The Lifecycle of Social Science Research Data:

Improved Discovery through better Metadata *and* Search Tools

One research data management output is metadata preparation and its exposure to external search tools

Tomorrow's precise, metadata-driven, research data-focused **Search & Discovery**

Research study is conceived and planned, methodologies selected, funding sources explored
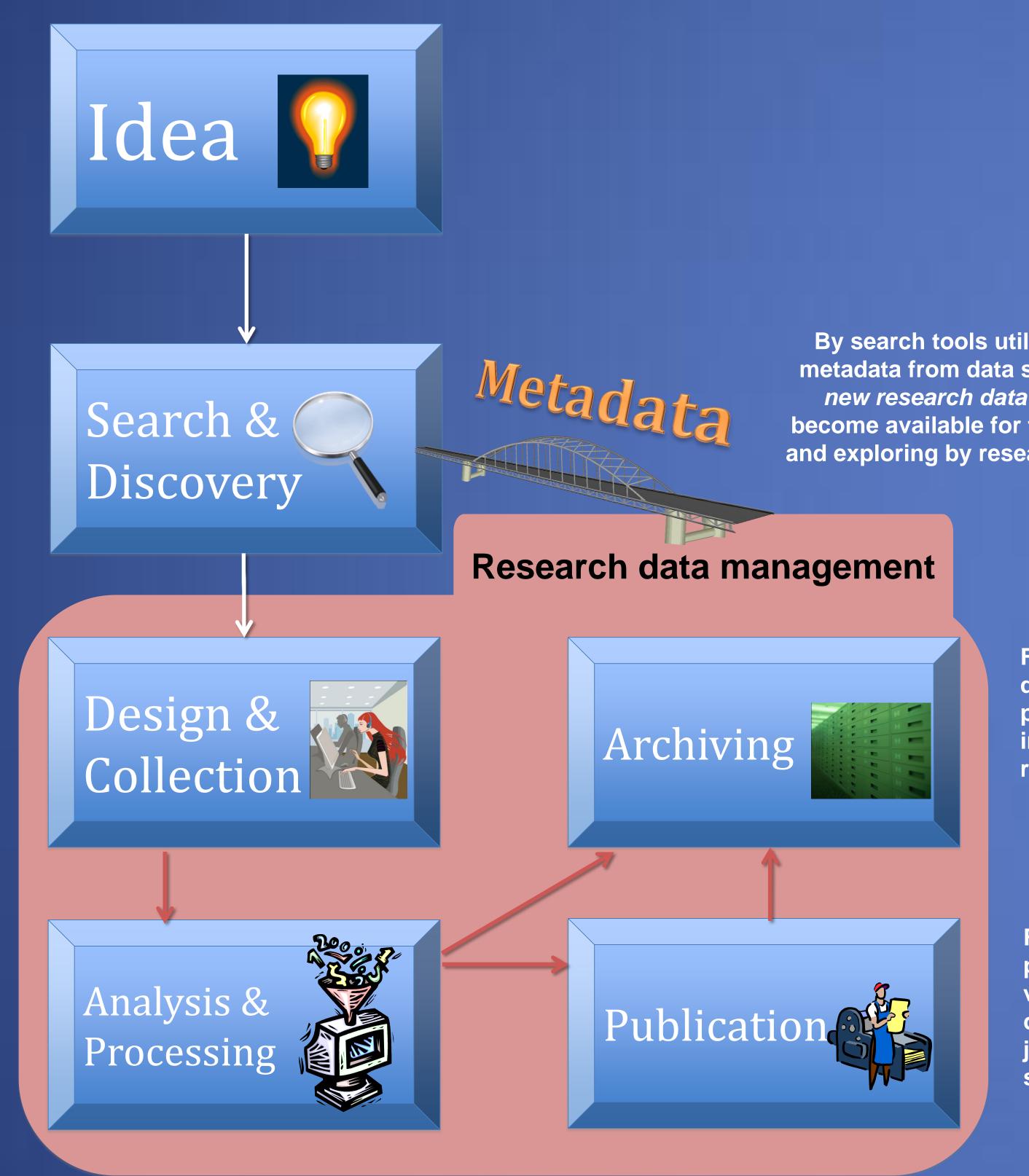
Idea

Existing data sources are sought and explored – also happens for basic research needs

Search & Discovery

Metadata

By search tools utilizing metadata from data stores, *new research data* **can become available for finding and exploring by researchers**

Research data management

Research instruments are designed; data are collected through surveys, interviews, etc. – and from existing data sources
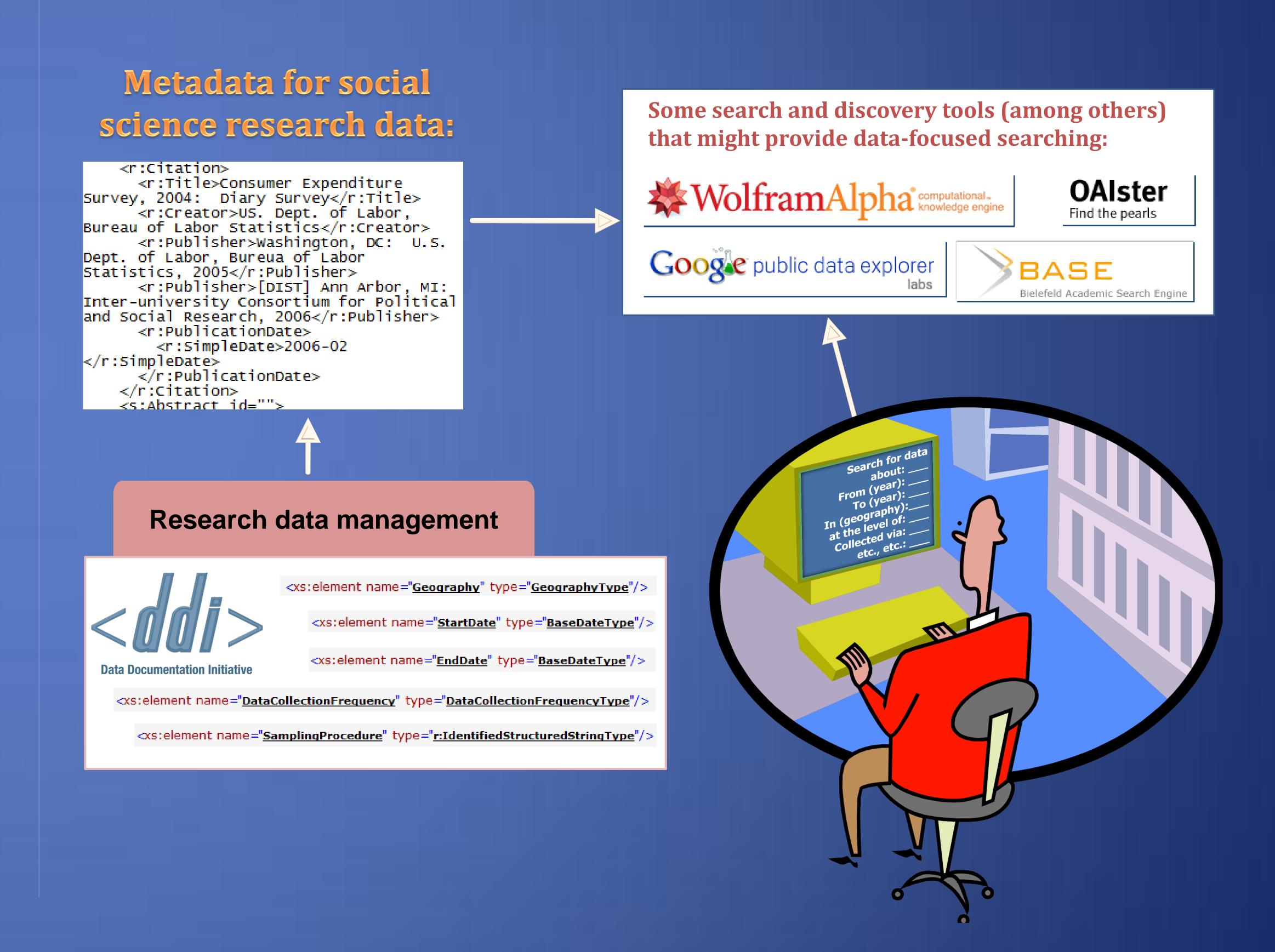
Design & Collection

Archiving

Final datasets are deposited for long-term preservation – e.g., into institutional or domain repository

Collected data are merged, cleaned, analyzed, sub-setted, coded, harmonized, linked, etc.

Analysis & Processing

Publication

Final datasets are made publicly accessible – e.g. via researcher's and/or department's and/or journal publisher's web site

Metadata for social science research data:

```
<r:Citation>
    <r:Title>Consumer Expenditure
Survey, 2004: Diary Survey</r:Title>
    <r:Creator>US. Dept. of Labor,
Bureau of Labor Statistics</r:Creator>
    <r:Publisher>Washington, DC: U.S.
Dept. of Labor, Bureua of Labor
Statistics, 2005</r:Publisher>
    <r:Publisher>[DIST] Ann Arbor, MI:
Inter-university Consortium for Political
and Social Research, 2006</r:Publisher>
    <r:PublicationDate>
        <r:SimpleDate>2006-02
</r:SimpleDate>
    </r:PublicationDate>
</r:Citation>
<s:Abstract id="">
```

Some search and discovery tools (among others) that might provide data-focused searching:

WolframAlpha computational knowledge engine

OAIster Find the pearls

Google public data explorer labs

BASE Bielefeld Academic Search Engine

Research data management

```
<ddi>
Data Documentation Initiative

<xs:element name="Geography" type="GeographyType"/>
<xs:element name="StartDate" type="BaseDateType"/>
<xs:element name="EndDate" type="BaseDateType"/>
<xs:element name="DataCollectionFrequency" type="DataCollectionFrequencyType"/>
<xs:element name="SamplingProcedure" type="r:IdentifiedStructuredStringType"/>
```

Search for data about:
From (year): ___
To (year): ___
In (geography): ___
at the level of: ___
Collected via: ___
etc., etc.:

William C. Block Director	Stefan Kramer Research Data Management Librarian	Jeremy Williams Programmer/Analyst
Cornell Institute for Social and Economic Research (CISER)		

Cornell University
Cornell Institute for Social and Economic Research