An Innovative Approach to Standards Analysis: Center of Gravity Analysis for International Standards Published by ISO/IEC JTC1 SC 36 Information Technology for Learning, Education and Training

Dr. Bruce E. Peoples
Laboratoire Paragraphe (EA349), Université Paris 8
Saint-Denis, France
brucepeoples02@gmail.com

Abstract

Standards make a positive contribution to the world we live in. They facilitate trade, spread knowledge, disseminate innovative advances in technology, and share good management and conformity assessment practices. There are a multitude of standard and standard consortia organizations producing market relevant standards, specifications, and technical reports in the domain of Information Communication Technology (ICT). With the number of ICT related standards and specifications numbering in the thousands, it is not readily apparent to users and developers how these standards inter-relate to form a basis of technical interoperability. There is a need to develop and document a process to identify how standards form a basis of interoperability in multiple contexts at a general horizontal technology level that covers all technology domains, and within specific vertical technology domains and sub-domains. By analyzing which standards inter-relate through normative referencing, key standards can be identified as technical center of gravity standards. These normatively referenced standards are specific standards required for the successful implementation of standards that normatively reference them, and form a basis for interoperability. This paper gives an overview of a methodology for determining center of gravity standards utilizing International Standards published by ISO/IEC JTC1 SC 36 Information Technology for Learning, Education and Training as a basis of analysis.

Keywords—standard analysis; ISO/IEC JTC1 SC 36; center of gravity standards; standards; ISO; IEC

1. INTRODUCTION

An ISO International Standard represents global consensus on the state of the art in the subject of that standard [1]. One of the purposes of an ISO standard is to promote interoperable implementations of a technology. It is not readily apparent how standards inter-relate to form a basis of interoperability across multiple standards representing a horizontal technology level, and numerous vertical technology levels.

Currently, identifying inter-relationships between standards take the forms of (1) focus of a specific standards sub-committee, (2) production of multi-part standards within a sub-committee, and (3) through examining normative references cited by a *single* specific standard. There is a need to clearly identify those standards that promote interoperability on a larger scale, standards that form a basis of interoperability across numerous standards that form a horizontal technology level such as learning, education and training, and vertical technology levels such as metadata, collaborative technologies, etc.

In researching standards, I have designed and implemented a methodology to discover how standards inter-relate through normative referencing across several published International Standards (IS) to form a basis of interoperability. The normatively referenced standards cited by a majority of standards under study in a dataset are what I term potential *center of gravity* standards. Key to the methodology is the use of normative references. A normative reference, is a conditional element in a standard document that lists cited documents that are indispensable for the application of the standard [2].

By identifying center of gravity standards, new patterns can be discovered and utilized. These patterns include (1) finding the critical technologies that promote interoperability among standards, (2) identifying/predicting shifts of center of gravity standards over time, (3) identifying/predicting the creation of new center of gravity standards over time, (4) identifying/predicting growing, stable or shrinking technology domains, (5) identifying specific technology verticals that share linkages with other technology verticals, and (6) improving management of the standards development and maintenance processes.

2. BASIC METHODOLOGY

The methodology developed was used to scientifically discover center of gravity standards utilizing both visual and mathematical techniques afforded by Social Network Analysis (SNA) based on Graph Theory, and implemented using Knowledge Discovery in Databases (KDD) techniques.

The KDD process utilized is a 5 stage iterative process, based on the Cross Industry Standard Process for Data Mining (CRISP-DM) developed by the CRISP-DM SIG (Special Interest Group), funded as an EU Project under the ESPRIT initiative [3]. The 5 stage process was used to (1) select a dataset, (2) pre-process the data contained in the dataset, (3) transform data in the dataset to a usable form for SNA tools, (4) mine the data looking for center of gravity patterns utilizing a custom algorithm and SNA analysis techniques, and (5) evaluate and interpret patterns found through analysis.

NodeXL, an open source SNA tool was used to import data, create a graph visualization of the dataset, and provide high level graph metrics [4]. UCINET, a SNA tool developed by Analytic Technologies, was used to produce supporting metrics to confirm the existence of center of gravity standards [5].

A. Dataset

The dataset used in testing the methodology consisted of International Standards published as of 2012 by the International Organization of Standards/International Electrotechnical Committee; Joint Technical Committee 1, Sub-committee 36 Learning, Education and Training (ISO/IEC JTC1 SC 36), and their normative references.

In the test dataset, there are 18 IS produced by ISO/IEC JTC1 SC 36. The 18 published IS are considered the horizontal technology level. ISO/IEC JTC1 SC 36 has implemented a Working Group (WG) structure to facilitate effective production of IS. Each WG focuses on standards for specific technology domains such as collaborative technologies, and management and delivery technologies. The IS produced by WGs are considered vertical technology levels. There are also 28 standards and technical reports in the dataset produced outside of ISO/IEC JTC1 SC 36, normatively referenced by ISO/IEC JTC1 SC 36 IS contained in the dataset.

B. Metrics Utilized

The methodology implemented a metric developed specifically for the use in identifying center of gravity standards, and four metrics calculated by NodeXL and UCINET to confirm the existence of center of gravity standards:

- In-Degree - Edges pointing into a vertex. For this methodology, In-Degree edges are edges where the head of the edge points to a normatively referenced standard [6].
- Threshold Percentage - A calculation developed specifically for the methodology to identify center of gravity standards:

  Threshold Percentage = X / Y * 100

  X is the number of times a specific standard is normatively referenced by a published ISO/IEC JTC1 SC 36 IS (In-Degree count) and Y is the number of published ISO/IEC JTC1 SC 36 IS (18) contained in the dataset. The quotient is multiplied by 100 to create a percentage. The percentage is the Threshold Percentage. The percentage must be greater than 50%, for a normative reference to be considered a center of gravity standard. The greater than 50% threshold indicates the normative reference is utilized by over half of the published ISO/IEC JTC1 SC 36 IS.
- In-Degree Centrality - The number of nodes that an In-Degree focal node is connected to, and measures the involvement of the In-Degree node in the network [7]. In-Degree Centrality measure supports identification of center of gravity standards in the dataset. There should be a pattern of decreasing In-Degree Centrality values as In-Degree count decreases.
- In-Degree Closeness Centrality - The inverse of the sum of the shortest distances between an In-Degree vertex and all other vertices reachable from it [7]. In-Degree Closeness Centrality measure supports identification of center of gravity standards in the dataset. There should be a pattern of decreasing In-Degree Closeness Centrality values as In-Degree count decreases.
- In-Farness - The sum of the lengths of the geodesics from an In-Degree vertex to every other vertex. Farness is the reciprocal value of Closeness Centrality [7]. In-Farness measure supports identification of center of gravity standards in the dataset. There should be a pattern of increasing In-Farness values as In-Degree count decreases.

C. Graph Created

NodeXL was utilized to generate the visual graph shown in Fig. 1. The dataset was entered into NodeXL as ordered pairs where an ISO/IEC JTC1 SC 36 produced IS formed the first value of the ordered pair, and a normative reference utilized in the IS was entered as the second value of the ordered pair. The ordered pairs were organized by groups, designating WGs that produced the IS containing normative references. A group identified as NON-SC36 contained standards not produced by ISO/IEC JTC1 SC 36. The NON-SC36 group contains only normatively referenced standards.

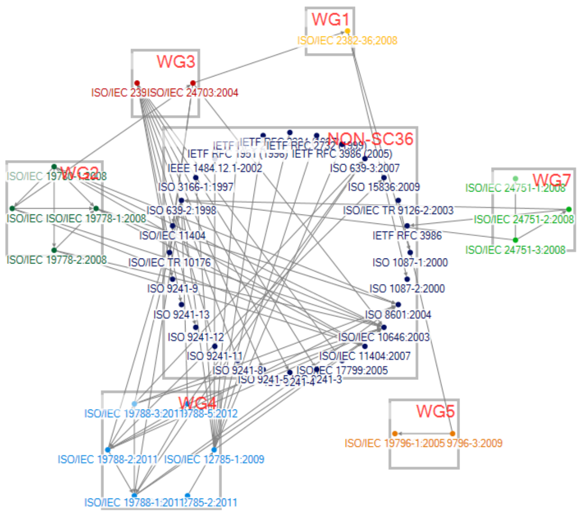

Fig.1 NodeXL generated graph

3. GENERATION AND ANALYSIS OF METRICS

Generation and analysis of metrics was performed at two levels, the horizontal technology level, the ISO/IEC JTC1 SC 36 level, and vertical technology levels, the ISO/IEC JTC1 SC 36 Working Group levels.

A. ISO/IEC JTC1 SC 36 Horizontal Technology Level

NodeXL was utilized to generate accurate In-Degree counts. The top five In-Degree standards, normatively referenced standards, were used as an indicator of center of gravity standards. The Threshold Percentage calculation was then utilized to determine if any of the top five In-Degree standards exceeded the 50% value. As shown in Table I, one standard did exceed the 50% value, ISO/IEC 10646:2003 *Information Technology – Universal Coded Character Set*. ISO/IEC 10646:2003 was normatively referenced by 56% (rounded) of ISO/IEC JTC1 SC 36 standards under consideration and can be considered a center of gravity standard for ISO/IEC JTC1 SC 36 produced standards at the horizontal technology level. ISO/IEC 10646:2003 specifies the Universal Character Set (UCS). The standard is applicable to the representation, transmission, interchange, processing, storage, input and presentation of the written form of the languages of the world as well as additional symbols. The standard covers 110,181 characters from the world's scripts. This standard is key in forming interoperable implementations of systems for the exchange of data, information, and knowledge.

TABLE I. Threshold Percentages

Normatively Referenced Standard	In-Degrees	Total SC 36 Standards	Threshold Percentage
ISO/IEC 10646:2003	10	18	56%
ISO 639-2:1998	7	18	39%
ISO 8601:2004	5	18	28%
ISO/IEC 11404:2007	4	18	22%
ISO 639-3:2007	3	18	17%

Metrics supporting the identification of ISO/IEC 10646:2003 as a center of gravity standard were generated and analyzed. The matrix of ordered pairs created in NodeXL was exported in a UCINET DL format, and then imported into UCINET. Normalized In-Degree Centrality, In-Degree Closeness Centrality, and In-Degree Farness were calculated utilizing UCINET. Table II shows the expected patterns of decreasing In-Degree Centrality values as In-Degree count decreases, increasing In-Farness values as In-Degree count decreases, and decreasing In-Degree Closeness Centrality values as In-Degree count decreases. This supports identification of ISO/IEC 10646:2003 as a center of gravity standard at the ISO/IEC JTC1 SC 36 level, the horizontal technology level.

TABLE II. In-Degree Centrality Measures

Normatively Referenced Standard	In-Degrees	NrmInDeg	InFarness	InCloseness
ISO/IEC 10646:2003	10	22.222	1576	2.855
ISO 639-2:1998	7	15.556	1711	2.630
ISO 8601:2004	5	11.111	1713	2.627
ISO/IEC 11404:2007	4	8.889	1890	2.381
ISO 639-3:2007	3	6.667	1891	2.380

In Fig. 2, ISO/IEC 10646:2003 appears in red with red edges showing WGs developed IS that normatively references ISO/IEC 10646:2003. There is a visual pattern showing standards developed by WG2 Collaborative and Intelligent Technology (collaboration vertical technology level), and WG4 Management and Delivery of Learning, Education and Training (metadata vertical technology level) normatively references ISO/IEC 10646 more than any other WGs.

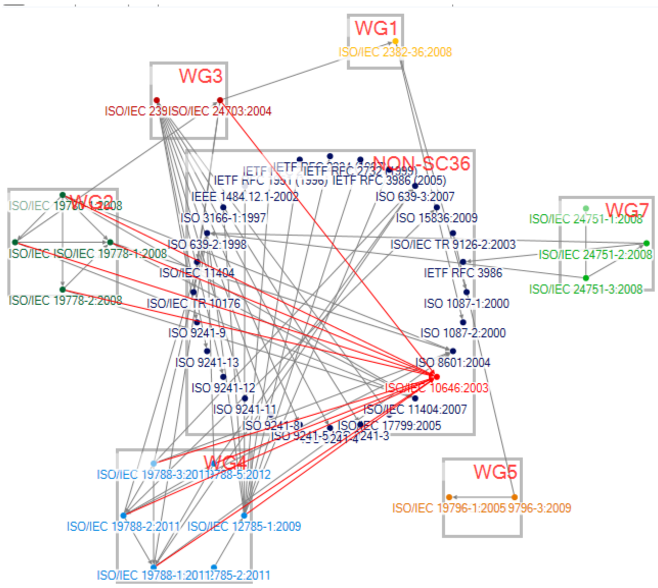

Fig. 2 NodeXL generated graph with ISO/IEC 10646:2003 in red

B. ISO/IEC JTC1 SC 36 Work Group Vertical Technology Levels

At the WG levels, vertical technology levels, NodeXL was utilized to generate accurate In-Degree counts. The top 5 In-Degree standards, standards that were normatively referenced, were used as an indicator of center of gravity standards.

The Threshold Percentage calculation was then utilized to determine if any of the top 5 In-Degree standards within each WG exceeded the 50% value. Supporting metrics were not calculated for the vertical technology levels.

Threshold Percentages were generated on the basis of highest In-Degrees counts for normative references and total number of IS developed by each WG. Table III contains identified center of gravity standards based on Threshold Percentage calculations. The table shows different and in some cases identical center of gravity standards identified for each WG vertical technology level. It is interesting to note the Threshold Percentages calculated at the WG vertical technology levels are higher for identified center of gravity standards than the Threshold Percentage for the center of gravity standard at the ISO/IEC JTC1 SC 36 horizontal technology level. This indicates higher Threshold Percentages for normatively referenced standards form stronger center of gravities within a vertical technology level.

TABLE III. WG Level Center of Gravity Standards

Work Group	Normatively Referenced Standard	In-Degrees	Total WG Standards	Threshold Percentage
WG1	ISO 1087-1:2000	1	1	100%
WG1	ISO 1087-2:2000	1	1	100%
WG2	ISO/IEC 10646:2003	4	4	100%
WG2	ISO/IEC 11404:2007	4	4	100%
WG2	ISO/IEC 19778-1:2008	3	4	75%
WG4	ISO/IEC 10646:2003	5	6	83%
WG4	ISO 639-2:1998	5	6	83%
WG7	ISO 639-2:1998	2	3	67%

There is another interesting pattern that can be seen in Table III. WG2 Collaborative and Intelligent Technology (collaboration vertical technology level), and WG4 Management and Delivery of Learning, Education and Training (metadata vertical technology level) shares ISO/IEC 10646:2003 *Information technology -- Universal Multiple-Octet Coded Character Set* as a center of gravity standard. Additionally, WG4 Management and Delivery of Learning, Education and Training (metadata vertical technology level), and WG7 Culture, Language, and Individual Needs (learner disability vertical technology level) shares ISO 639-2:1998 *Codes for the representation of names and languages – Part 2: Alpha-3 code*, as a center of gravity standard. This may indicate a stronger interoperability basis for some implementations based on WG2 and WG4 produced standards and WG4 and WG7 produced standards.

4. CONCLUSIONS

The developed methodology has shown center of gravity standards can be identified at both the horizontal technology level, in this context the ISO/IEC JTC1 SC 36 level, and vertical technology levels, in this context the ISO/IEC JTC1 SC 36 WG levels. These identified center of gravity standards allows identification of specific normatively standards across multiple technology contexts in forming a basis of interoperability between published ISO/IEC JTC1 SC 36 IS. Identified center of gravity standards can also be utilized as a basis for other analysis techniques and applications.

- As a basis for a recommender system to recommend potential normative references for standards development within technology verticals.
- As a basis for time series analysis to visualize how center of gravity standards form, expand, contract, and shift over periods of time.
- As a basis for predicting the formation of center of gravity standards and related shifts in technology verticals.
- As a basis for standards maintenance activities utilizing the created graph to visualize potential impacts of changes to normatively referenced standards on standards that utilize the normatively referenced standards.
- As a basis to create ontologies for advanced pattern analysis of center of gravity standards.

5. FUTURE RESEARCH

Although the dataset utilized in the developed methodology is small, the methodology was created to allow scaling to larger datasets. The methodology should be used on a larger dataset such as all standards produced by ISO/IEC JTC1. In this context, ISO/IEC JTC1 would be considered the horizontal technology level and ISO/IEC JTC1 Sub-committees would be considered vertical technology levels.

6. ACKNOWLEDGEMENT

I am deeply grateful to Renaud Fabre and Imad Saleh, co-directors of my PhD Thesis for their guidance, support and inspiration in defining and researching the concept of center of gravity standards.

7. REFERENCES

[1] ISO Central Secretariat. (2010). *ISO Strategic Plan 2011-2015*. Geneve: International Organization for Standards.

[2] ISO/IEC Technical Management Board. (2011). *ISO/IEC Directives, Part 2*. Geneve: ISO/IEC.

[3] Azevrgo, A. S. (2008). KDD, SEMMA and CRISP-DM: a parallel overbiew. *IADIS European Conference on Data Mining 2008*. Amsterdam: IADS 2008.

[4] Smith, M. M.-F. (2010). *NodeXL: a free and open network overview, discovery and exploration add-in for Excel*. Retrieved from CodePlex: http://nodexl.codeplex.com/

[5] Borgatti, S. E. (2002). *UCINET for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.

[6] Sedgewick R., W. K. (2011). *Algorithms, 4th Edition*. Pearson Education, Inc.

[7] Analytec Technologies. (n.d.). *UCINET 6 Help for Windows*. Retrieved from analytictech.com: http://www.analytictech.com/ucinet/help/webhelp.html