Efficient Visual Navigation of Hierarchical, Educational Standards Crosswalks Between Multiple Collections

Challenge Overview
https://www.mindsumo.com/contests/educational-data-visualization

Abstract—**In K-12 U.S. education, the learning goals that document the expected knowledge, skills and other characteristics for learners at various grade levels are typically referred to as standards. Education standards have a subject focus (e.g. math, ELA) and a varying scope of applicability (e.g. national vs. state). State Departments of Education, district, school administrators, contractors and educators are held accountable to several standards collections as they prepare and select curriculum, build lesson plans, identify learning resources and design assessments. Standards alignment is a process where standards describing similar concepts across multiple standards collections are correlated through a crosswalk.A challenge exists in efficiently presenting standards and the resultant crosswalks so they can be effectively navigated and reviewed. Recent advancements in interactive information visualization, in particular graph and network visualization seem to be a promising approach to address some of the related challenges. This paper outlines the criteria for designing an acceptable solution to such approach.**

CONTENTS

LIST OF FIGURES

I. INTRODUCTION

In the late 1980s, policymakers, educators and reformers began advocating the publication of high, rigorous standards that would apply to all students and that would clearly communicate to students, teachers and others the high level of achievement expected [1]. This policy effort led to the Goals 2000: Educate America Act of 1994 that encouraged the use of:

> "... two types of standards: content standards that would convey what students should learn, and performance standards that would indicate how well students would be expected to perform with respect to those content standards ... an overarching vision statement for each subject; content and performance standards describing knowledge, skills and understandings as well as the levels of competence expected; and school delivery standards that were intended to ensure opportunity to learn..." [2]

This focus on the development of standards drove several initiatives, e.g.

- several U.S. governors and corporate leaders founded Achieve, Inc. as a bipartisan organization to raise academic standards and graduation requirements, improve assessments, and strengthen accountability in all 50 states.
- state leaders led the development of state-specific standards collections, e.g. Virginia Standards of Learning[1]
- the National Governors Association (NGA) convened a group of people to work on developing the "common core state standards".

A. Education Standards

As an education standards collection example, the Common Core State Standards (CCSS) is one of the most prevalent. These standards were released for mathematics and English language arts on June 2, 2010, with a majority of states adopting the standards in the subsequent months [3]. States were given an incentive to adopt the Common Core Standards through the possibility of competitive federal Race to the Top grants. Though the Common Core State Standards do not

[1] http://tinyurl.com/qjl39gr

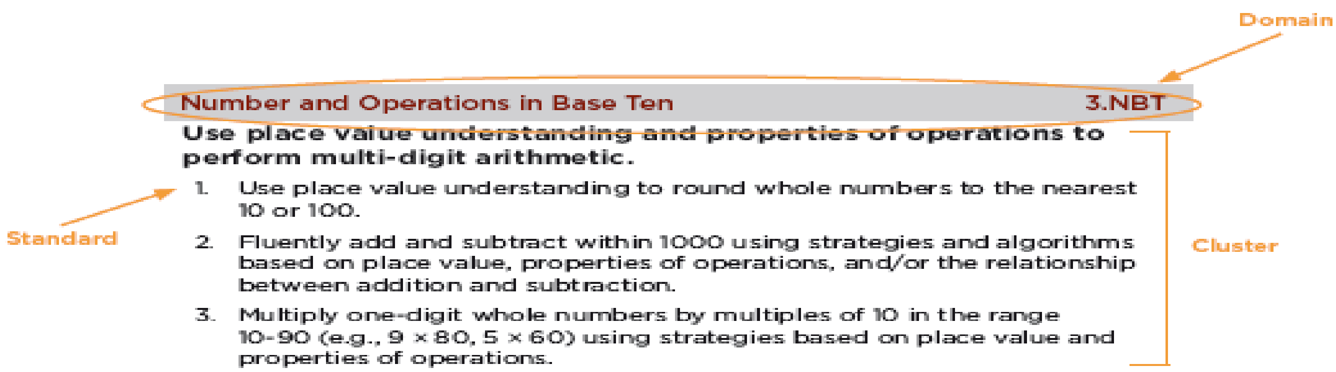

Fig. 1. Common Core State Standards - Math Levels

Fig. 2. CCSS Math Standards Snippet

cover science and social studies content standards, the Next Generation Science Standards were released in April 2013 and have been adopted by many states.

Educational standards such as common core state standards for math[2] can be visualized as a hierarchical graph structure. The common core math standards use a structure of Subject (Math) $\rightarrow$ Domains $\rightarrow$ Clusters $\rightarrow$ Standards. From a top level graph "node" representing the "math" subject there are second level nodes that represent domains, e.g. "Counting and Cardinality" or "Number and Operations in Base Ten". Domains are further structured into clusters and clusters contain node instances that represent an individual standard, For example, see Figure 1.

Domains
: Domains are larger groups of related standards. Standards from different domains may sometimes be closely related.

Clusters
: Clusters summarize groups of related standards. Note that standards from different clusters may sometimes be closely related, because mathematics is a connected subject.

Standards
: Standards define what students should understand and be able to do.

A depiction of such graph example that proceeds from Subject (Math) $\rightarrow$ Domains $\rightarrow$ Clusters $\rightarrow$ Standards is shown in Figure 2. In this example we can see a node-link diagram with the root at the top. The nodes represent standards data and the edges represent $\{parent, child\}$ relationships e.g. $[\{n_1, n_{1.1}\}, \{n_{1.1}, n_{1.1.1}\} \ldots]$.

We can imagine another hierarchical collection of standards that incorporates an enumeration of math standards as shown in Figure 3. Again, this is depicted as a node-link diagram with $\{parent, child\}$ relationships e.g. $[\{n_2, n_{2.1}\}, \{n_{2.1}, n_{2.1.1}\}, \{n_{2.1.1}, n_{2.1.1.1}\} \ldots]$.

State Departments of Education, district, school administrators, contractors and educators typically need to comprehend several hierarchical collections of standards and critically they often need to understand how each collection relates to or is aligned with others.

B. *Crosswalks*

Standards alignment is a process where standards describing similar concepts across multiple standards collections are correlated into a crosswalk. The literature describes the activities of several working groups, committees, boards, review teams [4], [5], [6], [7] leveraging various crosswalk methodologies, e.g. WestEd methodology [8].

This process "can be a very time-consuming activity" [9] which translates into an expensive effort. Technology can help

[2] http://www.corestandards.org/Math/

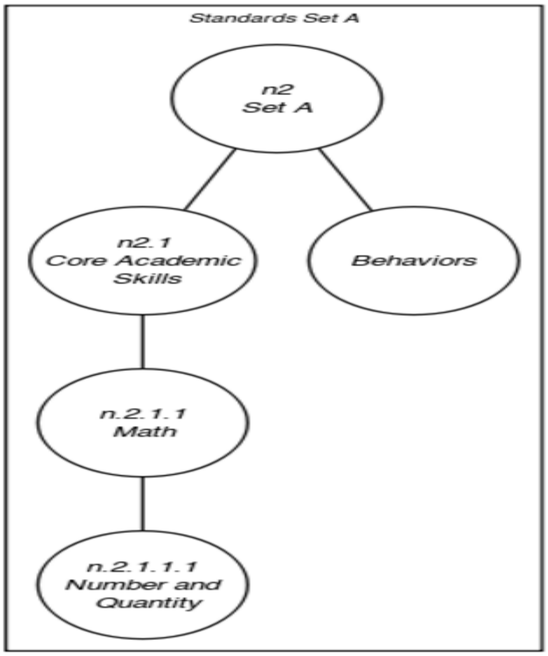

Fig. 3. Standards Set A Snippet

make the process of developing crosswalks more efficient and can be used to create effective interactive visualizations of the resultant data investment.

Efforts such as the Achievement Standards Network (ASN) national repository of comprehensive machine-addressable achievement standards [10] has enabled the development of applications serving the education community including search engines, metadata generation tools and other third party services. For example, various researchers have developed automated or semi-automated tools to help suggest/build these crosswalks. Yilmazel and Liddy at the Center for Natural Language Processing (CNLP) in the School of Information Studies at Syracuse University [11] built the Standards Alignment Tool (SAT). This was part of the Computer-Assisted Content Standard Assignment and Alignment (CASAA) project and used natural language processing to suggest possible alignments between Achievement Standards Network (ASN)[10] standards statements. In another example, WGBH - a non-commercial educational PBS member television station located in Boston, Massachusetts - used ASN standards data in their Teachers Domain intermediary application that generates its alignment mappings dynamically through use of a controlled vocabulary performing the switching functioning[3]. Through a member's Teachers Domain profile, the system maps all retrieved educational resources to the controlling standards in the member's state.

Standards crosswalking offers many advantages. It helps viewers understand the overall matchup between two collections, viz. where there might be gaps or disagreements in structure. It also implies transitive applicability of prior mapping of learning resource of the following form: [9]

TABLE I
USES OF TRANSITIVE CROSSWALKS

Premise 1: Standard P is similar to standard Q.
Premise 2: Learning resource X aligns with standard Q.
Conclusion: Learning resource X aligns with standard P.

This can unlock remediation content for a standards collection by mapping it to another collection that has significant open education resource entries.

Using the node-link standards samples we have listed so far (Figure 2 and 3) we can represent the simplest possible crosswalk example in Figure 4. This graph diagram shows the two standards collection graphs side-by-side with a crosswalk link between one node in the left collection and one in the right collection: $\{source, target\}$ relationships e.g. $[\{n_{1.1.1}, n_{2.1.1.1}\} \ldots]$.

Typically crosswalk data is presented as a static table of data in a printed report. While this method of data presentation has a long history of use in many applications the rise of "visual thinking" has created several new means to enable information sharing, navigation and storytelling [12].

II. Data Visualization

Node-link diagrams of the type we have shown as an alternative to a table of data are prevalent today and have a long history of use. For example, 13th century work of Ramon Llull, who drew diagrams of this type for complete graphs in order to analyze all pairwise combinations among sets of metaphysical concepts [13]. These diagrams can effectively encode standards data and provide a visual understanding of the hierarchical structure. Given the naive node-link presentation of the crosswalk though (Figure 4) we can anticipate problems as this approach is applied to more realistic data. For example, there may be hundreds of crosswalk links between source and target, not to mention the potential requirement to concurrently present multiple target collection for a single source all in one view.

Researchers from computer science, cognitive science, data analytics, human factors and usability including specialized communities such as the Visual Analytics Science and Technology (VAST)[4] have all contributed to a better understanding of how to choose efficient and effective methods for visualizing data.

A. Graph and Network Visualization

There are several aspects from this body of research to consider as we try to find alternative, effective solutions to visualize and navigate crosswalks between a source and multiple targets.

[3]http://www.pbslearningmedia.org

[4]http://vacommunity.org/tiki-index.php

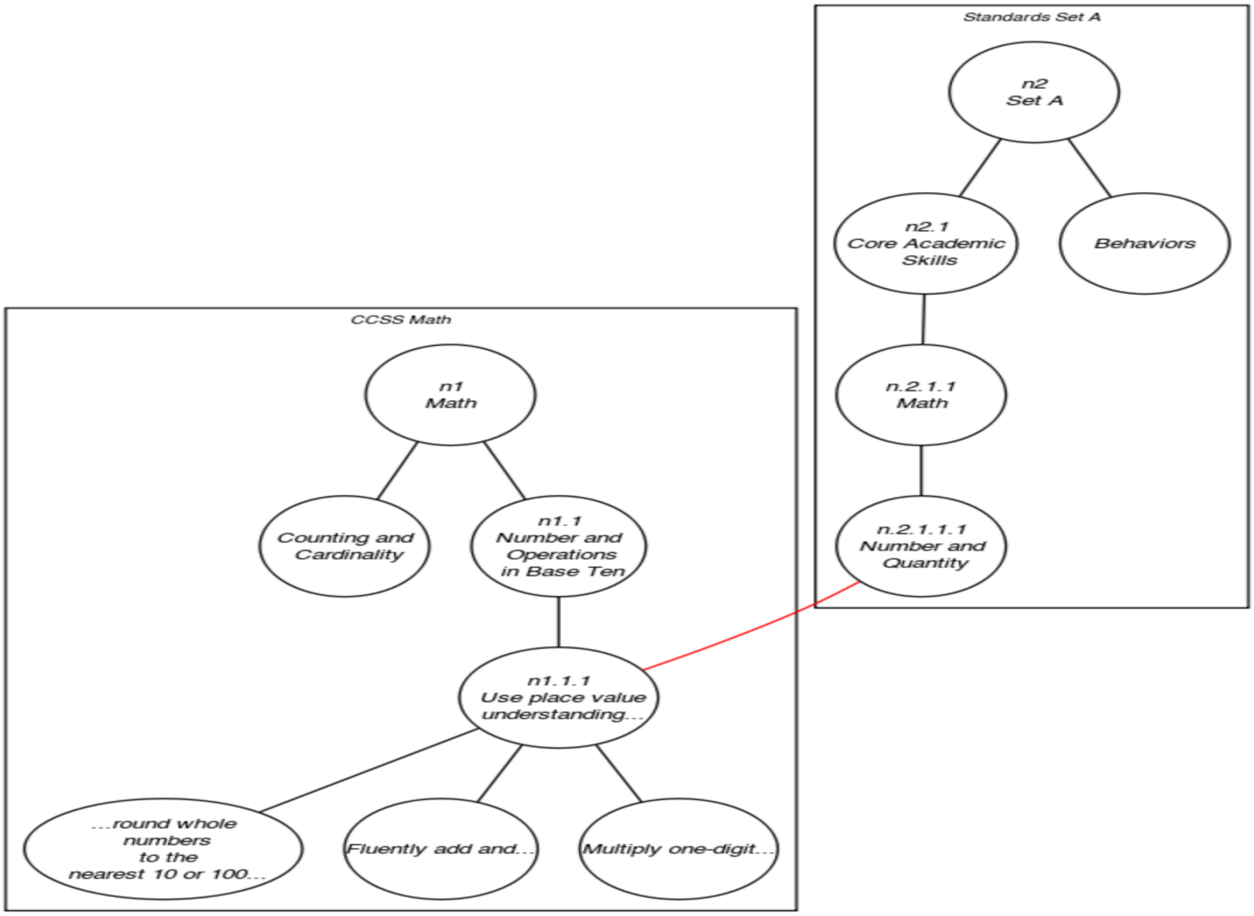

Fig. 4. CCSS – Standards Set A Crosswalk Snippet

1) Representation Scheme: There are two main representation schemes when we consider solutions for visualizing hierarchical data:

- Node-link
- Space-filling [14]

Representation scheme choices offers different tradeoffs/advantages:

- Presentation of detailed content information
- Presenting structural information
- Manual traversal, zooming/panning/scrolling/expanding/-collapsing
- Effort required by a viewer to formulate a mental model
- Scaling up to large collections
- Efficient use of available display space
- Maintaining an overall view

An important consideration is how a solution is made to move beyond simply a graph layout to graph visualization, i.e. a navigation solution that leads to comprehension over time. How can a solution enable a viewer (whether expert, intermediate or non-expert) to query, visit, or move around a graph? For example, changing focus may require different renderings.

2) Layout: There are several options to choose from when considering layout techniques:

- Hierarchical
- Force-directed
- Circular
- Clustered
- Attribute-based
- Matrix

3) Aesthetic Factors: Aesthetics are an important consideration in designing an efficient solution. Various research studies have examined which of the aesthetic factors matter most and/or what kinds of layout/vis techniques look best [15], [16], [17], [18] These results appear to be mixed although, e.g. consideration for edge crossings do seem important.

Here are several aesthetic considerations that may be relevant to evaluate for a solution:

- Crossings—minimize towards planar
- Total Edge Length—minimize towards proper scale
- Area—minimize towards efficiency
- Maximum Edge Length—minimize longest edge
- Uniform Edge Lengths—minimize variances
- Total Bends—minimize orthogonal towards straight-line

Shneiderman's often cited "NetViz Nirvana" points should also be consulted when evaluating the aesthetics of a solution:

1) Every node is visible
2) For every node you can count its degree
3) For every link you can follow it from source to destination
4) Clusters and outliers are identifiable

III. Goal

Given this context, the proposed goal is to:

> Design an interactive data visualization solution that presents a source standards collection hierarchy and multiple target collection hierarchies, along with the source-target crosswalks, that allows users (e.g. teachers, administrators) to efficiently navigate the information to better comprehend the mappings while e.g. designing lesson plans, making curriculum decisions or searching for learning content.

A. Example: ASN, Sankey approach

The ASN platform presents a public sankey[5] approach for visualizing crosswalks[6]. In their solution the source and target hierarchies are collapsed into a left and right hand lists. This solution requires a significant amount of vertical scrolling and it is hard to follow the connections as viewers get into the middle of the representation. Additionally this solution would be a poor candidate if we need to introduce multiple source-target crosswalks in a single visualization.

B. Design Parameters

For a proposed solution, we expect the following parameters:

1) Must be web-based
2) Must work in the full range of all modern browser releases, e.g. Chrome, Firefox, Safari, IE
3) Can leverage free, open source frameworks, tools (e.g. D3, etc.)
4) Optimize the representation, layout and aesthetics
5) Prefer to be able to easily identify what a node or link is about (e.g. hover over etc.)
6) Preference is for solutions that provide efficient navigation and can address advanced features, e.g. keyword search, filtering, automatic layouts, etc.

In order to provide a more focused target of who to design for and what tasks, a challenge response should be able to address the following scenario.

> The response should take the perspective of a teacher that needs access to a source set of standards (e.g. common core state standards) which are hierarchically arranged and concurrently needs access to their target state standards which are also hierarchically arranged. There may even be a third set of target hierarchical standards that are also relevant for their analysis. The source and targets have all been crosswalked by some committee or prior effort and the job of the teacher is to navigate these structures to perform the following analysis tasks while they are e.g. designing lesson plans, making curriculum decisions or searching for related learning content:

1) Browse detailed content information
2) Browse structural hierarchical information
3) Query, search and visit a graph or other visual representation
4) Perform interactive, manual traversal, possibly zooming/panning/scrolling/expanding/collapsing
5) Comprehend crosswalk linkages even at various states of traversal (e.g. expansion/collapsed states)
6) Identify gaps (i.e. where no alignment exists)
7) Perform tasks as either an expert, intermediate or non-expert
8) Potentially highlight areas of interest

C. Requested Deliverables

The challenge provider will distribute sample datasets for this challenge. The solution should allow for the inspection of webpage/pages that demonstrate an efficient interactive, information visualization solution to the problem.

[5]https://en.wikipedia.org/wiki/Sankey_diagram
[6]http://toolkit.asn.desire2learn.com/node/73/flow

References

[1] L. B. Resnick and D. P. Resnick, "Assessing the thinking curriculum: New tools for educational reform," in *Changing assessments*. Springer, 1992, pp. 37–75.

[2] D. Koretz and L. S. Hamilton, "Testing for accountability in k-12," in *Educational Measurement*. Prager Publishers, 2006, pp. 531–578.

[3] C. C. S. S. Initiative *et al.*, "Common core state standards for mathematics," *Retrieved September*, vol. 15, p. 2010, 2010.

[4] C. McGaughy and A. de Gonzalez, "California diploma project technical report i: Crosswalk study–crosswalk of the intersegmental committee for the academic senate statements of competencies to the common core state standards." *Educational Policy Improvement Center (NJ1)*, 2012.

[5] K. L. Shapley and J. Brite, "Aligning mathematics assessment standards: New mexico and the 2009 national assessment of educational progress (naep). rel technical brief. rel 2008-no. 011." *Regional Educational Laboratory Southwest (NJ1)*, 2008.

[6] D. T. Conley, "Crosswalk analysis of deeper learning skills to common core state standards." *Educational Policy Improvement Center (NJ1)*, 2011.

[7] C. Flowers, D. Browder, and L. Ahlgrim-Delzell, "An analysis of three states' alignment between language arts and mathematics standards and alternate assessments," *Exceptional Children*, vol. 72, no. 2, pp. 201–215, 2006.

[8] M. Timms, S. Schneider, C. Lee, and E. Rolfhus, "Aligning science assessment standards: New mexico and the national assessment of educational progress (naep). issues & answers. rel 2007-no. 021." *Regional Educational Laboratory Southwest (NJ1)*, 2007.

[9] R. Reitsma, B. Marshall, and T. Chart, "Can intermediary-based science standards crosswalking work? some evidence from mining the standard alignment tool (sat)," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 9, pp. 1843–1858, 2012.

[10] S. A. Sutton and D. Golder, "Achievement standards network (asn): an application profile for mapping k-12 educational resources to achievement standards," in *International Conference on Dublin Core and Metadata Applications*, 2008, pp. 69–79.

[11] O. Yilmazel, N. Balasubramanian, S. C. Harwell, J. Bailey, A. R. Diekema, and E. D. Liddy, "Text categorization for aligning educational standards," in *null*. IEEE, 2007, p. 73b.

[12] M. Friendly, "Milestones in the history of data visualization: A case study in statistical historiography," in *Classification—the Ubiquitous Challenge*. Springer, 2005, pp. 34–52.

[13] D. E. Knuth, "Two thousand years of combinatorics," in *Combinatorics: ancient & modern*, R. Wilson, J. J. Watkins, and R. Graham, Eds. Oxford University Press, 2013.

[14] B. Johnson and B. Shneiderman, "Tree-maps: A space-filling approach to the visualization of hierarchical information structures," in *Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on*. IEEE, 1991, pp. 284–291.

[15] H. Purchase, "Which aesthetic has the greatest effect on human understanding?" in *Graph Drawing*. Springer, 1997, pp. 248–261.

[16] C. Ware, H. Purchase, L. Colpoys, and M. McGill, "Cognitive measurements of graph aesthetics," *Information Visualization*, vol. 1, no. 2, pp. 103–110, 2002.

[17] M. Ghoniem, J.-D. Fekete, and P. Castagliola, "A comparison of the readability of graphs using node-link and matrix-based representations," in *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on*. Ieee, 2004, pp. 17–24.

[18] F. V. Ham and B. E. Rogowitz, "Perceptual organization in user-generated graph layouts," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 14, no. 6, pp. 1333–1339, 2008.