



Brett L. Abrams
Senior Archivist
Electronic Records Division

E-ARK Conference

December 2016

NARA & SIARD 2



ELECTRONIC RECORDS DIVISION

NATIONAL
ARCHIVES

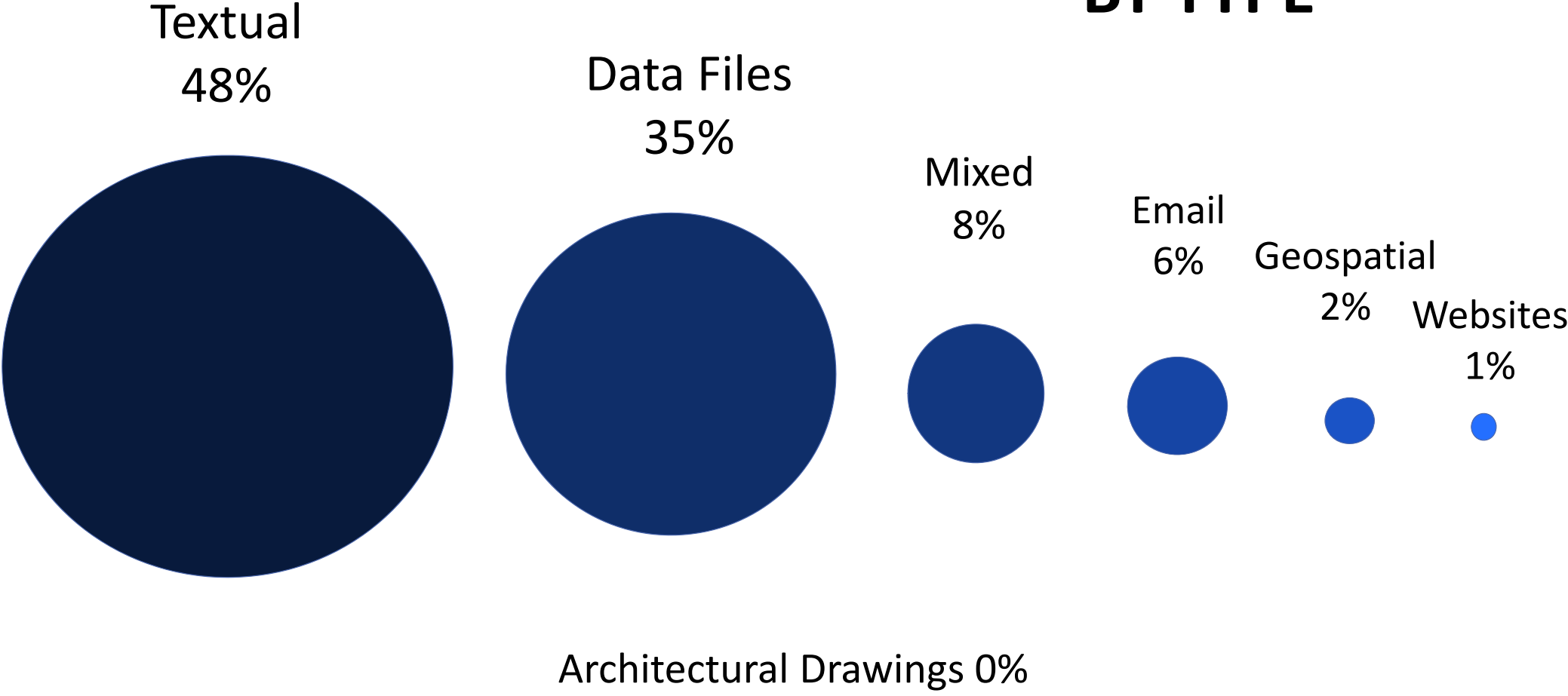
Started in 1969

Accessions, processes, arranges for preservation, and provides access to born-digital federal records scheduled for permanent retention at the National Archives

Records include: databases, presentation formats, geospatial, email, web records in a variety of specified formats

Holdings: 1,118 series from over 100 federal agencies; over 800 million unique files; over 397 Terabytes; 943 series contain data files

ACCESSIONS IN PROGRESS BY TYPE



DATA & DATABASES

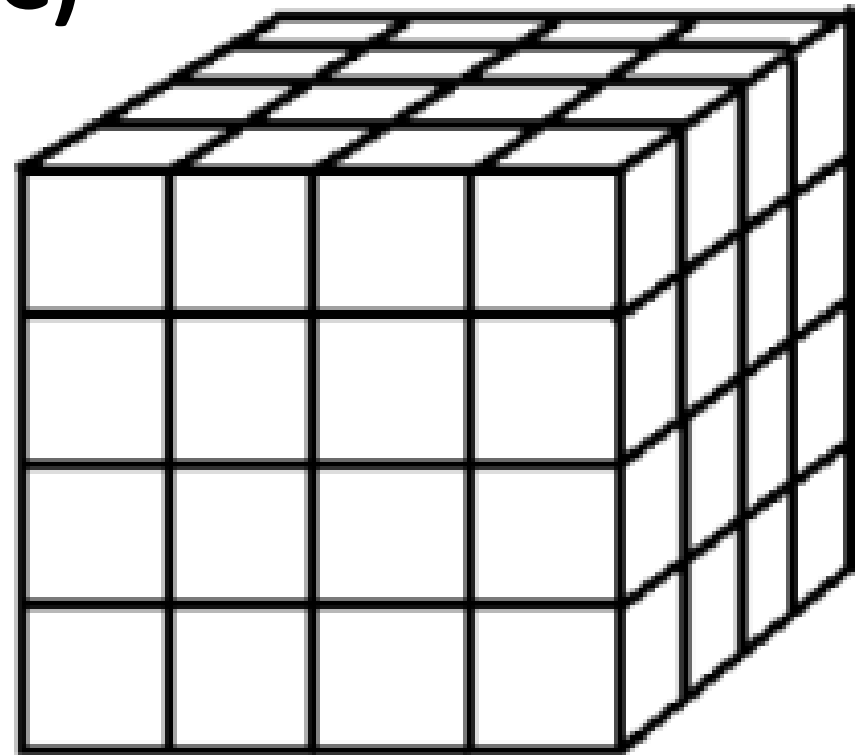
Code of Federal Regulations 36 1235 and NARA 2014-14 lay out the requirements for federal agencies transferring records to NARA

CURRENT CHALLENGES

- Conversion to flat files
 - Time consuming process for agencies
 - Loss of relational structure
- Documentation of the data files
 - Reliant upon separate documentation to explain individual files

ARCHIVAL ELECTRONIC RECORDS INSPECTION AND CONTROL (AERIC)

- The Archival Electronic Records Inspection and Control (AERIC) system provides automated verification of records
- Compare the actual content of the electronic records to description of the content of those records as represented by the documentation



VERIFICATION OF FLAT FILE RECORDS

Fixed-Length Verification

Counts the number of records in which the width of the value is less than the width of that field

Range Verification

Matches the values in a field against a consecutive set of values

Distinct Values Verification

Counts the Frequency of all values in a field

Code List Verification

Matches the values in a Field against a list of acceptable values

Min/Max Verification

Identifies the lowest and highest value in a field

Mandatory Verification

Counts the number of records in which the field is blank

Primary Key Verification

Identifies the number of non-unique values in field or combination of fields

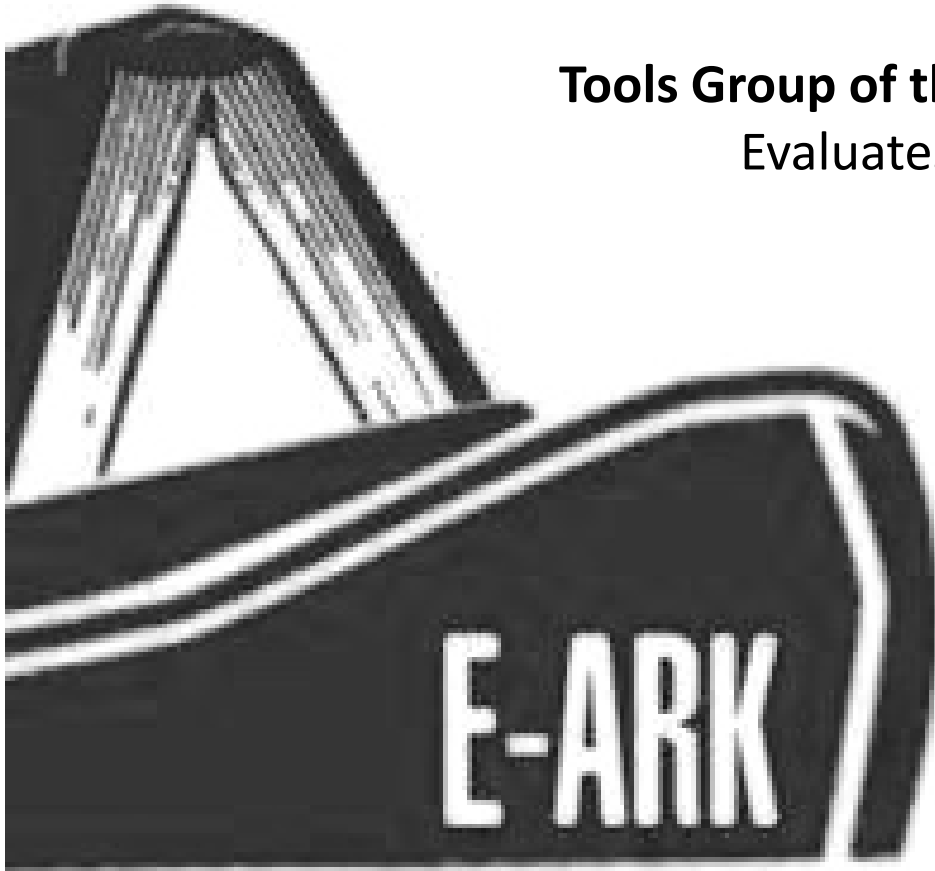
Datatype Verification

Shows the values in a field that do not match the data type designated in the AERIC Record Layout

TESTING SIARD 2.0

Tools Group of the Electronic Records Division

Evaluates tools to enable us to verify variety of records



Looked at SIARD 1 and now evaluating SIARD 2.0

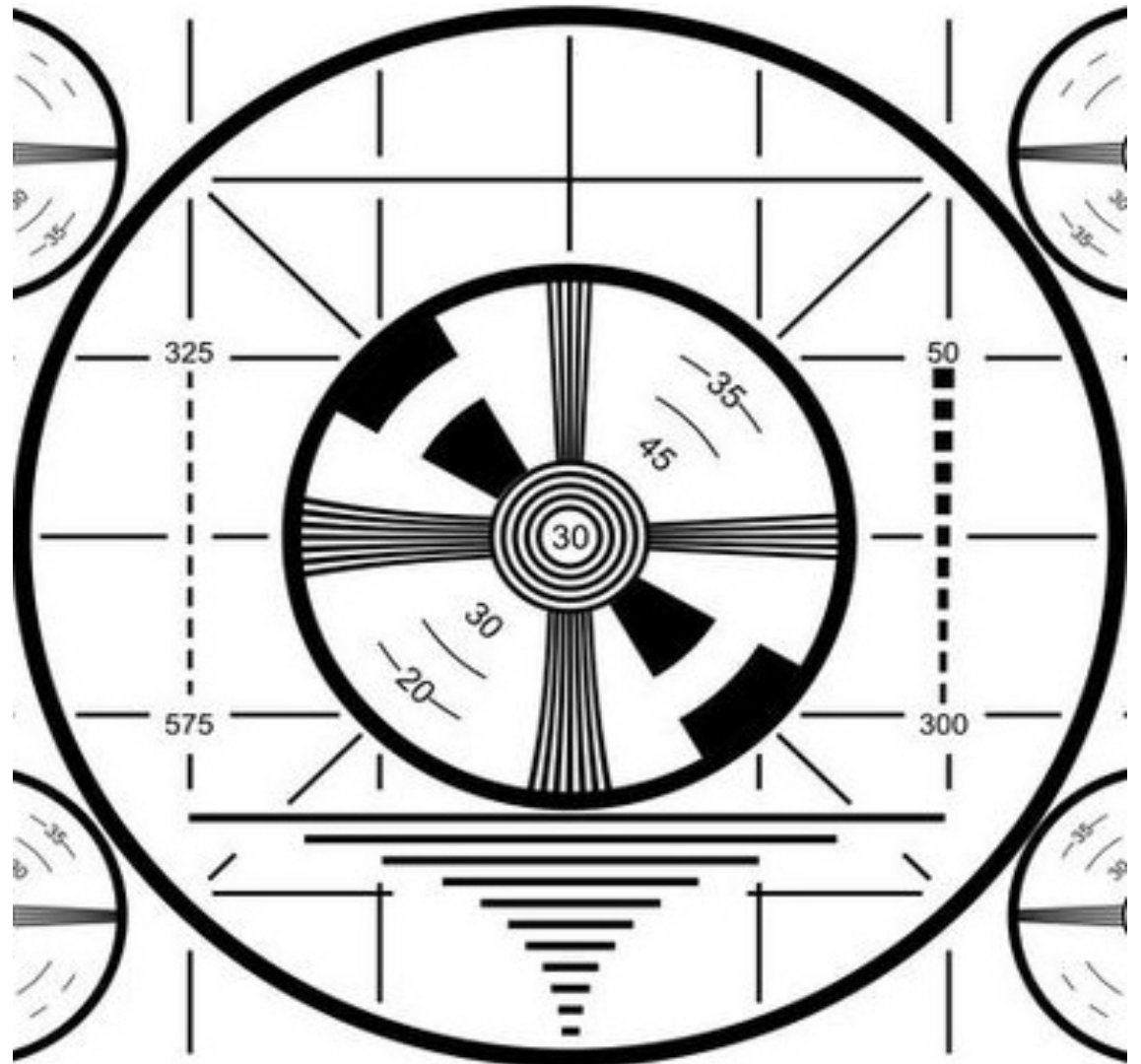
Hopes for the format format:

1. Maintain relationships of tables
2. Easier to export
3. Easier to access

PILOT

Comparing validation, preservation, verification and access steps of relational databases in SIARD Format against those records in ASCII flat format

- Ease of Transfer and Creation
 - Ease of Conversion
- Preservation Process
 - Ease of Access



THANK YOU