated URIs and URLs recognizing that the search may yield unranked nonunique results" [2, pag. 196, Sec. VII-C, par. 1, itm. 4]; "*Publish resource cross-references:* Perform a lookup of a registered resource by label or tag and return any cross-references identifying the resource in other systems unrelated to PORTAL-DOORS" [2, pag. 197, Sec. VII-D, par. 1, itm. 6]. [23, Tab. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6] "The PORTAL-DOORS System (PDS), collectively comprised of these interacting network systems of PORTAL registries and DOORS directories, has been architected as a distributed system for registering resource entities and publishing metadata about them" [23, pag. 157, Sec. 1, par. 1].

All Wilkinson *et al.* FAIR quotes from [1, pag. 4, Box 2].

DOI: 10.1109/ECAI46879.2019.9042003

Table IV
WILKINSON *et al.* FAIR PRINCIPLES, TAGGED A, PARAPHRASED FROM TASWELL NPDS PRINCIPLES WITHOUT CITATION

Wilkinson *et al.* FAIR quotes	Taswell NPDS quotes
"A1. (meta)data are retrievable by their identifier using a standardized communications protocol"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "*Map label to location:* Perform a lookup for a resource labeled uniquely by URI (or IRI) and return the associated URLs (or IDNs) required to be resolvable Internet locations for: a) the primary site and any mirror sites for the resource itself (a mapping via the associated URLs from the URI label to the resource itself); b) the URI (or IRI) namespace directory containing associated metadata maintained by the resource owner with descriptions in RDDL (a more indirect mapping via the associated URLs from the URI label to the metadata at the namespace directory linking to the resource); or c) the contact information maintained by the governing PORTAL registry if neither the resource itself nor its URI namespace is maintained online by the resource owner (the most indirect mapping via the associated URL from the URI label to the metadata at the registry enabling contact with the owner of the offline resource)" [2, pag. 196, Sec. VII-C, par. 1, itm. 1]; "The semantic web will succeed analogously when a similar dynamic synergism can be created between a resource label system with registries and label servers, all of the appropriately optimized communications protocols," [2, pag.200, Sec. X, par. 3] "*Publish resource cross-references:* Perform a lookup of a registered resource by label or tag and return any cross-references identifying the resource in other systems unrelated to PORTAL-DOORS." [2, pag. 197, Sec. VII-D, par. 1, itm. 6]; [23, Tab. 1, Fig. 2, Fig. 3, Fig. 4]; "any PDS implementation must maintain the important requirement of uniquely identifying resources by the resource entity label which must be an IRI or URI" [23, pag. 177, Sec. 6.1, par. 13]; "Metadata can be associated with each of the five objects listed above. The following list summarizes the metadata for each of the five objects together with the design principles that govern software implementation for the database, web service, and interoperable messaging interface schemas for the PORTAL-DOORS System." [23, pag. 170, Sec. 5.1, par. 8].
"A1.1. the protocol is open, free, and universally implementable"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "PORTAL and DOORS could each be implemented as web services over http." [2, pag. 198, Sec. VII-E, par. 1]. "DOORS could be implemented as an extension of either DNS or CRISP protocols since both have mechanisms enabling extensions. However, PORTAL should be implemented as an extension of the CRISP protocol because it lies so naturally within the scope of the stated goals for CRISP" [2, pag. 198, Sec. VII-E, par. 2]; "The PORTAL-DOORS paradigm favors a flexible and modular approach promoting collaborative networks of cross-linking resources" [2, pag. 201, Sec. XIII, par. 1] [23, Tab. 1, Fig. 2, Fig. 3]; "PORTAL-DOORS allows complete freedom with an identification scheme for which globally unique labels are simply required to be URIs" [23, pag. 164, Sec. 4, par. 8]; "thus enabling essentially unrestricted choice of naming or labeling schemes for identification and thereby avoiding monopolistic control by any single organization." [23, pag. 164, Sec. 4, par. 9].
"A1.2. the protocol allows for an authentication and authorization procedure, where necessary"	[2, Tab. II, Fig. 1, Fig. 2,]; "provide identification and authentication: Include the provenance and signature of each resource record returned in the response to the lookup or query request" [2, pag. 197, Sec. VII-C, par. 1, itm. 1]. [23, Fig. 2, Fig. 3]; "The reference implementation adopts the following conventions: In author mode, the agent may edit only records initially entered by the agent. In editor mode, the agent may edit any records in the same registry. In administrator mode, the agent may edit any registry or directory records accessible via the same registrar." [23, pag. 172, sec. 5.3, par. 5].
"A2. metadata are accessible, even when the data are no longer available"	[2, Tab. II]; "These identifiers may specify either abstract or physical resources, neither of which are required to be accessible via the Internet." [2, pag. 194, Sec. VI, par. 1]; "Each specialist should be able to conduct cross-directory searches in related fields and find any relevant resource of interest whether a simple spread- sheet macro or an ontology-based expert system, regardless of location of the directory or registry governing the data record found for the resource metadata." [2, pag. 193, Sec. IV-C, par. 5]. [23, Fig. 2, Fig. 3]; "this approach assures that all metaresources about the same targeted resource can refer consistently to that resource yet be managed independently of it as the primary resource and of each other as the other secondary resources." [23, pag. 172, Sec. 5.3, par. 5].

All Wilkinson *et al.* FAIR quotes from [1, pag. 4, Box 2].

DOI: 10.1109/ECAI46879.2019.9042003

Table V
WILKINSON *et al.* FAIR PRINCIPLES, TAGGED I, PARAPHRASED FROM TASWELL NPDS PRINCIPLES WITHOUT CITATION

Wilkinson *et al.* FAIR quotes	Taswell NPDS quotes
"I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "the implementation of the formal DOORS record as a valid XML document containing within itself a valid RDF mini-document for the semantic description" [2, pag. 195]; "If specifications for the DOORS and PORTAL systems are implemented as extensions of the CRISP framework, then they should be derived as XML schemas that depend upon the CRISP protocols" [2, pag. 198]; "the PORTAL-DOORS framework is built upon the XML/RDF/OWL foundations of the semantic web" [2, p. 201]; [23, Tab. 1, Fig. 2, Fig. 3]; "A hybridized architecture with both XML Schemas and terminologies serving the original web and also RDF triples and OWL ontologies serving the semantic web to bridge and transition from the original web to the semantic web" [23, p. 180].
"I2. (meta)data use vocabularies that follow FAIR principles"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "As a bootstrap, the PORTAL-DOORS framework eschews debates about formal ontologies versus informal folksonomies and microformats" and instead "creates a hybrid with labels (URIs and IRIs) and tags (key word and phrase strings) for the original web, and with descriptions" [2, pg.200] [2, Sec. XII, par. 1]; "Section VII provides a detailed exposition of the design principles and requirements necessary for both DOORS and PORTAL server functions and data records to operate as an effective infrastructure for registering resource labels and tags" [2, pg.191]; "basic principles and requirements for data records and server functions are proposed here for a new infrastructure technology as an extension and analog of the existing IRIS-DNS framework" [2, pg.194]; "SANs integrated with digital libraries would enable investigators to cross disciplines and search fields outside of their main area of expertise without being required to know in advance key words such as the phrase 'diffusion of innovations' that would have been relevant to Kazic's search of the literature" [2, pg.200]; [23, Tab. 1, Fig. 2, Fig. 3]; "Guided by architectural principles intended to address all of these major problems, PDS has been designed to operate as a hybrid between the original web and the semantic web with mechanisms that serve both the original web and the semantic web simultaneously as well as either one independently of the other." [23, pg.157]; "These HDMM principles do require hierarchical control and distribution of metadata records, but do not require hierarchical identification of resources." [23, pg.164]; "A hybridized architecture with both XML Schemas and terminologies serving the original web and also RDF triples and OWL ontologies serving the semantic web to bridge and transition from the original web to the semantic web." [23, pg.167]; "Version 0.7: Completion and revision of lexical PORTAL functionality including interoperability with terminology tools. Version 0.8: Completion and revision of semantic DOORS functionality including interoperability with ontology tools." [23, pg.182].
"I3. (meta)data include qualified references to other (meta)data"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "*resource cross-references* with any globally unique identifiers permitted by the policies of the specific PORTAL registry type for identification of the resource in other systems unrelated to PORTAL-DOORS" [2, p. 195] [2, Sec. VII-A, par. 1, itm. 5]; "Cross-references may be any URIs or IRIs stored optionally in BioPORT for identifying the resource in other systems." [2, p. 199]; "This flexibility should encourage the development of applications that exploit DOORS string searches on resource labels and tags in addition to DOORS semantic searches on resource descriptions while maintaining cross-links between resources in PORTAL-DOORS and cross-references to other systems" [2, p. 199]; "The PORTAL-DOORS paradigm favors a flexible and modular approach promoting collaborative networks of cross-linking resources and inter-referencing ontologies" [2, p.201]; [23, Tab. 1, Fig. 2, Fig. 3]; "The author has pursued a new approach distinguished by its goal of building a distributed shared infrastructure" [23, p. 161]; "In the case of PORTAL-DOORS, the server infrastructure, the content control, and the content itself are all shared and distributed" [23, p. 183].

All Wilkinson *et al.* FAIR quotes from [1, pag. 4, Box 2].

 DOI: 10.1109/ECAI46879.2019.9042003

Table VI
WILKINSON *et al.* FAIR PRINCIPLES, TAGGED R, PARAPHRASED FROM TASWELL NPDS PRINCIPLES WITHOUT CITATION

Wilkinson *et al.* FAIR quotes	Taswell NPDS quotes
"R1. meta(data) are richly described with a plurality of accurate and relevant attributes"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; the required *record provenance* [2, Sec. VII-A, par. 1, itm. 3], required *record distribution* [2, Sec. VII-A, par. 1, itm. 4], permitted *record signature* [2, Sec. VII-A, par. 1, itm. 7]; [23, Tab. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6]; "*Entity metadata:* All metadata pertaining to the entity itself including tags, labels, locations and description of the entity as well as references to the owner and contact for the entity; corresponds to PDS schema element `EntityMetadata` and considered primary or Level 1 metadata about the entity itself" [23, pag. 170, Sec. 5.2, par. 9, itm. 1]; "*Record metadata*: All metadata pertaining to the stored records about the entity and the process of registering and managing the records including timestamps for creating and updating the records, references to the governing registries and directories, as well as references to the registrant and agents for the records; note that the registrant and agent for the records may be different from the owner and contact for the entity; corresponds to PDS schema element `RecordMetadata` and considered secondary or Level 2 metadata about the Level 1 metadata" [23, pag. 170, Sec. 5-2, par. 3, itm. 2]; "*Infoset metadata:* All metadata pertaining to the dynamic infoset about the entity assembled from the distributed stored records including status, validation timestamps if validated, and any entailments if inferred by a reasoning engine; corresponds to PDS schema element `InfosetMetadata` and considered tertiary or Level 3 metadata about the Level 1 and Level 2 metadata" [23, pag. 170, Sec. 5-2, par. 3, itm. 3].
"R1.1. (meta)data are released with a clear and accessible data usage license"	[2, Tab. II, Fig. 1, Fig. 2]; "the *resource label* with a globally unique URI (or IRI) required by the generic PORTAL registry type for identification of the resource in PORTAL-DOORS" [2, Sec. VII-B, par. 1, itm. 1]; "the *resource owner* with contact information for the personnel who own and manage the resource" [2, Sec. VII-B, par. 1, itm. 2]; "the DOORS *servers* with URLs (or IDNs) for the primary and secondary DOORS servers that publish the metadata not maintained at the PORTAL registry" [2, Sec. VII-B, par. 1, itm. 3]; "the *owner signature* with the XML-Signature of the owner permitted by the generic PORTAL registry type" [2, Sec. VII-B, par. 1, itm. 6]; "the *resource owner* with contact information for the personnel who own and manage the resource" [2, pag. 196, Sec. VII-B, par. 1, itm. 2]; "the *resource label* with a globally unique URI (or IRI) required by the generic PORTAL registry type for identification of the resource in PORTAL-DOORS" [2, pag. 196, Sec. VII-B, par. 1, itm. 1]; "Perform other standard requests of registrar/registry systems [69] such as a lookup that returns all resources registered by an owner and a lookup that returns the contact information for the resource owner, managing personnel, or any other associated metadata." [2, pag. 198, Sec. VII-D, par. 1, itm. 7]; PORTAL record must include "the *resource owner* with contact information for the personnel who own and manage the resource" and the inclusion of this same resource owner in the required *record provenance* of a DOORS record, as quoted in R1.2 below, combined with contact information in the *resource location* field described in A.1 above, serve the same purpose in that they allow a user to determine the owner of the resource and contact the owner to request license information. [23, Tab. 1, Fig. 2, Fig. 3] "the term resource agent refers to the person who registers, manages and edits the information about the resource and who is presumed to be acting on behalf of the resource owner" [23, pag. 172, Sec. 5.3, par. 2, itm. 4].
"R1.2. (meta)data are associated with detailed provenance"	[2, Tab. II, Fig. 1, Fig. 2]; the required *record provenance* [2, Sec. VII-A, par. 1, itm. 3], required *record distribution* [2, Sec. VII-A, par. 1, itm. 4], permitted *record signature* [2, Sec. VII-A, par. 1, itm. 7]; "the *record provenance* with identification of the: a) resource owner; b) authoritative master PORTAL registry; and c) authoritative primary DOORS server" [2, pag. 195, Sec. VII-A, par. 1, itm. 3]; "Include the provenance and signature of each resource record returned in response to the lookup or query request" [2, pag. 197, Sec. VII-A, par. 1, itm. 3]. "the resource tags, if registered at the governing registry, including a tokenized name and/or phrases enabling nonsemantic string queries of tags" [2, pag. 195, Sec. VII-A, par. 1, itm. 5]. [23, Fig. 2, Fig. 3]
"R1.3. (meta)data meet domain-relevant community standards"	[2, Tab. I, Tab. II, Fig. 1, Fig. 2]; "Metadata items listed in Sections VII-A and VII-B are considered required or permitted with respect to the generic PORTAL registry type, not with respect to a semantic domain-specific PORTAL registry type (see Fig. 2). Thus, the schema imposed by the PORTAL root server (for the generic type) is least restrictive while a schema imposed by a PORTAL master server (for a specific type) may be more restrictive. An item considered permitted with respect to the generic PORTAL registry type may be considered required with respect to a specific PORTAL registry type if declared by its policies. Distinct registry types serving different semantic domains of inquiry may have very different policies regarding the manner in which unique labels and optional tags are created for each resource when registered" [2, pag. 196, Sec. VII-B, par. 2]. [23, Tab. 1, Fig. 2, Fig. 3]; "A distributed network of registries and directories for resource metadata oriented by problem domain or specialist community rather than by technology format of the resource" [23, pag. 167, Sec. 5.1, par. 3, itm. 1]; "Depending upon the problem-oriented specialty domain of the PORTAL registry and its registration policies, examples may include persons, patients, investigators, authors, or organizations; online virtual entities or offline physical entities; data services, data storage tools, and data records (independent of and unrelated to any PORTAL-DOORS metadata record); analysis services and data processing tools; authored information, books, journals, papers, web sites, and web pages; and many other examples and categories within any field of interest defined by the administrators of the particular PORTAL registry." [23, pag. 169, Sec. 5.2, par. 1, itm. 1].

All Wilkinson *et al.* FAIR quotes from [1, pag. 4, Box 2].

DOI: 10.1109/ECAI46879.2019.9042003