UCLA

Describing archives a content standard

Multi level optimum + added value EAD / EAC descriptions in the absence of DACS guidance

Case references (DACS 14.23, 14.23A, 14.26)
Courts of the United States

level 1 (district / trial + variants) courts

US federated states (50, 4 formally titled Commonwealths | 94 courts)

federal district (1 | 1 of 94 courts)

Individual federated states (50, 4 formally titled Commonwealths | 1,912 courts)

unincorporated organized territories (4 | 3 of 94 courts)

unincorporated unorganized territories (~1 , 11 total)

Court of International Trade

United States Court of Federal Claims

military courts (unified system applicable to the Army, Navy-Marine Corps, Air Force and Coast Guard | 4 courts)

military commission (administered by the Department of Defense, Office of the Convening Authority, Office of the Chief Prosecutor, Office of the Chief Defense Counsel, Military Commissions Trial Judiciary | Guantánamo Bay Cuba)

level 2 (intermediate appeals + variants) courts

US regional circuits (12)

Individual federated states (50, 4 formally titled Commonwealths | 104 courts)

United States Court of Appeals for the Armed Forces (worldwide jurisdiction)

United States Court of Military Commission Review

level 3 (final appeal) courts

United States supreme court (1)

Individual federated states (50, 4 formally titled Commonwealths | 58 courts)

and all non US jurisdictions

The UCLA Civil Rights Project (collection 0787) has been processed using localized archival description rules for court cases. These rules adhere to a micro-descriptive scheme designed to enhance and further develop new methods and tools to

- **extract** and **assemble** archival authority descriptions
- **enhance** methods for **matching** and **combining** records describing the same case
- **develop** methods for accommodating descriptive data
- **add** geographic coordinates to places where court related documents circulate
- **develop** timeline or map rendering of chronological biographies or histories
- **enable** users to query social professional networks
- **develop** graphical displays of complex, dense networks
- **develop** graphical displays of organizational charts
- **develop** sequential displays of legal organizations / actions merging or dividing

This document discusses current limitations existing archival descriptive frameworks - DACS - and the work of UCLA staff to compensate for the gaps in the professional standard. We look forward to starting a conversation within the archival community (especially those currently working with litigation records) on revisions and amendments.

DACS of course aims to not only make the records more easily discovered and accessed but also, and at the same time, build unprecedented resources that provide better access to the socio historical contexts. DACS rules should be expanded and refined to include case naming conventions that show the people involved in the litigation, law firms and their attorneys, judges and their court appointed experts and advocacy groups - all of whom created the records.

We advocate entering into Memoranda of Understandings with NARA Federal Records Centers, law schools within affiliated campus systems (4 within the UC system) and a hub (Civil Rights Litigation Clearinghouse) to construct international standard authority records for cases centered within geographic regions. Using improved DACS standards, it is possible to construct linked data that will be an exciting and scalable experiment in extracting information from a cross collaborative collection in ways that are impossible with smaller, individualized collections.

Merging the social, the data and the substantive subject matter a civil rights collection can become one of the most exciting media collaboratories with instantaneous applications for teaching, data curation and public service.

The archives / special collections profession is committed to the discovery, location and use of distributed historical court case records. So is UCLA.

Persons use archival records as primary evidence for the lives and work of historical persons and the events in which they participated. Courts, and court systems produce enormous quantities of records which by their nature arrange parties (individuals + organizations) and issues of great historical / primary value. Most US court records are unavailable in digital form before 1998. Many federated states lag significantly behind the Federal courts in providing access to court records. Proprietary publishers currently impose significant monetary costs and physical limits on access to court records.

Each calendar year, closed court case files in the United States represent 367,000 archival appraisal opportunities.

UCLA

Figure 1
Cases opened + closed in the 3 levels of United States courts 2011

	opened cases	closed cases	ratio
US district courts 2011	289,252	303,158	1.05
US appeals courts 2011	55,126	57,357	1.04
US supreme court 2011	9,066	7,827	0.86

Source Administrative office of the US courts judicial business 2011

The predominant type of record the United States courts create and maintain is a case file, which contains a docket sheet and all documents filed in a case. Approximately 10% of the litigation in any given year in the United States is handled in the federal judicial system. 90% of United States litigation moves through 1.17% of cases filed in US district courts were adjudicated through a trial.

Court case paper and digital records reflect integral collaborations

The majority of court records in the United States are paper. These records are the product of localized aspects of interaction common in litigation. The paper records evolved hand-in-hand with organizational work practice common to court systems, which are governed by localized operating (procedural) rules.

Although paper is the ubiquitous artifact in support of collaborative work - even in “high-tech” research environments - archivists familiar with legal materials understand their inherent nature frequently disrupts the principle of respect des fonds as a basis of archival arrangement and description. As court cases accessions rely more on born digital materials (disk images that preserve the creator’s arrangement, and XML expressed documents), local repositories may adopt processing scheme that reassert the respect des fonds principle.

Collections that incorporate legal documents show that paper documents are written on, read from, stacked, and filed many times over the records continuum. The documents were brought to meetings, exchanged, distributed, and discussed. The legal communities in many nations (and especially the US and the Commonwealth countries) are a ‘club’ whose members are persons with: 4 years of general college education, 3 years of legal education in core and elective subject matter approved by a national professional association and state (bar) accreditation which licenses them to appear in courts.

These legal professionals have particular interest in the facts and circumstances surrounding the emergence of legally significant social imbalances or transgressions of enforceable rules. They develop and manipulate policies to overcome social imbalances and/or allegations of rule breaking. And they are trained in applying data to develop corrective policy criteria to induce courts to provide remedies (in civil matters, the remedy is based in fairness; in criminal matters, the remedy is based on justice).

ISO 15489 Information and documentation - Records management, defines record keeping metadata as: Data describing context, content and structure of records and their management through time.

Australia's localized version of ISO 15489 (widely considered to be 10-30 years ahead of the US with respect ti implementing archival standards) has a definition that adds the concept of 'domains' and operational terms:

[record keeping metadata] is structured or semi-structured information that enables the creation, management and use of records through time and across domains. **Recordkeeping metadata can be used to identify, authenticate and contextualise records and the people, processes and systems that create, manage, maintain and use them**.

Their work involves advice, research, investigations and data gathering which is incorporated into legal actions tailored to the jurisdiction and the particular problems presented in the case. The resulting legal records blend objective assessments and subjective arguments in documents - examples are briefs, memoranda - which are used to support motions (requests for judicial action). The courts - through its judicial officers - relies on these collaborative documents for its decision-making.

The production of legal documents is heavily dependent on the exercise of professional judgement. To ensure the adequacy of these judgements, the law firms and courts are organized so that these documents go through extensive and thorough review. Most archival attention is focused on published opinions and related documents. However, the publication rates range between 59%-93% in United States appeals courts (intermediate)[1]. The percentages are far lower in United States district courts.

Legal work supports the interweaving of discussion with activities centered around documents in a group situation, such as marking up or reading through. These social processes surrounding the discussion of the document are vital, but it is equally critical that interacting with legal documents not disrupt this complex social interaction, or force participants to break away from any ongoing collaborative activity. Legal documents also provide at-a-glance information so that people who are co present can discern the activities of others with respect to the document being discussed. Teams of reviewers (at many levels of a court, law firm or advocacy organization) can and do sit around desks arguing and discussing the documents in question, marking sections, pointing out parts, and exchanging pages in ways that reflect and support the local interaction processes.

Because of these interactions and collaborations - adequate archival processing of paper legal materials requires a sufficient archival encoded record (EAD / EAC) to capture the social dimensions of the records, but rules that ensure the process is as minimal and compact as possible (and free from the need for technical support and skills generally above that of a practitioner). The original use of the document helped their users to tell whether the person next to them was turning toward or away from the document, helped them to see approximately where in the document they were, and told them whether a colleague was flicking through pages, or is setting it aside. In part, it is the physicality of the paper medium and the ease with which it could be manipulated which gives a sense of the activities of others and helps the group coordinate and focus their discussion.

Of course the original interactions and its nuances are lost at the archival stages of a paper document's continuum, but a good set of descriptive rules (DACS) and a strong commitment to ultra simple, micro descriptive techniques importantly restores a measure of the social networks involved. In collections with blended digital content, these paper documents can be displayed alongside audio and video recordings of the proceedings.

What makes court cases distinctive among records series is the fact that large swaths of civil legal subject matter can be expressed in class actions (where a discrete number

1 Table S-3 Judicial Business 2011 - Types of Opinions or Orders Filed in Cases Terminated on the Merits (2011) - Administrative Office of the United States Courts

of named plaintiffs, generally less than 100 represent a class of persons that can number in the many thousands). In cases involving certain subjects (even where they are not procedurally a true class action) invite the participation of numerous third parties - individuals, corporations and advocacy groups or associations.

Civil cases invite multiple actual parties, and third parties

Many controversies brought to the judiciary affect more than the immediate adversaries. Through judicial recognition and ingenuity and persistence, third parties find representation in the judicial process in two major ways: sponsorship (initiating, financing, and conducting) litigation and the submission of friends of the court (amicus curiae) briefs. Private interest groups, federal, state, local and municipal governments and their elected / appointed components - School Boards, City Councils also participate. Collection 0787, and our prospective collaborations with other institutions / repositories involve cases in school segregation and reapportionment, fair housing and education - each is among the most important and complex domestic issues of the past 50 years. These cases, and their associated judicial documents are among those most at risk of remaining hidden in the Federal Records Centers. There are implementation gaps and serious cost barriers associated with the Public Access to Court Electronic Records PACER - a United States government legal documentation system. As a result, there is the potential to lose archival opportunities in connection with perhaps as many as 300,000 court cases each year[2].

UCLA's collection has micro-descriptive EAC records for **79** **school desegregation** cases (nature of suit code 440 and/or 448), 3 voting (nature of suit code 441), 1 equal employment (nature of suit code 442), **35** **fair housing** cases (nature of suit code 443), **22** **education cases** (nature of suit code 448). A review of other existing collections UCLA collections may produce additional examples of these (or other) case types.

The official clearinghouse for civil rights litigation[3] has acquired / coded **49 / 9** **school desegregation cases** (nature of suit code 440 and/or 448), 13 / 11 voting (nature of suit code 441), 2825 / 211 equal employment (nature of suit code 442), **124 / 0** **fair housing cases** (nature of suit code 443), **26 / 1** **education** cases (nature of suit code 448).

We propose using 2 cases, the Los Angeles school desegregation case (Crawford @ UCLA) and the San Francisco case (NAACP SF Branch) as starting points.

2 Approximately 10,000 cases per year in 2011 are classified in the aggregate as other statutory actions. The distinctive nature of suit codes relating to United States civil rights cases are 440 other civil rights 441 voting 442 employment 443 housing accommodations 444 welfare 445 disabilities employment 446 disabilities other 448 education. Where the case was not tried, the official court records are eligible for destruction subject to the review of an appointed independent expert committee who advises the Judicial Conference Committee on Court Administration and Case Management CACM (records subcommittee). The civil rights battlefields in public education has shifted away from remedying segregation to the preservation of affirmative action policies, especially in admissions practices and funding.

3 Housed at the University of Michigan Law School; the repository collects documents and information from civil rights cases in specific case categories across the United States. It is available to scholars, teachers, students, policymakers, advocates, and the public, to allow greater understanding of civil rights litigation. It has a faculty director appointed within the Law School, it receives technical support from MPublishing, a component of the University of Michigan Libraries, current support from Washington University Saint Louis Law School, historical support during the first 5 years from Washington University's Center for Empirical Research in the Law, and government support from the National Science Foundation grant SES-0718831, for a project titled "The Litigation Process in Government-Initiated Discrimination Suits," which funded document acquisition and coding for employment cases initiated by the United States Equal Employment Opportunity Commission from 1997 to 2006 (post PACER implementation).

Court records are distributed

A principle or rule established in a previous American legal case is either binding on or persuasive for a court or other tribunal when deciding subsequent cases with similar issues or facts.

In the United States (and in the 54 independent nations which were former British colonies, the Commonwealth of Nations), parties litigate disputed questions of fact before a tribunal. The tribunal is completely passive. It does not conduct independent investigations. It cannot recognize, without the parties' assistance, all information contained in the evidence presented. Instead, it determines the appropriate legal sanctions to impose based solely on the information identified and presented by the parties. This is the archive dynamic: information presented in the past continues to inform the judiciary across the decades through the application of an Anglo American legal doctrine.

When a federal case is filed, it is held in the US district court for approximately 15 years. Thereafter, the case is transferred to one of the Federal Records Centers in **17 cities**. The National Archives charges the courts a storage fee for holding these documents; in 2010 the fee was $6.2 million. In 2010, the **US Judicial Conference Committee on Court Administration and Case Management CACM**, estimated that half of the closed case files currently in storage at federal retention centers (for which storage fees are charged) would be immediately eligible for disposal under the existing retention schedule with the National Archives and Records Administration.

We discovered from primary sources (persons on the CACM) that **all suit code 440: Civil Rights – Other (411,699 existing cases) were provided a permanent proposed disposition**. These case files document the national experience as well as citizens' rights and government accountability in the arena of civil rights[4].

These records are held in the Federal Records Center[5] and to a lesser degree in archives and manuscript libraries, large and small. Persons looking to thoroughly research a case (or trying to assess the quantity of materials available) may need to search many different archives, following clues, hunches, and leads to find the records relevant to their topic (and it is likely that - based on the current state of archival description - many records will remain undiscovered). Staff at the National Archives and Records Administration have been engaged with the policies and operating processes of US court records since 2011[6].

With UCLA's collection 0787, it is possible to envision creating an International Standard Authority Record for individual court cases, and linking that data to other repositories. However, this work would be easier if DACS standards were improved upon.

---

4 Herren, Omar. Records disposition authority N1-021-10-2. The appraisal read "a random selection of cases from eight different courts around the country was examined for this suit code. Cases involved a variety of police brutality issues, treatment of the disabled and children, prison conditions, home schooling, bilingual education, reverse discrimination, privacy rights, separation of church and state, abortion rights, political solicitation, and additional areas of conflict between the government and the citizen. Large proportions of cases terminated both before and after pre-trial were seen as historically significant either in themselves or as part of a totality of similar cases, because they document important trends and patterns in American social and cultural experience. Issues that were later assigned specific suit codes, such as disability access and voting, were first covered here; undoubtedly, there are additional civil rights developments currently being captured here that will later have their own codes. Many of these issues are related to policy shifts, including treatment of the mentally disabled and physically handicapped, the relationship between privacy and technology, church-state relations, discrimination against a variety of minorities and/or majority populations, minority education, prisoner housing, and similar interactions among population groups and between individuals and groups and the government at all levels. While a portion of these cases may be seen as routine, retention of all of the cases is the only feasible method of documenting the judicial process with regard to the broad range of statutes in this area of civil rights, their use and their effect." Documentation provided by Margo Schlanger, director Civil Rights Litigation Clearinghouse, University of Michigan Law School on 08 31 2012.

5 The NARA Record Groups are: RG 21 Records of the District Courts of the United States, RG 116 Records of the Administrative Office of the United States Courts,, RG 123 Records of the United States Court of Claims, RG 172 Records of the United States Commerce Court, RG 205 Records of the Court of Claims Section, RG 267 Records of the Supreme Court or the United States, RG 276 Records of the United States Court of Appeals, RG 308 Records of the U.S. Tax Court, RG 321 Records of the U.S. Court of International Trade

6 Arian Ravanbakhsh, Electronic Records Policy Specialist

United States district court (level 1)

The case records of the U.S. circuit courts are organized into series according to the types of jurisdiction exercised by the courts. Each subset of case records contains case files and other jurisdictionally specific materials, including dockets, order and judgment books, and other documents related to that aspect of the courts' business. Case files contain the original papers submitted by the parties as well as other documents issued by the courts in those proceedings. Case files may also contain transcripts of testimony or exhibits introduced at trial, although the parties usually retained exhibits. Case files are arranged either chronologically by date of filing, numerically by case number, or in some cases, both chronologically and by case number. Case names and numbers can be found in docket books and related indexes. Within each case file, records may be organized in loose chronological fashion, by type of proceeding or document, or by some other method.The clerks' organizational schemes have been largely preserved in Record Group 21 at the various regional branches of the National Archives[7].

Appeals court (level 2)

The historical records of the twelve regional U.S. courts of appeals have been accessioned by the National Archives and Records Administration (NARA) and are deposited at its regional branches as part of Record Group 276. Newer records are located in the offices of the respective clerks of court, in the circuit libraries, or in one of the federal records centers. The records of the U.S. Court of Appeals for the Federal Cir- cuit have not yet been accessioned by NARA. Once accessioned, these records will make up Record Group 504 at the National Archives in Washington, D.C. The records of the U.S. courts of appeals are made up of general and administrative records, as well as case materials.

Court of appeals case records generally consist of three types of records: (1) transcripts of the official record from the lower court or federal agency, including printed copies of the proceedings of the trial or hearing as well as copies of papers that were filed in the trial; (2) documents submitted to the court of appeals by the parties, including briefs, appendices, petitions, and answers; and (3) papers produced during the appellate proceedings, including stipulations, agreements, motions, correspondence between the court and the parties, judgments, orders, decrees, mandates sent to district courts or received from the Supreme Court, and judicial opinions.

The organization of court of appeals case records varies significantly from court to court. Some clerks collected all of the various materials related to specific cases into "case files." Other courts filed the three types of case records listed above as distinct sets of records. Whether stored in integrated case files or by type of record, court of appeals case records are arranged by case number, which may be obtained from a court's alphabetically arranged indexes or docket.

7 Jonathan White. Guide to research in federal judicial history (2010). Federal Judicial Center Federal Judicial History Office at 36

Several courts of appeals separated matters related to the courts' original jurisdiction (such as the review of decisions by federal agencies or commissions, petitions for various writs or rehearings, and motions to vacate sentences or stay court orders) from the court's appellate jurisdiction records. Some courts also set apart materials from certain types[8].

Final appeals (level 3)

The historical records of the Supreme Court make up Record Group 267 at the National Archives in Washington, D.C. Record Group 267 contains general records, case records, the records of the clerk, and the records of the marshal of the Supreme Court. Nontextual records, including sound recordings, are held at the National Archives at College Park, Maryland.

The Supreme Court's appellate case records consist of case files, indexes, opinions, copies of mandates sent to lower courts, and miscellaneous papers in undocketed cases.

Appellate case files contain transcripts of proceedings from lower courts, petitions for writs of error or certiorari, briefs, motions, orders, judgments, decrees, mandates, agreements, bonds, depositions, writs, and other papers. Appellate case files from 1792 to 1933 are arranged in one numerical sequence (case numbers 1–38,700). Since 1934, appellate case files have been arranged by term and thereunder by case number.

The Supreme Court's appellate jurisdiction records include a numerically arranged collection of manuscript opinions and revised printed appellate opinions, including both majority and dissenting opinions. The National Archives published an Index to the Manuscript and Revised Printed Opinions of the Supreme Court of the United States in the National Archives, 1808–1873 (Washington, D.C.: The National Archives, 1965), which includes citations to both appellate and original jurisdiction cases[9].

Judicial collections (chambers papers)

The preservation of a judge's personal papers and records establishes a foundation for historical studies that document the role of the federal judiciary in American life. The personal papers of a judge offer a perspective and a level of detail not available in the official records of the courts. Research collections based on judges' papers are often the most valuable source for illuminating the judicial process. A significant portion of the historical record of the federal courts is lost because there are few memoranda of understanding that provide for a systematic transition of these papers to archival repositories. The existing collections are highly distributed.

The chambers papers of a federal judge remain the private property of that judge or the judge's heirs, and it is the prerogative of the judge or the judge's heirs to determine the disposition of those papers. Neither federal statute nor the policies of the Judicial

8 Jonathan White. Guide to research in federal judicial history (2010). Federal Judicial Center Federal Judicial History Office at 38-42

9 Jonathan White. Guide to research in federal judicial history (2010). Federal Judicial Center Federal Judicial History Office at 43

Conference of the United States make any provision for the preservation of federal judges' papers. Judges' staffs or the clerks of court cannot determine where the papers go, and the National Archives cannot accept the collections as part of the records of the courts. There are no court funds available for the preservation of judges' papers, and the federal records centers do not provide temporary storage of judges' chambers papers[10].

Current DACS rules pertain to courts as institutions, but do not address the naming conventions of court cases. How can the social identities and networks of litigation participants be identified in archival records?

Civil and human rights, environment, constitutional interpretation, energy, healthcare liability, insurance, financial regulation and shareholder actions, consumer protection, fair housing, foreclosures, bankruptcies, education, labor, employment, international trade, immigration, public records are all legal subjects that frequently invite the participation of third parties (persons not directly involved in the main case, but who are affected by the outcomes).

The relationship between access to information and access to justice is at the heart of civil litigation. Private parties (and public agencies who litigate) have many available points of access to their information when they decide to file a lawsuit. Intellectually and historically the connections between these parties and their information sources are opaque. Lawyers can talk to their clients and other willing witnesses. Documents can be gathered. Specific rules or practices may sometimes limit the information needed before a formal lawsuit is brought. Many controversies brought to the judiciary affect more than the immediate adversaries.

Two ways by which third parties find representation in the judicial process are through the sponsorship (initiating, financing, and conducting) of litigation, and the filing of amicus curiae briefs.

In describing these cases then, we paid careful attention to bringing each of the parties unique identities into the record (at least within a non viewable metadata field).

Description of social identities needs to be associated with itemized records

If rendered in a consistent form and included in electronic indexes standardized metadata terms can become a powerful tool for researchers to discover materials related to that topic. It is a local decision as to which names, terms, and concepts found in a description will be included as formal access points, but repositories dealing with court cases should provide them in all types of descriptions. Such indexing becomes increasingly important as archivists make encoded finding aids and digital content available to end users through a variety of repository-based and consortial online resource discovery tools using less precise full-text searches[11].

One of the archival principles on which DACS rests requires that, in processing legal collections, it is necessary to describe at an optimum level the court cases as well as the places these cases are adjudicated.

10 *See* A Guide to the Preservation of Federal Judges' Papers Second Edition (2009), Federal Judicial History Office Federal Judicial Center at 1.

11 DACS at xviii + xix

This fundamental principle is

Principle 6: The principles of archival description apply equally to records created by corporate bodies, individuals, or families.

The documents that are the product of the functions and activities of organizations may differ in extent, arrangement, subject matter, etc., from those that result from the activities of individuals or families. While there may be valid reasons to distinguish between them in the workflow of a repository, the principles of archival arrangement and description should be applied equally to materials created by individuals, families, or organizations.

The most important function of any descriptive rule is the ability to create access points for researchers who will search an archival finding aid electronically[12].

When using online finding aids, users prefer to find information quickly via search, rather than navigate the finding aid's hierarchy as archivists may have intended. In particular, a number of studies[13] show that many users prefer to find information immediately by using a search function rather than by navigating the finding aid's hierarchy. Because of "More Product, Less Process" imperatives in archival processing, realistically there will be only 1 opportunity in the processing of most collections to create a record that is useful for locating the social interactions.

Search engines such as Google produce results according to a type of statistical distribution where a high-frequency population is followed by a low-frequency population which gradually "tails off." This tail is often long, meaning that the total number of infrequently used keywords outnumbers the total of the top ten keywords.

We were particularly guided by Michelle Light's recommendation to provide only a level of description necessary to allow a user to reach within 3-4 feet of analog (container) materials based on search results[14].

In practice that involved establishing a simple division (and physical grouping) of case documents into pleadings, then supplemental case information. Archival processors use DACS to construct titles for series, subseries, and folders. Until the methods for producing born digital finding aids improve, the existing DACS rules set needs to address an ability to record numbers of long tail, unique keywords (in a metadata repository field) while producing title name segments that won't distract a human (who probably won't read the title much anyway, having found records relevant to a more specific term that was the basis of their search strategy).

---

12 O'English, Mark. "Applying Web Analytics to Online Finding Aids: Page Views, Pathways, and Learning about Users," Journal Western Archives volume 2.1 2011.

13 Prom, Christopher J. "User Interactions with Electronic Finding Aids in a Controlled Setting." American Archivist 67.2 (2004): at 257-258. MetaPress. Web. 20 May 2009; Nimer, Cory. "What Do You Mean It Doesn't Make Sense? Redesigning Finding Aids from the User's Perspective." Journal of Archival Organization 6.4 (2008): at 228. Informaworld. Web. 8 February 2008; Scheir, Wendy. "First Entry: Report on a Qualitative Exploratory Study of Novice User Experience with Online Finding Aids." Journal of Archival Organization, 3.4 (2005): 49-85. Informaworld. Web. 27 March 2009.

14 Light, Michelle. "EAD @ 10 The endangerment of trees," presentation to the California Archivists Conference 2008.

DACS, in its best form, will motivate processors to think carefully about the value of the description in an online environment before investing in the description itself.

A descriptive rule should produce simple text

Plain text (in software) means that any file will incorporate a sequence of characters with no extra typography that has to be interpreted by machines. The best method for entering plain text elements is to produce a string of characters using only a comma as a separator (comma separated values CSV). Adherence to relatively simple file format is widely supported by consumer, business, and scientific software applications. It also serves the archival common use of moving tabular data between programs that natively operate on incompatible (often proprietary and/or undocumented) formats. Almost all software programs support some variation of CSV at least as an alternative import / export format.

Google Instant, the current state of the art search query tool reduces the length of text to produce an effect where a user reads an intended search string faster than they type, which saves time. Predictions help guide users searches, and those predictions avoid pronouns, prepositions, conjunctions, interjections and most adjectives.

At UCLA, we have incorporated this understanding into both strands of the court case naming rules.

How we overcame a DACS deficiency

Court cases (the papers filed by the parties for consideration by the Court (pleadings) and the related supplemental information (everything else, but particularly attorney work product, data, legal research, and out of court communications) touch 2 of the 6 primary categories of access points: A documentary forms, B functions + activities.

These documents also touch on 5 DACS elements: Title Element (shared) (2.3) Extent Element (2.5), Scope and Content Element (3.1), Administrative/Biographical History Element (2.7, Chapter 10).

To work well as an access point, DACS asserts any terms associated with court cases should have these elements (though not necessarily in a simple inventory list without an associated metadata repository linked to those records).

We focused on the creation of metadata (a secondary DACS consideration). We based our construct on **ISO 25964 Thesauri and Interoperability with other vocabularies** and **ISO 11179-5 Information technology specification and standardization of data elements**.

This standard is slanted differently (towards machines rather than humans) than DACS. These standards favors the design of web pages that are machine readable as well as human readable, so that computers can query and connect web content automatically, thus improving our ability to exploit the information potential.

Arrangement of EAD / EAC metadata
smallest scale (container listing)

<unitid countrycode="us" label="docket number"> (federal cases)
<unitid countrycode="us.california" label="docket number"> (state cases)

<unittitle>National Association Advancement Colored People San Francisco (city) branch NAACP (+35 parties) against California Education Department CA ED (+24 parties): complaint</unittitle>

Identified plaintiff is first listed party
First names are not included
Party appears most frequently in procedural history lists (ex Lexis)
count number of parties total, these are enumerated in <odd>

substitute against for 'v.'

Identified government defendant is the largest administrative division
Identified corporate defendant is the largest administrative division
Identified person defendant is functionally described

For example 'Arnold Schwarzenegger" is California governor

<corpname encodinganalog="110" source="lcnaf">United States. District Court (California : Northern District)</corpname>
<address>
<addressline>37°46'52.94" N 122°25'04.62" W</addressline>
</address>
<unitdate type="pleading" normal="19880423">04 23 1988</date>

Arrangement of EAD / EAC metadata
intermediate scale (DACS primary description / finding aids)

```
<dsc type="court cases">
    <c level="series">
        <did>
        <unitid>Series 1</unitid>
        <unittitle>pleadings</unittitle>
        </did>
        <did>
        <unitid>Series 2</unitid>
        <unittitle>case file</unittitle>
        </did>
        <scopecontent>[...]</scopecontent>
        <bioghist>[...]</bioghist>
    </c>
```

This is where we stored all of the parties litigation identities, taking care to separate firstnames and last names and including occupational information with dates.

```
        <odd>[...]</odd>
```

This is where we stored all of legal authorities which were the basis of the parties claims.

```
        <mandates>
            <mandate>
            <term>[...] <term>
            <mandate>
        <mandates>
```

Some discussion about Global Legal Information Network GLIN (US Library of Congress) records

GLIN records use the base standard XML and are drafted using Resource Description Framework RDF specifications[15]. There are surprisingly few records available in GLIN - 7 judicial decisions dated 2006-2009. That suggests that the creation of these records is a very new phenomenon or has been discontinued because of resource constraints. We are mindful that over description can make processing a slow affair, and the method we used to create XML records (within Oxygen 14) relies on micro descriptive elements that are likely far less extensive than most users of DACS would agree are acceptable (DACS is aligned with Metadata Encoding and Transmission Standard METS headers).

We decided the GLIN adoption of the field 'title' was opaque (even in the same original case, parties often express the headers differently). We like the fact that our localized naming convention depends more on the frequency with which a party's name appears in proprietary databases like LEXIS. Until social network and archival content query tools become more available, the Google ngram viewer is a useful free starting point to assess the usefulness of a term. The 'Perry' in the example record (**footnote 15**) refers to the Texas governor. We are convinced most researchers would be more inclined to seek out functional titles - who has sued the Texas governor (and why?) rather than proper names.

We liked the division of parties into plaintiffs and defendants. The use of the loan abbreviation et al - very common among legal practitioners - nevertheless obscures by collapsing the networked identities of the numerous parties. Moving forward, enumerating those parties in the EAD tag <odd> is the best mechanism for making sure those names are associated with actual records.

We thought it important - if for no other reason that existing archival workflows are extremely unsettled, in many cases relying on exports of columned data - that proper names be divided into separate 'last name' + 'first name' fields - with the proviso that in many pragmatic searches, the first name can be omitted where processing speed is a factor.

---

15 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:skos="http://www.w3.org/2004/02/skos/core#"
xmlns:rdfs="c://rdf-schema#"
xmlns:glin="http://www.glin.gov/Ontology#">

<glin:JudicialDecision rdf:about="http://content.glin.gov/summary/186358">
<dc:identifier>186358</dc:identifier>
<dc:title>League of United Latin American Citizens v. Perry</dc:title>
<dcterms:coverage rdf:resource="http://content.glin.gov/jurisdiction#237" />
<dcterms:issued rdf:datatype="http://purl.org/dc/terms/W3CDTF">2006-06-28</dcterms:issued>
<glin:decisionNumber>05-204; 05-254; 05-276; 05-439</glin:decisionNumber>
<glin:court rdf:resource="http://content.glin.gov/court#1" />
<glin:plaintiffs>League of United Latin American Citizens, et al.; Travis County, Texas, et al.; Eddie Jackson, et al.; GI Forum of Texas, et al.</glin:plaintiffs>
<glin:defendants>Rick Perry, Governor of Texas, et al.</glin:defendants>
<dcterms:modified rdf:datatype="http://purl.org/dc/terms/W3CDTF">2006-11-22T21:22:58Z</dcterms:modified>
<glin:publicationInfo rdf:parseType="Resource">
<glin:publication rdf:resource="http://content.glin.gov/publication#141" />
<dcterms:issued rdf:datatype="http://purl.org/dc/terms/W3CDTF">2006-06-28</dcterms:issued>
<glin:issueNumber>Volume 548 (forthcoming)</glin:issueNumber>
<glin:fullText rdf:resource="http://content.glin.gov/summary/186358/fullText/143539" />
</glin:publicationInfo>
</glin:JudicialDecision>

We also agree that the field ‘decision number’ is equivalent to a docket number - and while meaningless to almost every researcher - this number is the best key for arranging all the records in a particular case.

On a secondary (2d link) - each GLIN summary record displayed permits users to download the RDF/XML copy. GLIN does not utilize a permanent URL schema, so that links may require significant maintenance if embedded in UC / OAC (or other) finding aids.

Some discussion about Civil Rights Clearinghouse CRC (University of Michigan Law School) records

CRC records have 6 segments: **case profile**, **issues + causes of action**, **case details**, **additional resources**, **documents** (dockets + general documents), **people** (judges, monitors / masters, lawyers (for each party).

CRC records are presented in a XHTML format that is produced + maintained by The University of Michigan Library's MPublishing group. We are thankful to the clearinghouse director, Margo Schlanger for providing us access to the site, and for agreeing to work with UCLA in the future to provide links, and perhaps a distribution point for authority records.

Because it is tailored to specifically this class of record, the CRC site is a gold standard. In just 1 example of the many user oriented features - the CRC allows users to register and create a ‘compartment’ for storing records that are of particular interest. If coupled with Google Analytics - this compartment can prove to be an invaluable asset for collection development and management.

CRC adopts the combined court (in Bluebook format) + docket number as one key to finding case specific information. UCLA adopted the docket number identifier - we created ‘humped’ expressions where letters are in lower case, omitting all punctuation.

The CRC also uses their own case identification scheme for example ED-CO-0001 to delineate the type of case (education), the state/territory (Colorado), and the number in the CRC sequence (first=0001). This enables them to gather social information, particularly among the attorneys and judges who are involved in dozens, if not hundreds of cases. We love this scheme - but the addition of another identifier, and the possible need to transform the scheme to different (smaller than states) administrative levels made us pause in adopting this, or a similar linking tag. Nevertheless, we will link back in the finding aid to these records (which are permalinked) using the Encoded Archival Context tag @xlink:href to designate absolute Uniform Resource Identifiers URI - to specify locators (URLs), names (URNs), or both.

We coded approximately 5% of the collection’s folder level content with this enhanced EAC metadata as part of phase 1 of the Arcadia project.

The CRC segment **issues + causes of action** are within UCLA’s scheme (at the intermediate level) under the EAC tag <mandates>.

The CRC segment **case profile** (**summary**) is within UCLA's scheme (at the largest - collection / finding aid - level) under the EAD tag <scopecontent>. We recognize the CRC's work by incorporating their summaries for the 11 cases that overlap with collection 0787, and we provide the CRC's preferred citation (based on the legal profession's Bluebook standard), for example Lauren Cutson. Case Summary 05 26 2005, Jackson v. Artisan (No. 88-C-64, E.D. Wis., July 19, 1990), available as document no. JC-WI-0002-0001, at http://clearinghouse.net.
The CRC segment **case details** is within UCLA's scheme (at the intermediate level) under the EAD tag <bioghist>. Note here the concept of 'creator' (in a narrow archival sense) actually means the collective contributors to the digital + paper (analog) case file.

The CRC segment **people** is within UCLA's scheme (at the largest - collection / finding aid - level) under the EAD tag <odd>.

We are mightily impressed with the CRC, and look forward to partnering with them in phase 2 of the Arcadia project (or in some other sustainable future capacity).