Check for updates

Achieving Transparency: A Metadata Perspective

Daniel Gillman*†

US Bureau of Labor Statistics, 2 Massachusetts Ave NE Washington District of Columbia 20212, United States

ABSTRACT

Transparency is vital to realizing the promise of evidenced-based policymaking, where "evidence-based" means including information as to what data mean and why they should be trusted. Transparency, in turn, requires that enough of this information is provided. Loosely speaking then, transparency is achieved when sufficient documentation is provided. Sufficiency is situation specific, both for the provider and consumer of the documentation. These ideas are presented in two recent US commissioned reports: *The Promise of Evidence-Based Policymaking* and *Transparency in Statistical Information for the National Center for Science and Engineering Statistics and All Federal Statistical Agencies*.

Metadata are a more formalized kind of documentation, and in this paper, we provide and demonstrate necessary, sufficient, and general conditions for achieving transparency from the metadata perspective: conforming to a specification, providing quality metadata, and creating a usable interface to the metadata. These conditions are important for any metadata system, but here the specification is tied to our framework for metadata quality based on the situation-specific needs for transparency. These ideas are described, and their interrelationships are explored.

1. INTRODUCTION

The term *transparency* appears in a lot of recent writing and initiatives. For example, Transparency International[①] is an effort to reduce corruption in governments around the world by exposing the need for

---

* The opinions in this article are due to the author and do not necessarily reflect the official policies of the US Bureau of Labor Statistics.

† Corresponding author: Daniel Gillman (Email: gillman.daniel@bls.gov; ORCID: 0000-0002-1638-3881). ① Transparency International—https://www.transparency.org/

accountability. The Office of the Director for National Intelligence[2] in the US promotes transparency of intelligence for the intelligence community to inform the public. Amazon[3] has a transparency initiative to reduce counterfeit products from reaching customers by using security and unique codes to identify those products.

Data are also the subject of transparency efforts. The focus of the work of the statistical agencies is data. Producing data is the purpose of these agencies, and much effort is expended towards ensuring all data have the highest possible quality[4]. With the publication of the report *The Promise of Evidence-Based Policymaking* by the US Commission on Evidence-Based Policymaking in September 2017 (Evidence Report) [1], a new push towards greater access to data has emerged in the US and elsewhere. Congress passed the Evidence Act in 2018, and this act codified much of what was included in the Evidence Report. Included in the Evidence Act were recommendations for greatly improving the way data are produced, described, disseminated, and managed. However, there remains much work to achieve the recommendations and adhere to the themes in the Evidence Report. Two of these themes are 1) to increase the emphasis in producing and maintaining metadata and 2) make data and methodologies behind those data more transparent. These themes are related, and we explore this in some detail.

Plans and efforts to make data transparent are currently discussed widely, and we understand the need and goal as generally desirable. A report sponsored by the US National Center for Science and Engineering Statistics (NCSES) and under the auspices of the US National Academies of Sciences, Engineering, and Medicine called *Transparency in Statistical Information for the National Center for Science and Engineering Statistics and All Federal Statistical Agencies* (Transparency Report) [2] was published in November 2021. This report addresses much of what is needed to make data and methodologies behind creation of the data transparent. Though the report focuses on the needs of statistical agencies, we feel many of the recommendations are sufficiently general as to apply to any organization producing data.

This paper reiterates some of what is in the Transparency Report and extends the ideas. The goal is to propose necessary and sufficient conditions for data or methodologies to be transparent. Metadata play a central role.

This paper is organized in the following way. First, we say what is meant by transparency. This moves into a discussion about metadata, schemas, and technical specifications. Conformance—satisfying all requirements in a technical specification—is a key idea. Also key are metadata quality, which is defined, and the usability of a system designed to support transparency. Throughout, we advance the argument as to why these criteria—conformance to a technical specification, metadata quality, and system usability—are necessary and sufficient to providing transparency.

---

[2] Director of National Intelligence—https://www.dni.gov/index.php/how-we-work/transparency

[3] Amazon—https://brandservices.amazon.com/transparency

[4] US Office of Management and Budget—https://www.whitehouse.gov/omb/information-regulatory-affairs/information-policy/

2. TRANSPARENCY

Dictionaries define transparency to mean "the condition of being easy to perceive or detect." Therefore, when it is easy to perceive or detect how to find, access, understand, and use data or methodologies they are transparent.

The Transparency Report [2] contains a slightly more useful description—

transparency is the provision of sufficiently detailed documentation of all the processes of producing official estimates.

This broad definition of transparency raises some important questions. In a specific circumstance, how is transparency insured? What documentation is sufficient and how is it delivered? Is it possible to detect when transparency is achieved through automated means or manual inspection?

Metadata are data in the role of describing other data, methodologies, or resources. Therefore, we know the answers will lie, in part, with the metadata that is available to a user. The ability to find, understand, and use data depends on the quality of the metadata available. And here the quality is a measure of how well the metadata provide all the information necessary to allow the user to complete the tasks set forth.

3. TRANSPARENCY AND METADATA

As defined above, transparency is a general notion, or concept. It is easy to define, but difficult to characterize generally. It is much easier to characterize when applied to specific circumstances or kind of application. Why is this the case? Is it possible to make sense of this reality?

Cognitive psychologists [3] have identified at least two kinds of concepts: entity and relational. Entity concepts are easier to characterize and more difficult to define. Relational concepts are the opposite. "Tennis ball" and "variable" are entity concepts. Each is fairly easy to characterize. The concept "guest" is an example of a relational, or role, concept. One can be a guest at a party, a guest in a hotel, or a guest user on a secure web site, for instance. Each kind is characterizable. However, these specific cases do not have much in common, so the generic concept of guest is not easily characterized. Role concepts must be refined, or specialized, to make them characterizable.

Transparency is a role concept. Consider the information needed to make a variable transparent to a user versus what is needed to make a data transformation process transparent. These are very different situations, and they require different characterizations.

In general, characterizing a concept provides a means to determine if a specific object or situation meets those criteria, i.e., if the object or situation *corresponds* to the concept. Specifically, the characteristics⑤ of

---

⑤ Characteristics differentiate concepts from each other.

a concept are the categories in which the properties© of objects that correspond to the concept belong. The properties of an object are descriptive of the object. References [4] and [5] discuss these ideas in detail.

To illustrate we will assume we can describe a variable to support some transparency needs with the following characteristics: name, defining concept, universe, question, datatype, and set of allowed values. See Table 1 below for a set of properties corresponding to the characteristics describing a marital status variable. Short descriptions of the characteristics are included to help those less familiar with describing variables. Each property corresponds to its characteristic in the following way: the characteristic takes the role of a question ("What are the allowed values for this variable?") and the property takes the role as the answer ("<1, single>, <2, married>, <3, separated>, <4, divorced>, and <5, widowed>").

Table 1. Describing a Marital Status Variable.

Characteristics	Description	Properties
Name	Name of the variable	MS01
Defining concept	Meaning of the variable	Legally defined marital state
Universe	Population the variable measures	Adults
Question	Question asked to capture data	What is …'s marital status? Choose one of the values.
Datatype	Computational aspects [6]	Nominal⑦
Allowed values	Valid values the variable may assign	<1, single>
		<2, married>
		<3, separated>
		<4, divorced>
		<5, widowed>

The properties in Table 1 are metadata, and their corresponding characteristics are elements of a schema for describing variables for that metadata. In all cases, each characteristic has a set of properties that correspond to it. These properties form the set of allowed values (or value domain) for the element in a schema resulting from the characteristic. So, if a particular concept is characterizable, these characteristics lead directly to a metadata schema. The elements and constraints of the schema turn the schema into a technical specification.

4. TECHNICAL SPECIFICATIONS

A technical specification is a set of expressions [7], which are of the following kinds:

- Statement: expression that conveys information
- Instruction: expression that conveys an action to be performed
- Recommendation: expression that conveys advice or guidance
- Requirement: expression that conveys criteria to be fulfilled

---

© Properties differentiate objects from each other.

⑦ The nominal datatype in statistics describes an unordered set of categories. The ISO/IEC 11404 equivalent datatype is state.

In our variable example, we can express the schema in tabular form as seen in Table 2.

Table 2. Schema for Describing Variables.

Characteristics	Rules
Name	Up to 16-character text
Defining concept	Term from a glossary entry
Universe	Set of possible units for observation
Question	Source of data from a questionnaire
Datatype	Nominal, Ordinal, Interval, Ratio, Date/Time, Lat/Log, Text
Allowed values	Either a list, a numeric range, a format®

This schema can also be turned into a set of expressions in natural language. These form the technical specification in Table 3 below.

Table 3. Technical Specification for Describing Variables.

Natural Language Expressions
A <u>name</u>, which shall be no more than 16 characters in length
A <u>defining concept</u>, which shall be either a link to an entry in an existing glossary or a term with well-known unambiguous meaning
A universe, a specialization of people, households, establishments, events (e.g., marriages, hiring), or outcomes (e.g., benefits, jobs) describing all possible units for observation
A question, which comes from the questionnaire used to collect data
A datatype, which shall be one of the following
o Nominal—unordered set of categories
o Ordinal—ordered set of categories
o Interval—numeric values where there is no meaningful zero (nothing)
■ E.g., Fahrenheit temperature scale
■ Expressed as a range
o Ratio—numeric values with a meaningful zero
■ E.g., Kelvin temperature scale
■ Expressed as range
o Date/Time—one of valid formats for date and time in ISO 8601
o Lat/Long—valid format for expressing latitude and longitude of a position on Earth
o Text—character string of no particular maximum length
The set of allowed values shall take one of the following forms:
o A list of categories and associated codes or terms
o A numeric range
o A specific format for date/time, latitude/longitude, or text

---

®   We recognize this list may not be inclusive. The author comes from the social science data community.

As the reader can see, we added more information to the technical specification in Table 3 than what was in Table 1. Though not complete, the additional formality allows us to determine whether the rules of the schema are obeyed by some application. By recognizing the kind of expression (statement, instruction, recommendation, or requirement) each rule exemplifies, it is possible to test (sometimes formally, but certainly informally) whether all the rules are adhered to.

5. CONFORMANCE

A metadata system (system, hereafter) comprises a database for metadata (or repository), update and retrieval capabilities, and a user interface. The system conforms to a technical specification if it satisfies all the requirements in that specification [7]. Sometimes requirements include other expressions that are not requirements themselves. For example, if there is a requirement that a particular algorithm be used, then the steps of that algorithm (expressions that are instructions) must be included as well.

Strict conformance means a conforming system satisfies all requirements and nothing more. Sometimes it is useful to extend a technical specification by adding new requirements, so a system conforming to the extended specification no longer strictly conforms to the original. Strict conformance is discussed in our metadata quality framework. If we consider the example of the marital status variable in Table 1 and the technical specification in Table 3, it is clear by inspection that the description of the marital status variable in Table 1 strictly conforms to the technical specification in Table 3. If we were to add a characteristic to Table 1, then it would just conform to the specification in Table 3.

Another consideration is based on the functionality of a conforming system. If the system can store a conforming instance of metadata, then that fulfills metadata instance level conformance. If the system can export, or write, a conforming metadata instance, it conforms as a metadata writer. If the system can read and store any conforming metadata instance written from another system, then it conforms as a metadata reader. Finally, if it can read a conforming metadata instance, store it, and write the same instance back out, the system has metadata repository level conformance. These ideas arise when we address usability. Many details about how systems conform to technical specifications can be found in [7]. These ideas work for strict conformance, too.

If we consider the example of the technical specification for describing a variable in Table 3, the rows are the characteristics of a variable. These correspond to the characteristics column in Tables 1 and 2. We are assuming here that the technical specification in Table 3 contains the schema elements necessary to make variables transparent. The properties column in Table 1 describes the marital status variable named MS01. These are metadata, and they strictly conform at the metadata instance level.

Continuing with the example, if we assume Table 3 contains a complete set of required characteristics needed to describe a variable, then the description in Table 1 is a complete description of MS01, since it strictly conforms to Table 3. If a user needs a description of the variable MS01, Table 1 makes that information transparent to the user through metadata writer strict conformance. Transparency is achieved by a system

strictly conforming to the specification in Table 3 as a metadata writer. This implies strict conformance to a specification is necessary for transparency. However, is this all that is needed?

There are two questions that naturally arise. The first question is "Are the metadata (the properties for MS01) correct?" This is the metadata quality issue. The other question is "Is there some system that the user can employ to find this information, or do they always have to turn to this paper?" This is the usability concern.

We address these in the following sections.

6. METADATA QUALITY

Metadata are data as well, so it is natural to apply data quality frameworks [8] to metadata to determine their quality. However, metadata are data in the role of describing some other resources. Therefore, a metadata quality framework needs to account for this descriptive aspect.

Metadata that are instances of a schema are a shorthand for the textual descriptions contained in traditional documentation. Consider the schema in Table 1. The schema elements and instance values for the variable MS01 generate a set of declarative sentences as follows:

- Name of the variable is *MS01*.
- Defining concept of the variable is *legally defined marital state*.
- Universe of the variable is *adults*.
- Question capturing the variable is *what is …'s marital status?*
- Datatype of the variable is *nominal*.
- Allowed values for the variable are *<1, single>, <2, married>, <3, separated>, <4, divorced>, <5, widowed>*.

These simple sentences may be combined, with a few editorial flourishes, into a textual description of the MS01 variable. The text might look like this:

The variable named MS01 measures the marital status of adults. In this case, marital status is defined as the *legally defined marital state*, and the possible values that may be assigned are *<1, single>, <2, married>, <3, separated>, <4, divorced>, <5, widowed>*. These are unordered categories, so the datatype is *nominal*. The question used to capture the marital status for each adult is "What is …'s marital status?"

It is evident that this textual description is equivalent to the combination of the sentences, and they in turn are derived directly from the schema for variables and the marital status variable instance. This means the schema and instance are equivalent to the textual description of the variable named MS01. This story approach is consistent with interview studies by Maron [9] of metadata experts, who indicate they consider the story metadata tell.

The question of the quality of metadata (the instance values of the schema) and the equivalent text can be broken into the language components syntax, semantics, and pragmatics; and we describe this further here.

Syntax, semantics, and pragmatics are commonly used criteria in information quality and metadata quality frameworks. Sundgren [10] and Price and Shanks [11], [12] used the framework to talk about information quality—the story that some data convey about society. Sundgren only briefly defined each component as representational, contents-oriented, and purpose-oriented, respectively. Price and Shanks defined syntax and semantics more carefully, they indicated syntax is about format rules, allowed values fitting validity criteria, and the correct code representing a category is assigned. Semantics refers to whether the correct meaning has been assigned to elements, so, to use a simple example, if an adult is married their marital status is recorded as such. Pragmatics was defined in terms of the user and their judgment as to how relevant some information is.

Myrseth *et al* [13] used the same framework to talk about metadata quality as the metadata relate to some underlying data. This is like the information quality approaches above. Syntax refers to validations with respect to some schema, semantics refers to the match between data and what they are intended to represent, and pragmatics refers to the perception by the user of the quality of information.

Price and Shanks [11] and [12] took a semiotic approach to recording data and metadata. This is based on the earlier work of C. S. Pierce [14]. The approach allows a systematic way of tying representations and meanings, especially in information systems. Myreth *et al* [13] adopted the same ideas. Interestingly, for Price and Shanks this leads to a definition of data from the semiotic perspective. Farance and Gillman [5] arrived at the same idea independently but extended the definition to include a datatype as described in ISO/IEC 11404 [6].

We base our metadata quality framework using the same terms, syntax, semantics, and pragmatics as above. The definitions we use for syntax and semantics are roughly the same, but we alter the use of the term pragmatics.

The approach considers both the schema instance and the derived story and asks if they are true. This addresses metadata quality from the point of view of description.

First, we need to make sure each instance value is valid under the rules of the technical specification. We refer to this as the syntax aspect of metadata quality. For example, typing errors, formatting mistakes (e.g., using more than 16 characters in the variable name, see Table 3, entering the wrong date/time format), and assigning codes not on the valid list are syntax errors.

The syntax rules are part of the technical specification that defines the schema, and they are among the requirements provided. Therefore, the syntax component of metadata quality is achieved through strict conformance to the schema. If an instance of the schema conforms, it must satisfy the syntax rules.

The next level is semantic. Do the instance values convey meaningful information? For example, consider the datatype in Table 1. Looking at the numeric codes from the list of allowed values, someone might decide this is an ordered list of categories and select the ordinal datatype. That contradicts the nature of marital status categories. They are not ordered.

In our example, we list 5 marital statuses, but maybe that was changed from an original list of 4: single, married, divorced, widowed. Separated was added later. Then, if the choices associated with the question do not match the allowed values, this is a semantic error also. In another example, the universe might say monks instead of adults. That could be a valid value, but for a marital status variable it does not make a lot of sense, because monks take a vow of celibacy and don't get married.

Referring back to the sentences derived from the pairs of schema elements and instance values, we ask if each is true. This is the semantic aspect of metadata quality. The truth may be defined in the same way Tarski did for formal languages [15]. The sentence *The datatype of the variable is nominal* is true iff (if and only if) the datatype really is nominal.

The highest level is the pragmatic aspect of metadata quality. Here, we focus on the story rather than each sentence. We ask the following questions of the story:

- Is the story complete? In other words, does it contain the whole truth?
- Is the story concise? In other words, does it contain nothing but the truth?

The schema element / instance value pairs might be true individually (i.e., correct), but they don't necessarily tell a complete and concise story. If part of the story is left out, the story is not complete. If part of the story is irrelevant or unnecessary, the story is not concise. Using the example in Table 1 and the story implied (see above), examples of possible missing and irrelevant characteristics are

1. Missing: leaving out the set of allowed values
   a. The description is incomplete
   b. This corresponds to whether the **whole truth** about the resource is provided.
2. Irrelevant: Include an estimate of the population of the Soviet Union
   a. The description of the variable is not concise because the Soviet Union no longer exists as a country and the included information is not relevant to understanding variables
   b. This corresponds to whether **nothing but the truth** about the resource is provided.

Pragmatics fails when the story fails in some way. From the point of view of the technical specification we conform to, we expect that specification to include a set of elements that are relevant and complete. In this way, the stories based on the schema element / instance value pairs (key/value pairs) are informative.

Using our example again, we limit transparency for variable MS01 if there are errors in the description. Conveying erroneous information may be worse that conveying none. An error associated with any of the quality aspects leads to this conclusion, though it is possible that some errors will be more harmful than

others. The important point is that conformance by itself does not imply quality, and lack of quality impedes transparency.

We summarize the metadata quality framework in Table 4 below:

Table 4. Metadata Quality Framework.

Metadata Quality Framework		
Components	Definition	Verification
Syntax	Instance values follow formatting and validity rules	Conformance of instances to metadata schema
Semantics	Each instance value is correct	Truth of each instance value
Pragmatics	Story generated by all element / instance pairs is complete and concise	Whole truth and nothing but the truth of generated story

We note previous authors defined the pragmatic aspect of metadata quality as something seen by the users of the metadata. Our framework takes a different perspective. It places pragmatics as the aspect of quality associated with the combination of each schema element / instance value pair into forming a narrative. It is the accuracy of this narrative with respect to all three quality aspects that characterizes our framework.

7. SYSTEM USABILITY

There are many articles and books about usability and human-computer interaction. A whimsical and leisurely explanation of usability is in [16]. [17] and [18] contain a more detailed look at human-computer interaction. The most important thing to understand is that any effective system must have a usable interface. Failure to achieve this makes systems hard to use. Hard to use systems don't deliver information easily or, often, completely. Transparency is reduced when information is difficult to get.

For us, usability is applied to the interface and functionality of the system. There are two main considerations:

- The system must provide adequate functionality to support transparency
- The system user interface must be usable.

The main functionality needed to support transparency is for the system to provide the necessary metadata. Usability of an interface is closely related to conformance to a metadata specification. It is possible that through an interface to a system that another system (usually a person) can determine whether the first system conforms to some specification. This is possible when the interface is usable. The usable interface allows the person to inspect the system to make sure all requirements of the specification are satisfied. So, by inspecting the output of the system, it is possible check whether it conforms at the writer level. If the system interface includes the ability to input metadata and those metadata are expected to conform to a metadata specification, then the system must conform at the reader level, too. If the input metadata are

delivered back to a user (the same or different than the one providing the metadata), the system must conform at the repository level.

The other main consideration is how easily a user can interact with the system interface. Placement of text and buttons; use of fonts and colors; naming of labels on buttons, pages, and links; and whether provided functions behave in the expected way are all areas that usability testing can improve. A usable system delivers the desired results in a predictable and straightforward way. This is achieved through affordances—the ability of the user to intuit which buttons and what functions to use [19].

This points to something interesting, though. The usability necessary to determine conformance to the technical specification implies an object described by a high-quality metadata instance of the technical specification is transparent. In this way, all our transparency criteria are inter-connected.

8. CONCLUSION

In this paper, we make the case that conforming to a technical specification for metadata, achieving metadata quality, and ensuring that any system interface is usable are necessary and sufficient to achieve transparency.

In the course of presenting the argument for our 3 criteria for transparency we laid out 3 components for assessing metadata quality—syntax, semantics, and pragmatics. We provided criteria for assessing each of the components: conformance for the syntactic component, the truth of each instance value for the semantic component, the whole truth for the completeness aspect of the pragmatic component, and nothing but the truth for the conciseness aspect of the pragmatic component.

We demonstrated how the criteria for transparency are interrelated, and these interrelationships means all the criteria need to be considered when planning and designing systems supporting transparency.

REFERENCES

[1] The Promise of Evidence-Based Policymaking, Commission on Evidence-Based Policymaking, Chaired by Katharine Abraham, September 2017. Available at: https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2019/03/Full-Report-The-Promise-of-Evidence-Based-Policymaking-Report-of-the-Comission-on-Evidence-based-Policymaking.pdf. Accessed 19 February 2022

[2] National Academies of Sciences, Engineering, and Medicine; Division of Behavioral and Social Sciences and Education; Committee on National Statistics; Panel on Transparency and Reproducibility of Federal Statistics for the National Center for Science and Engineering Statistics. Transparency in Statistical Information for the National Center for Science and Engineering Statistics and All Federal Statistical Agencies (2021). Washington, DC: The National Academies Press. https://doi.org/10.17226/26360.

[3] Gentner, D., Kurtz, K.: Learning and using relational categories. In: W.K. Ahn, R.L. Goldstone, B.C. Love, A.B. Markman & P.W. Wolff (Eds.), Categorization Inside and Outside the Laboratory. Washington, DC: APA (2005)

[4]   ISO 704: 2000. Terminology Work—Principles and Methods. International Organization for Standardization, Geneva.

[5]   Farance, F., Gillman, D.: Foundational Metadata in Cross-Domain Integration: Detailed Model. Ch2, pp. 9–21. Available at: https://ddialliance.org/Specification/ddi-cdi (2020)

[6]   ISO/IEC 11404: 2007. Information technology—General purpose datatypes. International Organization for Standardization, Geneva.

[7]   ISO/IEC 20944-1: 2013. Information technology—Metadata Registries Interoperability and Bindings (MDRIB)—Part 1: Framework, common vocabulary, and common provisions for conformance. International Organization for Standardization, Geneva.

[8]   Ehling, M., Korner, T. (eds): Handbook on Data Quality Assessment Methods and Tools. European Commission, Weisbaden (2007)

[9]   Maron, D. The Pragmatics of Metadata: From Practice to Concept to Practice, IEEE Tech. Comm. Digit. Libr. (2018)

[10] Sundgren, B.: Guidelines for the Modelling of Statistical Data and Metadata. Published as Guidelines from the United Nations Statistical Division, New York (1995)

[11] Price, R., Shanks, G.: Empirical Refinement of a Semiotic Information Quality Framework. In: Proceedings of the 38th Hawaii International Conference on System Sciences, p. 216a (2005)

[12] Price, R., Shanks, G.: A Semiotic Information Quality Framework: Development and Comparative Analysis. In: Willcocks, L.P., Sauer, C., Lacity, M.C. (eds): Enacting Research Methods in Information Systems. Palgrave Macmillan, Cham (2016)

[13] Myrseth, P., Stang, J., Dalberg, V.: A data quality framework applied to e-government metadata: A prerequsite to establish governance of interoperable e-services; 2011 Conference on E-Business and E-Government (ICEE) (2011)

[14] Pierce, C.S.: Collected Papers. Harvard University Press, Cambridge, MA (1931–1935).

[15] Tarski, A.: "Truth, Meaning, and the Indeterminacy of Translation—The Semantic Conception of Truth" in Philosophy of Language: The Central Topics, Chapter 2.

[16] Krug, S.: Rocket Surgery Made Easy. New Riders (2009)

[17] Cowley, B., Filetti, M., Lukander, K.: The Psychophysiology Primer—A Guide to Methods and a Broad Review with a Focus on Human-Computer Interaction in Human-Computer Interaction. Now Publishers, Inc. (2016)

[18] Dix, A., et al.: Human-Computer Interaction. 3rd Ed. Pearson (2003)

[19] Norman, D.: The Design of Everyday Things. MIT Press, Cambridge, MA (2013)

AUTHOR BIOGRAPHY

Daniel Gillman works in the Office of Survey Methods Research at the U.S. Bureau of Labor Statistics (BLS). His research interests include metadata, standards, terminology, and transparency. At BLS, Gillman led the effort to build a taxonomy of terms describing all time-series data and was a member of the team to build a glossary of BLS technical terms. He is consultant to several program initiatives at BLS to build metadata repositories in support of public data releases and internal processing efficiencies, to the BLS output database redesign effort, and to data-governance modernization effort sat the U.S. Department of Labor. He is chair of an interagency ad hoc group to develop guidance on metadata management for the statistical agencies. Over the years, Gillman has chaired numerous metadata standards groups under the US federal government, ANSI/INCITS, SDMX, UNECE, and the DDI-Alliance. He is a member representative to the DDI-Alliance and is a key developer of the DDI-4 (Cross-DomainIntegration) model driven standard. In addition, he has participated in numerous efforts across many standards-development organizations, and he has participated as an expert consultant to several metadata modeling efforts outside BLS. Prior to working at BLS, he worked at the U.S. Census Bureau.

ORCID: 0000-0002-1638-3881