

NCS Research Day Abstract

Title: A semantically-aware approach to NCS data comparability and interoperability

Jay Greenfield, PhD ¹; Daniela Smith ¹; Michael Keller, PhD ¹; Sophia Kuan ¹; Ranjana Srivastava, PhD ¹; Alexandra Shlionskaya ¹

¹Booz Allen Hamilton, Rockville, MD

Introduction

The National Children's Study (NCS) is building a standards-based Metadata Repository (MDR) in which definitions of data elements, value sets and forms are stored and versioned over time. This approach allows the NCS to construct and maintain logical consistency of the Study metadata, provide proper versioning among data element specifications, and promote general ease of use in accessing and using the study definition and study data in applications and workflows.

Methods and Results

Through programming transformations, we have imported the Master Data Elements Specification (MDES spreadsheet) into the Metadata Repository, and mapped a subset of study variables derived from the MDES to reference terminology (see Figure 1). These concepts are displayed in a navigation page as part of a Data Dictionary, such that they can be browsed and exported into various formats for use. Because the Metadata Repository includes all versions of the study data elements, the Data Dictionary includes a Change Log that compares each new version with the previous one, noting changes in the meaning of data elements among versions.

Conclusions

Establishing consistency, versioning, and meaningful representation of the study metadata allows Study Centers to use local informatics while preserving the comparability and interoperability of study data. The Metadata Repository supports the NCS as an integrated semantically-aware effort, allowing meaningful linkages of data and establishing connections to standards organizations.

Figure 1: NCS variables catalogued in the MDES are mapped to reference concepts and identifiers, making the Metadata Repository ISO-11179 compliant and semantically-aware.

