



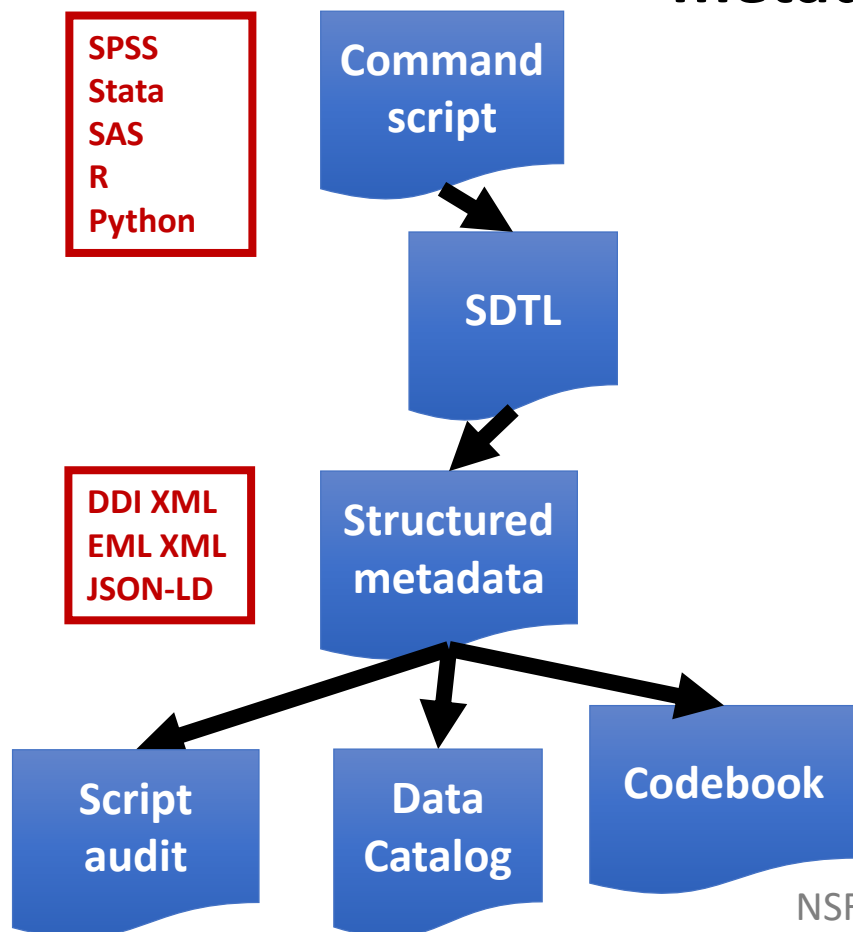
Award #: ACI-1640575

CSSI Element: Continuous Capture of Metadata for Statistical Data

PI: George Alter, Co-Pis: Jack Gager, Pascal Heus, Jeremy Iverson, Hosagrahar V Jagadish, Jared Lyle, Ornulf Risnes, Dan Smith

Institutions: University of Michigan, Algenta Technologies, Metadata Technologies North America, Norwegian Centre for Research Data

C²METADATA
Continuous Capture of Metadata



Automating the creation of variable-level provenance metadata from scripts used with statistical software.

1. Data management scripts are translated into Structured Data Transformation Language (SDTL).
SDTL is non-proprietary and machine actionable.
2. SDTL is translated into natural language.
3. A history (“derivation”) for each variable is added to a structured metadata file.
4. Variable histories are included in data catalogs, codebooks, and tools for auditing scripts.