COMMENTS REGARDING
DRAFT MAINE DIGITAL COURT RECORDS ACCESS RULES

Chief Justice Saufley, Senior Associate Justice Alexander, and Associate Justices Mead, Gorman, Jabar, Hjelm, and Humphrey:

The Maine Judicial Branch (the "MJB") has recognized its role in balancing the public's right to access information related to the justice system against the expectations of privacy held by those individuals who interact with the judicial system to resolve disputes and seek justice. Section 8-C of Title 4 recognizes the inherent authority of the Maine Supreme Judicial Court (the "SJC") to issue rules that "determine any other processes or procedures appropriate to ensure ***adequate preservation, disposition, integrity, security, appropriate accessibility and confidentiality*** of the electronic records."[1] Pursuant to this authority, the SJC has proposed new rules to govern the public's access to digital court records (the "Rules").[2]

As proposed, the Rules fail to construct the "comprehensive framework for public access to digital state court records" they set out to provide,[3] and unnecessarily create risks of privacy harm for persons who come to the court seeking to protect their rights. More specifically, the proposed Rules improperly burden litigants with the responsibility to mitigate disclosure risk and lack well-established data privacy protections. Simply put, the Rules fail to provide a comprehensive framework for public access to digital court records and unnecessarily create risks of privacy harm for persons who come to the court seeking to protect their rights.

As discussed in greater detail below, the MJB should ensure that the Rules (1) appropriately require the MJB to assume the responsibility of mitigating privacy harms resulting from unauthorized disclosure of personal information and (2) adopt, or require the MJB to adopt, well-established data privacy principles and procedures.

We prefer the MJB delay implementation of the Rules to further research and revise the Rules in light of the issues raised in these comments. At a minimum, the MJB should adopt a phased implementation plan that allows this important evolution of court administration to continue while also providing additional time to minimize the significant harm to Maine citizens and others who avail themselves of the Maine Court System that may ensue under the MJB's current approach. It is in the best interest of justice that any rules adopted by the MJB to govern access to digital records put forth a truly comprehensive framework.

[1] 4 M.R.S. § 8-C (emphasis added).

[2] *See* Draft Maine Digital Court Records Access Rules.

[3] Id. at Rule 1.

I. THE RULES PLACE THE BURDEN OF PROTECTING PRIVACY INTERESTS ON FILING PARTIES AND CREATE BARRIERS TO ACCESS

In attempting to protect privacy interests, the Rules would rely on a series of broad categorical approaches to define filing and disclosure that may ultimately unnecessarily restrict the public's access to court records. For example, while the MJB and Maine Legislature have recognized value in making court records accessible to the public through digital media, the distinction inherent in the Rules' definition of "Court Record" between files maintained by the judicial branch in digital form versus files maintained in paper form creates barriers to access certain files based solely on how they are maintained by the court rather than their content.[4]

As another example of the strain on balancing privacy with access considerations that arise from the Rules, the Rules would place the burden of protecting privacy interests on private parties, including lay people with no prior knowledge of statutory or legal privacy protections. At the same time, the Rules omit any remedies for individuals who suffer unauthorized use or disclosure of their personal information.

The Rules also fail to include mechanisms to hold the MJB accountable to Maine citizens for failing to take appropriate security measures to protect personal information. Such accountability mechanisms are a key facet of protecting the personal information of Maine citizens without unnecessarily restricting access to information. The Rules currently would place primary accountability for protecting the personal information of Maine citizens on filing parties.

The risk of requiring that filing parties redact confidential information and mark pleadings according to whether they may be further disclosed through the digital court records system will in particular create significant challenges to achieving the balance between privacy and access that the MJB seeks to champion.

Under Rule 9 of the Rules, Filing parties bear the responsibility of designating which documents may be disclosed publicly.[5] Filing parties would need to "conspicuously mark" any files related to cases designated as sealed, impounded, or nonpublic with "NOT FOR PUBLIC DISCLOSURE."[6] Placing the burden of identifying what information must be protected from disclosure on filing parties, rather than on the MJB (for example, by relying on technological solutions implemented by the MJB), creates risks that documents not intended for disclosure will be disclosed (or vice-a-versa). It also creates a risk that,

[4] Id. at Rule 2(g) ("'Court record'" means any file, document, information, or data received or maintained by a state court in digital form. . . .").

[5] Id. at Rule 9 ("It is the responsibility of the filing party to ensure that sealed, impounded, or nonpublic cases, documents, and information are redacted and/or submitted to the court in accordance with this rule.").

[6] Id. at Rule 9(a).

pursuant to rule 9(d), filing parties may have their filings rejected if they are improperly labeled.[7]

Because it places the burden of protecting privacy interests on filing parties, and threatens sanctions for parties who fail to meet that burden, Rule 9 would create a barrier to using the digital records system that disproportionately would impact unrepresented parties.

II. THE RULES DO NOT REFLECT APPROPRIATE USE OF TECHNOLOGY SOLUTIONS TO REDUCE BARRIERS TO ACCESS AND SUPPORT THE PROTECTION OF PERSONAL PRIVACY

Some of the above-noted deficiencies in the Rules may stem from the chronic understaffing of the state court system and related concerns about making effective use of limited court resources. For that reason, it is imperative that the MJB thoroughly explore the potential benefits of automated redaction software, and that it do so before finalizing and implementing the Rules.

In a recent white paper, the National Center for State Courts (NCSC) succinctly summarized the important relationship between court privacy policy formulation and the availability of effective technology:

> If a court has no technology capability and few resources, then it must close many of its case types and rely on filer liability (again excepting case types that are closed by statute). If a court has some automated redaction capability, then it can open a number of case types and document types. If it has an advanced automated redaction capability that can reliably protect all specified confidential information in any type of document, then it can open a maximum amount of public case information to public access.[8]

In this context, good technology can facilitate better policy. For many years, cost-effective automated redaction solutions were not available to courts, but NCSC focus groups have determined that the latest generation of redaction software shows great promise.[9] Assuming the MJB were to reach the same conclusion after adequate time and opportunity to explore current technology, the implications would be significant for address some of the Rules' present limitations.

We note also that Tyler Technologies, Inc., a company already contracted by the MJB to assist with computerization of records, recently incorporated automated redaction tools

---

[7] Id. at Rule 9(d) ("If any filed document does not comply with the requirements of these rules, a court shall, upon motion or its own initiative, order the filed document returned, and that document shall be deemed not to have been filed.").

[8] "Best Practices for Court Privacy Formulation," National Center for State Courts (July 2017), at 5, available at https://cdm16501.contentdm.oclc.org/digital/collection/tech/id/876.

[9] Id. at 4 & Appendix B.

into solutions it offers the State of Texas.[10] Tyler Technologies indicates that its redaction capability is "best-of-breed" and "tightly integrated" with its court-focused software solution, Odyssey, [11] to protect data that shouldn't be exposed to the general public. However, it is unclear in the Rules whether this redaction capability is a part of the Odyssey solution chosen by the MJB and, if it is, to what extent the MJB intends to use this solution in protecting the privacy of parties to a proceeding.

III. THE RULES DO NOT REFLECT ESTABLISHED PRIVACY PRINCIPLES THAT WOULD MITIGATE OR ALLOW THE MJB TO RESPOND TO CYBERSECURITY THREATS

Court systems are high-level targets for cyberattacks precisely because court records contain valuable personal information related to individuals and businesses.[12] Unauthorized access to such personal information could cause significant harm to the same, and the lack of safeguards may undermine the public's confidence in the ability of the MJB to protect their sensitive information, deterring such constituents from accessing the court to seek justice.

In parallel with efforts to explore automated redaction technology, the MJB should consider how to adopt and abide by vital data processing principles – such as transparency, data minimization, storage limitation, security, and accountability. Such principles, which inform recent or contemplated privacy legislation in Europe, California, and Congress, are rapidly becoming the standard against which privacy practices are judged. And their importance in this context is particularly significant; as the Joint Technology Committee formed by NCSC, the Conference of State Court Administrators (COSCA) and the National Association for Court Management (NACM) has observed, "[i]f EU data privacy standards were applied to US courts, the sensitive nature of court data would warrant the most stringent protections," and courts should therefore have "a game plan for preparing to comply with similar legislation in the US." [13]

The need for the Rules to reflect well-established privacy principles is not academic, but instead, grounded in the reality of and disruption caused by cyberattacks.[14]

---

[10] Press Release, "Tyler Technologies Enhances eFileTexas and re:SearchTX Portals to Protect Sensitive Case Information: Redaction tool protects sensitive information for filers and Texas court clerks" (Dec. 20, 2018), available at https://tylertech.irpass.com/Tyler-Technologies-Enhances-eFileTexas-and-re:Sear.

[11] Odyssey Case Manager Overview Brochure, https://www.tylertech.com/Portals/0/OpenContent/Files/3294/Odyssey-Case-Manager-Overview-Brochure_.pdf last visited (March 26, 2019).

[12] Judge Herbert B. Dixon, Jr., "Cyberattacks on Courts and Other Government Institutions," ABA Groups, Judicial Division (Jan. 17, 2019), ¶ 15, available at https://www.americanbar.org/groups/judicial/publications/judges_journal/2018/summer/cyberattacks-courts-and-other-government-institutions/.

[13] *See, e.g.*, "GDPR for US Courts," Joint Technology Committee Resource Bulletin (Sept. 19, 2018), at 4, available at https://ncsc.contentdm.oclc.org/digital/collection/tech/id/876/.

[14] *See, e.g.*, "Information Systems and Cybersecurity – Annual Report 2018," Administrative Office of the U.S. Courts, available at https://www.uscourts.gov/statistics-reports/information-systems-and-cybersecurity-

Cyberattacks on courts and other public institutions are well documented.[15] These attacks typically fall into one of four categories - denial of service attacks, phishing, ransomware, and spyware – any one of which would compromise the principles and goals enumerated in the Rules.[16] Direct access to the MJB is not the only avenue for cyberattacks; judicial records may also be compromised through other government branches. Courts around the nation have faced many of the privacy and protection issues now before the MJB,[17] and the MJB would do well to learn from them.

The lack of Rules to guard against any of the four common types of attacks compromises the Courts enumerated goals and the ability of the Judicial Branch to credibly safeguard the personal information under its control. As it stands, conspicuously absent from the Rules is any mention of use of "processes or procedures appropriate to ensure ***adequate preservation, disposition, integrity, security, appropriate accessibility and confidentiality*** of the electronic records" that the Legislature has recognized are within the

---

annual-report-2018; "JTC Resource Bulletin: Responding to a Cyberattack," Joint Technology Committee, NCSC (Feb. 17, 2016), available at https://www.ncsc.org/~/media/Files/PDF/About%20Us/Committees/JTC/JTC%20Resource%20Bulletins/Responding%20to%20Cyber%20Attack%202-26-2016%20FINAL.ashx.

[15] *See, e.g.*, Dixon, *supra* note 11; "2018 Verizon Data Breach Investigations Annual Report" at 41, Verizon (2018), , available at https://enterprise.verizon.com/resources/reports/DBIR_2018_Report.pdf (Globally, the public sector faced over 22,000 security incidents with 304 confirmed data disclosures. Personal information accounted for 41% of the data compromised.); Laila Kearney, "With Paper and Phones, Atlanta Struggles to Recover From Cyberattack," Reuters (March 31, 2018), available at https://www.reuters.com/article/us-usa-cyber-atlanta/with-paper-and-phones-atlanta-struggles-to-recover-from-cyber-attack-idUSKBN1H70R0 (cyberattack in which the City of Atlanta's computer network was infiltrated and crippled by malicious actors); Kieran Nicolson, "State Juror Pool Data Breach Exposed Social Security Numbers," Denver Post (Aug. 8, 2017), available at https://www.denverpost.com/2017/08/08/state-juror-pool-data-breach-exposed-social-security-numbers/ (external exposure of information held by the Washington State Administrative Office of the Courts); "Washington State Courts Office Suffers Data Breach," Government Technology (May 9, 2013), available at https://www.govtech.com/security/Washington-State-Courts-Suffers-Data-Breach.html (external exposure of jury files held by the Colorado Judicial Department containing names and other data of 41,140 individuals).

[16] *See* Brian McLaughlin, "Cybersecurity: Protecting Court Data," PA Times (May 26, 2017), available at https://patimes.org/cybersecurity-protecting-court-data/. (1) Denial of Service attacks usually overwhelm servers to a specific site, preventing legitimate users from accessing services or records. (2) Phishing is one of the more common attacks and solicits personal information from unsuspecting users through e-mail that appears legitimate and requests users to enter items such as user names or passwords to compromise accounts. (3) Ransomware infects software and locks access to data until a ransom is paid. Cyberattackers access vulnerable systems through phishing e-mails, drive-by downloading, and unpatched system vulnerabilities. (4) Spyware infects a computer by producing pop-up ads, re-directing browsers and monitoring a user's internet activity. To the extent that the MJB's system is interconnected with other government systems, the risk of exposure to attacks increase.

[17] *See, e.g.*, "Judicial Branch's Computer System Attacked With Ransomware," NBC Connecticut (Mar. 9, 2018), available at https://www.nbcconnecticut.com/news/local/Judicial-Branchs-Computer-System-Attacked-With-Ransomware-476402943.html; Brian McLaughlin, "Cybersecurity: Protecting Court Data," PA Times (May 26, 2017), available at https://patimes.org/cybersecurity-protecting-court-data/; Ricardo Lopez, "Minnesota Courts Cyberattack Underscores Growing Threat," Star Tribune (June 25, 2016), available at http://www.startribune.com/minnesota-courts-cyberattack-underscores-growing-threat/384398871/.

ambit of MJB authority and responsibility.[18] For example, the Rules do not require the MJB to publish a privacy notice informing Maine citizens about how it uses and discloses personal information. Nor do the Rules indicate whether the MJB is required (instead of merely permitted, at its discretion) to adopt security measures to protect such information. The Rules also do not establish, or require the MJB to establish, a response protocol in the event of suspected or actual unauthorized access to personal information. The lack of reference to, or use of, these well-established data security protocols in the Rules flies in the face of generally established practices across all industries, as well as practices specific to court administration.[19]

The obvious deficiencies in the Rules raise significant questions about the process by which the MJB formulated the Rules. Our measured research unearthed several significant issues, the most pressing of which we outlined above. The Rules do not describe the research or authorities relied upon by the MJB in developing these Rules or the SJC's philosophy, and there is no information on the MJB website that would assure the citizenry that the MJB fully recognizes its responsibility to safeguard the personal information under its control, even as the MJB embraces its role in balancing the public's right to access court information and the expectations of individual privacy. As a cursory matter, some questions that come to mind – and for which neither the Rules nor the MJB website provide answers – are as follows:

- How does the MJB plan to address actual literacy and technology literacy deficiencies in potential users of Odyssey, the MJB's chosen software solution?
- Why has the MJB chosen to shift the risk of unintended disclosure to Odyssey users, particularly given that some users will not have the actual or technology literacy to use the system proficiently?
- Does to MJB intend to engage a cross-section of stakeholders during Odyssey implementation to ensure that indigent, rural, and other disenfranchised or low-use users of legal services continue to have a clear and accessible path to justice?
- How does the Court plan to safeguard against the specific types of cyberattacks most likely to occur?
- What procedures and plans are in place to allow the Court to continue to function if (when) a cyberattack is successful?

[18] 4 M.R.S. § 8-C.

[19] *See, e.g.*, "Information Systems and Cybersecurity – Annual Report 2017," Administrative Office of the U.S. Courts, available at https://www.uscourts.gov/statistics-reports/profile-administrative-office-us-courts-annual-report-2017. (In this report, the Administrative Office of U.S. Courts informed Congress that it had "developed and launched a mandatory IT security 'scorecard,' enabling courts to conduct annual IT security self-assessments. This resource helps court units identify IT security vulnerabilities, channel resources to address them, and bolster the Judiciary's overall IT security posture.")

- What incident response mechanisms are in place to allow affected individuals to mitigate any undue harm resulting from unauthorized access to the personal information they entrusted to the MJB?
- What technology mechanisms are in place to track individuals' access to the court system that would aid in identifying potential perpetrators of cyberattacks if (when) they occur?
- What incident reporting protocols are in place to track and learn from any unauthorized access and create a body of knowledge to support effective court practice in this area?

The Rules fall far short of providing a comprehensive approach and unnecessarily creates the risk of harm for persons who come to the court seeking to protect their rights. . This barebones approach to such an important evolution of court administration in Maine does not appear to leverage the existing body of knowledge of effective court practice.

IV. Recommendations and Conclusion

The following recommendations are modest actions that we urge the MJB to consider to mitigate or prepare an effective response to the issues set forth above.

1. The MJB should delay implementation of the Rules to further research and revise the Rules in light of the issues raised in these comments or adopt a phased implementation plan to allow this important evolution of court administration to continue while also providing additional time to minimize the significant harm to Maine citizens and others who avail themselves of the Maine Court System that is inevitable under the MJB's current approach.

2. Rule 9 should be amended to redistribute the burden for ensuring adequate labeling of filings containing sealed, impounded, or nonpublic information on the MJB. The Rules also should implement an accountability mechanism that requires MJB to adequately protect the personal information of Maine citizens.

3. To the extent MJB intends to use the automated redaction technology offered by Tyler Technologies, the MJB should revise the Rules to particularly state how and when it intends to leverage the benefits of automated redaction or issue an order that requires the MJB to adopt and maintain a privacy policy that does the same. The Massachusetts Supreme Judicial Court has issued such an order, attached here as Appendix I for reference.

4. The Rules should set forth, or require the MJB to adopt, a privacy policy and well-established privacy procedures, including an annual audit to identify system and process weakness. The NCSC has published best practices for courts in drafting

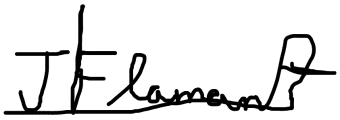
privacy policies, including a model privacy policy, [20] attached here as Appendix II for reference.

While the Rules are intended to further an interest in public access to court records, they also recognize the importance of protecting personal privacy. The Rules fail to balance those two interests because they improperly burden litigants with the responsibility of mitigating the risks that personal information will be disclosed without authorization and fail to incorporate well-established data privacy mechanisms. We urge the SJC to further consider the key challenges to balancing access and personal privacy highlighted in these comments before adopting its final rules.

In offering the above comments and recommendations, we are acting solely in our personal capacities as attorneys specializing in, among other areas, privacy law. We are not submitting these comments on behalf of any client, any organization, or our respective law firms.

Respectfully,

Krystal D. Williams, Esq.
Pierce Atwood, LLP



Julian B. Flamant, Esq.
Hogan Lovells US LLP

Vivek J. Rao, Esq.
Pierce Atwood, LLP

[20] Thomas M. Clarke, et al. "Best Practices in Court Privacy Policy Formulation," NCSC (2017).

THE COMMONWEALTH OF MASSACHUSETTS

Suffolk, ss. **Supreme Judicial Court**

ORDER

Order Re: Protection of Personal Information

Introduction. Massachusetts General Laws c. 93H provides that the judicial branch shall adopt rules or regulations to safeguard certain nonpublic personal information relating to residents of the Commonwealth, the improper or inadvertent disclosure of which could create a substantial risk of identity theft or fraud. This Order governs the security and confidentiality of personal information as defined by c. 93H in the Judicial Branch. It is designed to safeguard the personal information of all individuals, including nonresidents. It shall apply to the appellate courts, trial courts, court administrative offices and court affiliates, which shall be in compliance by September 1, 2010.

Definition. Under G. L. c. 93H, personal information consists of a resident's "first name and last name, or first initial and last name, in combination with any one or more of the following data elements that relate to such resident:

a. Social Security number;

b. driver's license number or state-issued identification card number;

c. financial account number, or credit or debit card number, with or without any required security code, access code, personal identification number or password, that would permit access to a resident's financial account.

Chapter 93H provides that personal information "shall not include information that is lawfully obtained from publicly available information, or from federal, state or local government records lawfully made available to the general public."

Information Security Program. Each appellate court, the Trial Court and any court affiliate that owns, stores or maintains personal information about an individual shall develop, implement, maintain and monitor a comprehensive, written information security program

applicable to any records containing such personal information. The information security program shall govern the collection, use, dissemination, storage, retention and destruction of personal information. The program shall ensure that courts and court affiliates collect the minimum quantity of personal information reasonably needed to accomplish the legitimate purpose for which the information is collected; securely store and protect the information against unauthorized access, destruction, use, modification, disclosure or loss; provide access to and disseminate the information only to those who reasonably require the information to perform their duties; and destroy the information as soon as it is no longer needed or required to be maintained. Such information security program shall contain administrative, technical, and physical safeguards to ensure the security and confidentiality of such records.

Every information security program shall include:

(1) A requirement for notice to the Chief Justice for Administration and Management in the case of a trial court, and to the appropriate Chief Justice in the case of an appellate court, in the event of any incident involving a breach of security[1] of personal information.

(2) Regular monitoring to ensure that the information security program is operating in a manner reasonably calculated to prevent unauthorized access to or unauthorized use of personal information; and upgrading information safeguards as necessary to limit risks.

(3) A regular review, at least annually, of the scope of the security measures. Such review also must be conducted whenever there is an incident involving a breach of security and when there is a material change in business practices that may reasonably implicate the security or integrity of records containing personal information.

(4) Documentation of responsive actions taken in connection with any incident involving a breach of security, and actions taken, if any, to make changes in practices relating to protection of personal information.

Departmental reviews. Each appellate court, court department and court entity shall review the type of personal information it collects and maintains with the goal of identifying any personal information that need not be collected or maintained. Each department will report the results of

---

[1]G. L. c. 93H defines breach of security as "the unauthorized acquisition or unauthorized use of unencrypted data or, encrypted electronic data and the confidential process or key that is capable of compromising the security, confidentiality, or integrity of personal information, maintained by a person or agency that creates a substantial risk of identity theft or fraud against a resident of the commonwealth. A good faith but unauthorized acquisition of personal information by a person or agency, or employee or agent thereof, for the lawful purposes of such person or agency, is not a breach of security unless the personal information is used in an unauthorized manner or subject to further unauthorized disclosure."

this review to the Chief Justice for Administration and Management, or, in the case of the appellate courts and affiliated agencies, to the Chief Justice of the Supreme Judicial Court, within six months.

Computer systems. If personal information is stored electronically, the information security program shall include provisions that relate to the protection of personal information stored or maintained in electronic form. Such provisions shall be developed with the Courts' Chief Information Officers.

Contracts. All contracts entered into by the Judicial Branch shall contain provisions requiring contractors to notify the court of any incident involving a breach of security of personal information, and to certify that they have read this Order, that they have reviewed and will comply with all information security programs and policies that apply to the work they will be performing, that they will communicate these provisions to and enforce them against their subcontractors, and that they will implement and maintain any other reasonable and appropriate security procedures and practices necessary to protect personal information to which they are given access as part of the contract from unauthorized access, destruction, use, modification, disclosure or loss.

MARGARET H. MARSHALL)
)
)
RODERICK L. IRELAND)
)
)
FRANCIS X. SPINA)
) Justices
)
JUDITH A. COWIN)
)
)
ROBERT J. CORDY)
)
)
MARGOT BOTSFORD)
)
)
RALPH D. GANTS)

Dated: January 7, 2010

Best Practices for Court Privacy Policy Formulation

Thomas M. Clarke, Ph.D.
Jannet Lewis
Di Graski
July 2017

Funding for this project was provided by the State Justice Institute Award SJl-16-P-131. The points of view expressed are those of the National Center for State Courts and do not necessarily represent the official position or policies of the State Justice Institute.

Acknowledgements

This project was funded by a grant from the State Justice Institute. The National Center for State Courts (NCSC) is grateful to the participants in the two focus groups who provided input on this white paper. Those participants are not responsible for the views expressed in the white paper and in some cases may not agree with them.

Privacy Policy Focus Group Members

John Bell
Washington State Courts

TJ BeMent
National Association for Court Management

Alan Carlson
Orange County, CA Courts

Tom Clarke
National Center for State Courts

Paul Embley
National Center for State Courts

Di Graski
National Center for State Courts

Rebecca Green
William and Mary Law School

Richard Hoffman
Council for Court Excellence

June Kress
Council for Court Excellence

Ben Moser
Council for Court Excellence

David Slayton
Texas Courts

Automated Redaction Focus Group Members

Jason Bergbower
Colorado Courts

Alan Carlson
Orange County, CA Courts

Tom Clarke
National Center for State Courts

Chad Cornelius
Colorado Courts

Nancy Crandall
Justice Connections

Paul Embley
National Center for State Courts

Di Graski
National Center for State Courts

Kevin Iwersen
Idaho Courts

Henry Sal
Computing System Innovations

Tom Trobridge
Teradact

Table of Contents

Privacy and Public Access Policies
April 2017

How This Report Should Be Used

The National Center for State Courts (NCSC) conducted two facilitated focus groups to produce this report. One focus group considered revisions to the original 2002 COSCA guidelines white paper on privacy and access policies. The second focus group reviewed the status of automated redaction capabilities and assessed the impact of redaction strategies on policy decisions. The membership of the two focus groups only partly overlapped.

NCSC judged that the relationship between policy and redaction capability was a key one and consequently structured this report around it. Readers will still find a separate section that explicitly recommends revisions to the original policy white paper, but this report deliberately asserts the view that policies and redaction capabilities should be considered simultaneously.

This position, and several others in the report such as a strong rejection of "practical obscurity" strategies, are not shared by all the focus group participants. The report is not based on a universal consensus of the focus group members on all issues. It is instead the position of NCSC. In the same vein, it is not endorsed by the Consortium of State Court Administrators (COSCA) as one of its official white papers. Readers should be aware that the report takes a point of view that not all may share.

Background

As state and local courts progressively convert their business processes from paper to electronic formats, policies around remote electronic access to court case information by the public become ever more important. COSCA last addressed this issue comprehensively in 2002 with a report authored by Martha Steketee and Alan Carlson that proposed a model policy for public access[1]. At that time, few courts had implemented electronic filing, so the model policy addressed both manual and electronic access. In the fifteen years since then, courts have learned a lot about living in an electronic world and providing remote access to their case data and documents. Consequently, there is a need to update what we know about this topic and revise the model policy.

---

[1] "Developing CCJ/COSCA Guidelines for Public Access to Court Records: A National Project to Assist State Courts," Martha Wade Steketee and Alan Carlson, October 18, 2002, State Justice Institute (http://cdm16501.contentdm.oclc.org/cdm/ref/collection/accessfair/id/210).

Summary of Current Policies

NCSC has consistently recommended that courts create electronic public access policies before they embark on electronic filing and e-court projects. Once courts have case information in electronic format, the public inevitably wants access to it. Unfortunately, many states only begin the process of creating such policies after they have implemented e-filing. Thus, a few states have electronic case information but still do not have appropriate access policies.

A recent review of the existing state electronic public access policies confirmed that the situation that existed several years ago persists[2]: states exhibit almost no consistency in their policies across most of the key policy decisions and one can find a wide range of policy decisions for almost all the policy aspects. So, a model policy is still relevant. There are some areas of growing consensus and an updated model policy can report that. In other areas, the courts have consolidated around two or three different policy solutions and the model policy can report that as well[3]. For other policy aspects, NCSC is advocating that public access and privacy policies be considered in light of new technology capabilities: autoredaction software that uses machine learning to help courts better balance their twin (and often competing) public policy goals, increased public access to court case records and increased public safety in a "cyber" world.

The Center for Legal and Court Technology at the William and Mary Law School partnered with NCSC for years on a quasi-annual conference on court privacy policies. As the years went by, a gulf slowly opened between what the policies required and what courts could actually, reliably implement, especially using technology. No issue illustrates this problem more than redaction.

Most state policies close a broad range of case types and document types to public access, usually justifying this significant retreat from stated preferences for openness by the difficulty and expense of reliably redacting information that should remain confidential. A few courts redact such information using court staff or county clerk staff, but for most courts that strategy is prohibitively expensive. Likewise, a few courts use automated redaction to at least partly replace human

---

[2] The Council for Court Excellence (CCE) and the National Center for State Courts (NCSC) administered a survey on privacy and public access policies to COSCA in the fall of 2016 and reported on the results at the December 2016 COSCA conference. *See also* a compilation of state court access policies at http://www.ncsc.org/Topics/Access-and-Fairness/Privacy-Public-Access-to-Court-Records/State-Links.aspx (click "Privacy Policies for Court Records").

[3] See Appendix A for the updated model policy for electronic public access to court case records.

reviews, but early vendor products were both expensive and only partly successful in supporting redaction policy requirements.

That left most courts with no alternative than to put the redaction function and resultant liability onto filers. Although widespread, this approach has serious flaws[4]. Audits have found that compliance is not very good. As the proportion of court cases involving self-represented litigants has grown over the last decade or so, the probability that filers will fully comply has correspondingly dropped. That leaves most courts with a very undesirable tradeoff: open case records to the public with significant occurrences of confidential information being disclosed or close an excessive proportion of case records to the public.

Approach to the Problem

This SJI-funded project held two focus groups to address these problems: one concentrating on an update of the model policy and one to assess the state of the art for automated redaction. The two groups had a small proportion of overlapping participants in recognition of the linkage between the two topics. The deliberations of the two groups very strongly reinforced NCSC's belief that what can and should be specified in electronic access policies is constrained or enabled by what can be done well using automated redaction.

To fully understand why this is the case, consider what courts are trying to do. Case documents and associated data never contain information that is all confidential (except of course when they are entirely closed by statute). Some subset of the data or document contains information that should not be released to the public. Protected information may be formally structured, like a social security number, or unstructured, like the name of a crime victim. Similarly, a case document may itself be formally structured with the confidential content in a reliably predictable place and format, or totally unstructured, in which protected information could appear anywhere.

Early versions of automated redaction worked fairly reliably with structured content in structured documents, but otherwise were not very reliable. Consequently, courts had no recourse except to close case types and document types or specify policies that risked revealing confidential information. Being risk averse by nature, courts consistently opted for the former strategy.

For a while this approach seemed to work, but over the last ten years the environment and public expectations have changed dramatically. First, many

[4] "A Contrarian View of Two Key Issues in Court Records Privacy and Access," Tom Clarke, 2016 Future Trends in State Courts (http://www.ncsc.org/~/media/Microsites/Files/Trends%202016/Contrarian-View-Trends-2016.ashx).

government agencies have opened considerable amounts of their data to the public on both the federal and state levels. Executive branch agencies gradually opened up their records in response to FOIA and state public disclosure laws. The federal courts incrementally opened more and more data as well. Second, both for-profit and non-profit organizations have steadily increased pressure on all government agencies to release their data. Third, by putting records on-line, the public has access to them beyond traditional brick-and-mortar hours, and this has provided additional value: it reduces traffic to government facilities while allowing the public more convenience by being able to access files with less disruption to their personal and work schedules.

As experience shows the public benefits of doing so, the public has become more comfortable with this trend. As with many aspects of privacy policy, the public can be remarkably fickle in its desires. As the saying goes, "When they are my data, I want privacy. When they are your data, I want access." Everyone using the Internet (which is everyone) knows that individuals regularly give up aspects of their personal privacy for business benefits that are valued at a few dollars. We often do so even when we know that a business may use our information in ways we would prefer that they don't or we simply don't understand exactly how they will use our data.

This tells us that public attitudes toward privacy and openness are not absolute, but are conditioned by perceptions of value tied to how the data will be used and what we get for allowing them to be used. Some organizations will certainly create products and services with open court data that the public will find valuable and support[5]. The role the media play in public accountability is but one example. So, pressure to open ever more court records continues to build. A separate SJI-sponsored focus group on Courts Disrupted found that such pressure would very likely become overwhelming in the near future[6].

Given this situation, courts desperately need a cost-effective technology solution in the form of automated redaction that can reliably support their policy requirements. Until quite recently that technology was not available, but the latest generation of redaction software is now showing signs of being capable of doing so[7]. Courts in several counties in Florida, Pennsylvania, and Michigan are successfully using the technology, several state courts have pursued Requests for Information or conducted Proofs of Concept, and additional projects are underway or planned in several state court systems to verify the capabilities.

---

[5] *See, e.g.,* https://thistoo.co, an on-line tool for divorcing couples in Ontario offering, among other services, "Real case data to help you quickly understand how your case will resolve."

[6] "Courts Disrupted," Joint Technology Committee (JTC) Resource Bulletin (to be published at http://www.ncsc.org/About-us/Committees/Joint-Technology-Committee/Publications-and-Webinars.aspx).

[7] See Appendix B for an overview of current automated redaction capabilities.

One can imagine a relationship between policy and technology that can be characterized by several maturity levels. If a court has no technology capability and few resources, then it must close many of its case types and rely on filer liability (again excepting case types that are closed by statute). If a court has some automated redaction capability, then it can open a number of case types and document types. If it has an advanced automated redaction capability that can reliably protect all specified confidential information in any type of document, then it can open a maximum amount of public case information to public access.

Unfortunately, highly capable automated redaction products are still relatively new, so their cost is not insignificant. Current vendors mostly use a transaction fee model, so higher volumes of cases incur higher costs. It is highly likely that those transaction costs will decrease dramatically as more courts implement the technology and national volumes rise. Until then, cost will be a barrier for many courts.

One possible solution for the cost problem is to recall that value is ultimately what matters. A very serendipitous characteristic of highly effective automated redaction products is that they can extract from filings almost any information a court might specify. This capability opens a new and potentially very useful business strategy to courts. Such data extraction could be used to drive many different kinds of automated workflows in court business processes, making courts significantly more cost effective[8]. That in turn would mitigate the up-front cost of the redaction software.

The potential for automated workflows to reduce court costs is quite large. Other industries have been able to extract as much as 95% of their labor costs from very similar business processes. NCSC informally estimates that up to 85% of what court clerks traditionally do could be automated in this way. Several recent court reform projects have identified new business processes for case triage and case management that could also be fully automated, adding yet more efficiency. Even the best electronic courts today have barely tapped into this potentially huge pool of cost savings. Through e-payment, e-filing, e-bench, and other technologies, leading courts have reduced their labor costs by at most 10% or 15% to date.

If courts could emulate other industries using data extraction software and automated workflows, one can imagine a quantum leap in value for court customers at the same time courts are reaping big savings that can be partly reallocated to providing better service in other ways. It could be nothing short of a revolution in the service provided to the public. This would come none too soon, since courts are ~~already losing case filings at~~ a rapid rate and seeing significant decreases in public support and legitimacy because of their operational failings.

---

[8] See Appendix C for a discussion of some of the workflows that could and should be automated in this way.

Assuming that courts begin to move up the maturity scale for automated data extraction and workflows, they will quickly recognize that the policy formulation approach used to date will be completely inadequate to the task. Heretofore, courts have convened ad hoc groups to consider and recommend electronic public access policies. Those groups have typically taken months and sometimes years to produce policy recommendations. Those recommendations then enter a court rules process that usually takes at least a year and sometimes longer. At the end, the adopted policies specify in considerable detail the exact requirements as if the capabilities of the court will never change.

This is clearly not agile enough by a large margin in an environment with rapidly changing technology capabilities and even more volatile public expectations. One solution would be to respect that reality by writing rules at a higher conceptual level and moving much of the technical policy detail to locations that can be more readily updated. A corollary strategy would be to make the rules process itself more agile, although how to do so is unclear and certainly outside the scope of this project.

If courts modified their rules-making processes to be more agile, it would pay off in other ways. Major national projects in civil and domestic relations reform have made multiple recommendations for revising court case processes and will probably continue to do so over the next years, as what we know to work grows and technology matures[9]. In many states these processes are enshrined at least partly in court rules that are subject to the same rigid procedures at a time when courts are trying to become more agile. So the benefits of being able to change rules more easily when appropriate would be broadly felt.

Many state courts also need to approach the area of public access policy formulation more broadly than they have in the past. Judges and court administrators often perceive public access as something completely different from policies regarding access by lawyers, case parties, or other justice agencies. They may create yet other policies aimed at use of court data by researchers or third-party data companies. Yet on the technology side, all these policies are implemented by specifying and enforcing business rules using a common technical infrastructure regulating access according to roles and data types. It would be useful for policy makers to become more aware of how their various policies get implemented and ensure that a

---

[9] For a discussion of civil reform recommendations, see the SJI-sponsored Civil Justice Initiative's project website at http://www.sji.gov/civil-justice-initiative-executive-summary/. For an easily accessible summary of the business rules that could be automated, go to "Automated Civil Triage and Caseflow Management Requirements," November 30, 2015 (http://www.ncsc.org/~/media/Microsites/Files/Civil-Justice/Automated%20Civil%20Triage%20and%20Caseflow%20Management%20Requirements%202015-11-30.ashx).

coherent approach is taken that supports successful implementation across the board[10].

Conclusion

The wedding of policy and technology is not unique to public access websites: court leaders pursuing important court innovations in procedure and Government-to-Citizen technology (like Online Dispute Resolution, Fines/Fees/Bail Reform, and Case Triage and Tracking) absolutely depend upon their policymaking bodies to better understand – and reflect on their definitions of court business processes – the art of the possible. This is a challenging era to be running a court system for both good and bad reasons. There are many exciting opportunities to improve court operations and services for the public and also to make being a court employee more interesting and meaningful. There are also high and ever rising expectations by the public that we will make significant improvements as an institution.

Court electronic public access policies both reflect and illustrate these two trends. Courts face big challenges in reliably redacting confidential case information and providing safe, open access to the public, but the ability to do so will pay off in other ways that will greatly help the courts do a good job overall. That obviously creates both opportunities and issues. As courts implement useful new capabilities, other courts will want to take note and leverage what is learned in a timely way to move forward as quickly as possible.

[10] The Florida courts have done a good job of taking a more unified approach to their access rules. *See* http://www.flcourts.org/resources-and-services/court-technology/technology-standards.stml ("Standards for Access to Electronic Court Records" and "Access Security Matrix").

Appendix A
Revised Model Policy for Electronic Public Access to Court Case Records

The 2002 guidelines[11] remain an excellent starting point for a revised model policy. So much of the original document is still valid that it makes the most sense to revise the model language in that report rather than create an entirely new model policy. To ensure that the original report was correctly understood and interpreted, the focus groups included one of the original authors.

Summary of Changes to Model Policy

Although the actual revisions with commentary will be presented below in full, a summary of the changes is provided here to indicate the scope and nature of the changes. The section numbers refer to the original 2002 document. Some sections are renumbered in the revised model policy.

Introduction: Retains the openness principle. Replaces the fundamental distinction between paper and electronic records with a distinction between remote and courthouse access. Asserts a new principle that access should be the same whether remote or in person.

Section 1, Purpose: Reduced the number of objectives to the most important ones and added rationales for each of them.

Section 2, Access by Whom: Revisions were made to focus on public access only. The commentary stresses the need for a common technical infrastructure and coordinated policies for access to court information by various roles.

Section 3, Access to What: The definitions in section 3.10 are still valid and useful. Minor revisions were made to focus the policy on court case records, leaving court administrative records to a separate policy. The remainder of Section 3 was simplified, based upon the assumption that public access is remote, electronic access.

Section 4, Applicability of Rule: The section was extensively revised and combined with Section 3 to (1) identify information where there is a consensus to protect, (2) make explicit the connection between openness and redaction capabilities, (3) move conceptually from document-centric to information-centric approaches, and (4) eliminate "practical obscurity."

---

11 http://cdm16501.contentdm.oclc.org/cdm/ref/collection/accessfair/id/210

Section 5: Renumbered to section 4.0, "Timing of Public Access." The public expects remote access 24/7/365. A separate but important issue is how soon after filing courts make information available to the public remotely. This is one of several areas where policy is closely tied to redaction strategies.

Section 6: Renumbered to section 5.0, "Access Fees." The section was revised to describe the three most common funding strategies and the rationales for using each strategy.

Section 7: The entire "Obligation of Vendors" section was deleted. Several states have developed good contract language for vendors. The Joint Technology Committee (JTC) and the Court Information Technology Officers Consortium (CITOC) will consider developing a model contract.

Section 8: The entire section "Obligation of the Court to Inform and Educate" was deleted. Courts do not need to adopt formal policy guidance on educating litigants, judicial officials, and court staff. Instead, see the new section 3.6 and its commentary, describing the best practice of providing a "one-stop shop" for all public transactions related to court case records (accessing, sealing, expunging, correcting, etc.).

For section content that remains the same, the original commentary is still valid and should be consulted in the original document. Commentary in the revised model policy focuses only on new or revised content.

Assumptions

The world has changed dramatically since 2002. Many courts now operate with completely electronic case records. The revised model policy is designed explicitly to support that new reality and is based upon these assumptions:

1. Courts require electronic filing of all case related information.
2. Courts manage all case related information in case management, document management, and content management systems.
3. Remote public access is supported via Internet and cell phone networks.
4. Remote public access is available essentially around the clock nonstop.

Revised Model Policy for Electronic Public Access to Court Case Records

Introduction

This policy is based on two fundamental principles:

1. Court records are presumptively open to public access.
2. Public access should not change depending upon whether access is remote or at the courthouse.

Section 1.0 – Purposes of the Policy

a. Maximize accessibility of court case records[12].
b. Protect users of the court from harm.
c. Make effective use of court resources.

Commentary: Accessibility is maximized for several reasons: to enhance public trust and confidence, to be accountable, to be transparent, to improve customer service, and to reveal common law. Protection from harm includes individuals, business organizations, government agencies, and the public at large. When balancing openness against potential harm, courts should make the rationales for their decisions explicit. Remote public access is part of a much larger strategy to provide court services online to improve access and convenience and to reduce cost. Cost and efficiency considerations refer to both user costs and court operational costs.

Section 2.0 – Who Has Public Access

a. Every member of the public should have the same access to court case records.
b. The public is defined to include:
    a. Any person, business, or non-profit entity;
    b. Any governmental agency for which there is no existing policy defining that agency's access to court case records;
    c. Any media organization; and
    d. Entities that gather and disseminate information for whatever reason.
c. The public does not include:
    a. Court employees;
    b. Entities who assist the court in providing court services;
    c. Governmental agencies whose access to court case records is defined by another statute, rule, order, or policy; and
    d. Parties to a case or their lawyers regarding access to the court record in their case (except possibly when access to information about opposing parties might pose a safety concern as with some domestic violence cases).
d. Public access is synonymous with anonymous access.

Commentary: Enhanced access outside the public role may be partly addressed by establishing requirements for identification and authenticated access. Business rules for non-public access may be quite complex and best expressed by defining roles, relationships, and the specific scope of access by case type, document type and data type. When properly implemented, the public is one of many roles whose

[12] See Appendix D for definitions of court records, case records, administrative records, and other terms.

access is enforced by a common technical infrastructure. One version of the official case record is maintained and different levels of access are enforced using virtual redaction and masking. One interesting recent issue is that there may be a significant level of attempted access by non-human requestors.

Section 3.0 – Applicability of the Policy

Section 3.1 – General Access Rule

a. Information in the court case record is accessible to the public except as prohibited by section 3.5 or 3.6.
b. In general, there should be a public indication of the existence of case information in a record to which access has been prohibited, but that indication should not disclose the nature of the protected information.
c. If harm may be done by indicating the existence of case information, then no indication of that existing record should be public.

Commentary: If a court hides the existence of case information or the case itself to prevent harm, it should make explicit the rationale it uses to determine when and why such protected information is hidden from the public.

Section 3.2 – Remote Access

All public court case records are presumptively accessible remotely.

Commentary: This section eliminates the ability to recreate "practical obscurity" by making all public court case records available at the courthouse but only a subset of those records available remotely. The principle underlying this part of the rule is that records are either public or not. The method of access should not affect that determination. In order to prevent harm, some court case records that were previously public may need to be closed. Improvements in automated redaction may mitigate that need.

Section 3.3 – Requests for Bulk Distribution of Court Case Records

a. Bulk distribution of information in the court case record is permitted for public records.
b. Requests for bulk distribution of information not publicly accessible can be made to the court for purposes with a public benefit. Courts have discretion to refuse such requests, to charge fees reimbursing the court for the cost of distribution, and to impose conditions on the requestor for access.

Commentary: If data are public, they are accessible even if in bulk form. The court has the right to make the requestor pay the cost of assembling and distributing the data in bulk form if they do not already exist in that format. The court may make

non-public data available for public purposes, but only if court users are protected from harm by imposing appropriate restrictions on access, use, and data retention. Bulk requests are often made by data aggregators and resellers. It is important that they provide to their customers only the most current versions of the court case record. A best practice is to require such users to "ping" the court database in real time to check for any changes.

Section 3.4 – Requests for Compiled Information from Court Case Records

a. The public may request access to public court case records that are not normally compiled in the requested format. The court has the right to make the requestor pay the cost of compiling and distributing the data.
b. Requests for compiled distribution of information not publicly accessible can be made to the court for purposes with a public benefit. Courts have discretion to refuse such requests, to charge fees reimbursing the court for the cost of distribution, and to impose conditions on the requestor for access.

Commentary: Requestors of compilations of non-public case information are typically barred by the court from selling the data to third parties or using the information to sell a product or service. Courts may impose additional restrictions to prevent harm. Model contracts are useful for ensuring both consistent policy use and comprehensive protection from potential harm.

Section 3.5 – Court Case Records Excluded from Public Access

a. Court case information may not be made accessible to the public if barred by federal law, state law, court rule, or relevant case law.
b. Court case records may also be excluded from public access if the court determines that harm would ensue, per the objective in section 1.0(b).

Commentary: Except for federal law, the details of what court case records are excluded from public access will vary from state to state and even from court to court in decentralized court systems. It is hard to predict how often case law might drive changes in what is public.

Common case types that are typically closed because of concerns about harm may include juvenile, family and probate. Document types typically closed include those that routinely include confidential personal information (such as financial disclosures) or potentially injurious but unsubstantiated assertions about opposing parties (such as divorce pleadings). Data types that are typical closed include identities and contact information of jurors, juveniles, witnesses, victims and other potentially vulnerable populations; financial account numbers; physical and mental health records; social security numbers; and other government identification numbers.

Consult the relevant National Institute of Standards and Technology (NIST) security standards on personally identifiable information (PII) that should be protected[13]. A best practice is to redact information in the most focused way that is technically and reliably possible. Thus, ideally, specific data elements should be masked by automated redaction. When that is not possible, then specific document types should be closed. When that is not possible, then specific case types should be closed.

Section 3.6 describes the desired business process for case-specific requests to access information otherwise barred by this section.

Section 3.6 – Requests for Exceptions to Access Policy

The courts will provide a standard process for requests to (a) prohibit access to certain public court case records, (b) allow public access to certain closed court case records, and (c) correct erroneous information in court case records. Court responses to such requests will balance the policy objectives in section 1.0.

Commentary: Considerations of harm should include (1) the risk of injury to individuals, (2) individual privacy rights and interests, (3) proprietary business information, and (4) public safety. The court should also consider applicable constitutional, statutory and common law. Where possible, explicit standard legal tests should be applied to such decisions.

It is an implementation best practice to provide the public with one centralized, easy-to-use website. The same website should support searches of public court case records, requests to expunge cases[14], and requests for bulk or compiled case records.

---

[13] *See* NIST Special Publication 800-122, "Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)," Erika McCallister, Tim Grance, and Karen Scarfone, April 2010 (http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf). *See also* "Guide to Protecting Personally Identifiable Information," Shirley M. Radack, April 28, 2010 (https://www.nist.gov/publications/guide-protecting-personally-identifiable-information).

[14] The Uniform Law Commission formed a Drafting Committee on Criminal Records Accuracy in 2014 and presented its first draft of uniform legislation in July 2016 (http://www.uniformlaws.org/Committee.aspx?title=Criminal%20Records%20Accuracy).

Section 4.0 – Timing of Public Access

a. Remote access to public court case records is essentially available at all times, subject to publicly scheduled downtimes for system maintenance and unforeseen technical issues.
b. Courts should make public court case records available in a reasonable time after filing. Courts should also respond within a reasonable time to requests for access to bulk or compiled case records and for requests governed by section 3.6, and inform the requestor when the bulk or compiled records will be available for dissemination.

Commentary: Remote access should essentially be 24/7/365. With electronic filing and reliable automated redaction, case records should become available for public access in near real-time after filing. Court responses to requests regarding public access should be "reasonable," i.e. comparable to response times by other government agencies to similar requests.

Section 5.0 – Access Fees

a. Any fees charged should be reasonable for the services provided.
b. If fees are charged, there should be a process for requesting indigency waivers, except for bulk and compiled requests.

Commentary: There is no national consensus on the charging of fees. Courts may or may not charge fees for (1) remote public access to court case records, (2) bulk access, and (3) compiled information. There are currently three fee models used by courts: no fees (there should be no monetary barriers to publicly accessible information), fees that only cover the cost of providing access, and fees that exceed the cost of provision and provide additional revenue to the court. Requests for fee waivers based upon indigency should be made available as part of the same "one-stop shop" website that is recommended in the commentary to section 3.6 above.

Section 6.0 – Operational Requirements

Access policy provisions must be supported and implemented in a cost-effective, reliable and enforceable manner.

a. Best practices should be used to protect court case records not open to the public.
b. Search capabilities for public court case records should support reasonable flexibility.
c. Search capabilities should not impose an undue operational burden on court systems.
d. Persons or organizations granted access beyond what is available to the public should be managed by role and required to identify and authenticate using best practices.

Commentary: The best policy in the world does not adequately protect confidential information contained in court case records if a court does not also implement good security practices. The National Institute of Standards and Technology (NIST) identifies cybersecurity practices and processes in a series of national standards[15]. One of many examples is encryption of confidential data in the court database.

If courts offered complete flexibility in searches for publicly accessible data, it would be tantamount to giving the public the database. That would be expensive and risky. Thus, courts must decide what search parameters to support. That should depend partly what the public most often wants to search on and partly on what searches minimize the operational burden on court systems. Finally, public access is by definition anonymous access, so there is no identification of users. This is true for information available without modification. Requestors for bulk or compiled data may be required to identify themselves and comply with other requirements. Non-public access should be controlled using appropriate best practices for well identifying and authenticating other roles that have legal but limited access to non-public case records.

---

[15] *See especially* NIST Special Publication 800-53, "Security and Privacy Controls for Federal Information Systems and Organizations," Revision 4, April 2013 (http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf). NIST is currently working on Revision 5 (http://csrc.nist.gov/publications/drafts/800-53r5/draft_sp800-53-rev5_update-message.pdf). *See also* NIST's "Framework for Improving Critical Infrastructure Cybersecurity," February 12, 2014 (https://www.nist.gov/sites/default/files/documents/cyberframework/cybersecurity-framework-021214.pdf). NIST is currently updating its Cybersecurity Framework: a draft version 1.1 was released on January 10, 2017 (https://www.nist.gov/cyberframework).

Appendix B
State of the Art for Automated Redaction

Manual redaction of paper records is a time-consuming process. Many courts still manually redact paper files. As more courts are implementing electronic document management systems, or receive files through e-filing systems, there is a growing need to have technology provide redaction solutions in the digital environment. Many vendors have created platforms that have built-in electronic redaction capabilities, or allow for redaction and other related capabilities to be added on as a component provided by a selected vendor solution. The world has been going digital for some time. This will surely increase the demand for not only redaction capabilities, but other related process improvements as well.

Redaction of electronic files starts with the software going through the process of learning patterns to determine areas that have a probability of containing information that should be redacted. Machine learning uses statistical modeling methods to predict targets, and accomplishes this by analyzing a large volume of information. The initial analysis and learning is a human/machine process. The more volume the software learns, the more accurate it becomes at targeting desired information. Other techniques, such as algorithm based natural language processing, is used to extract information from semi-structured and unstructured text. Natural language processing (NLP) is a component of artificial intelligence (AI) combined with computational linguistics, and uses methods that allow the computer to understand and process human language rather than traditional programming language. Some examples of how NLP is used are autocorrect, speech to text, and language translation.

The perspective of court CIOs is gradually shifting away from case documents to information. That information may be standalone data, metadata, or content within documents. It may even take the form of digital evidence, including videos. Redaction software needs to be capable of handling this range of targets in a sufficiently granular way, and many vendors are working towards that goal.

As courts move toward automated workflows, the supporting software needs to seamlessly support and implement those kinds of business requirements. That means redaction software must integrate with e-filing, case management, and document management software. In the near future, it must also integrate with digital recording software and digital evidence databases. Even that daunting degree of software integration may not be enough, since some courts also utilize vendors for related tasks like file analysis, data loss, records retention, data masking, and e-discovery.

Evaluation Criteria

As courts move forward with pilot tests of automated redaction, it will be very useful to collect consistent evaluation data. Some evaluation criteria are suggested here:

- Accuracy and reliability
    - Structured expressions (like case numbers) in structured documents (like forms)
    - Structured expressions in unstructured documents (like scanned pleadings)
    - Unstructured expressions (like a victim's name) in structured documents
    - Unstructured expressions in unstructured documents
- Affordability and pricing structures
- Core functionality
    - Easy specification of redaction targets
    - Easy configuration of redaction requirements (such as reliability thresholds)
    - In-line redaction
    - Ability to train on test documents or data sets
- Integration capabilities
    - Public APIs
    - Integration with third party electronic filing service providers
    - Integration with court electronic filing software
    - Integration with court case management software
    - Integration with court document and content management software

Cost and Benefits

There are various costing models for redaction and other related enhanced features, and vendors have tried to offer some flexibility that takes into account the platform, volume, and level of functionality that the court will need. A common model is the transaction-based fee model, so volume matters. Other costing options might be site based licensing that considers estimated volumes. Several vendors expect court case management companies to offer comparable capabilities as part of their off-the-shelf products in a few years.

To justify the cost of using automated redaction, courts must make an argument for the value of the capability. In the absence of costly liability lawsuits, it is difficult to make a direct argument for the value of automated redaction targeted solely at removing confidential information. If such redaction enables a court to safely open case types and document types that would otherwise have to remain closed, and if that increased openness were perceived as valuable to outside organizations, then there may be political reasons to implement automated redaction.

As the software becomes highly accurate in its identification and understanding of specific target information, it opens opportunities to use this capability in other ways. For example, the software may allow for the information extraction that can support automated workflows and thereby save the court significant time and money that may create a direct business case for adequate value. Deriving that kind of value requires

much more than just implementing the redaction software itself. A court must carefully think through its workflows, identify appropriate automation targets, and often reorganize their administrative organization (staff and skill sets) to support a new way of doing business.

As courts explore new technologies, they should consider the variety of capabilities and their related benefits now available on the market:

- In addition to Optical Character Recognition (OCR), some vendors offer Intelligent Character Recognition (ICR) that can be used for handwritten court case records. A similar capability for audio and video files is developing rapidly.
- In addition to automated data identification, some vendors offer an index that can be used to compare newly filed information to existing court case data (for example, does the party's name on the incoming case submission match the party's name in the court's case management system?).
- Some electronic content management systems offer configurable workflow engines: once the processes of OCR/ICR, automated data identification, and indexing are complete, the ECMS will apply the court's business rules and automatically route the filing to the appropriate next step in the court's workflow such as automated case entry and docketing.
- Enhanced functionality may allow for extracting data from electronic documents and automated entry in case managements systems and/or other databases.

Corrections:

The previous version of this appendix listed Mentis as being part of an Arkansas study. This reference has been removed, but this correction is to clarify that Mentis was not part of the Arkansas study.

A specific vendor listing for redaction software has been removed. The various approaches, integration methods, technology capabilities, and pricing models used by vendors cannot be suitably characterized in a simple list of redaction vendors. It is better for courts to research each solution to determine which approach, capabilities, and pricing structure best fits their platform and needs.

Appendix C
Automated Workflows Using Automated Extraction

Vendors of automated redaction software rightly point out that the ability to arbitrarily locate specified content in court case records potentially enables courts to use that information to automate court business processes in ways that can make courts markedly more efficient. Thus, redaction software might be better thought of as *data extraction software*. Such software enables courts to gather data at the time of filing for use later in a case. This is part of a larger paradigm shift from business processes built around case files and documents to processes based on data[17].

Business Values of Data Extraction Software

Once a court starts down the path of using data extraction to power its business processes, several business goals become achievable:

- Shorten the processing times for court filings and case dispositions.
- Reduce the number of court staff needed to process court filings, manage cases, implement records retention and archiving policies, and respond to records requests.
- Provide more granular public access to court case information.
- Provide appropriately redacted court case records in near real-time, reducing the lag time in publishing new case filings to the media.
- Reduce the risk of exposing confidential court case information to the public.
- Expand the scope of legacy court case records that are available for remote public access, while automating enforcement of retention and archiving policies.
- Improve the quality of court data, from the moment of filing.
- Support more sophisticated analytics of court case information.

Case Management Improvements with Extracted Information

Right now, electronic filers must input data about the filing into a so-called "envelope" so that a court can process it. Some of this "metadata" (data about data) could be extracted directly from the documents being filed, eliminating a data entry step. An important example of such metadata is the document type. Filers must currently either know or select from lists a correct document type, which is usually then checked again manually by a clerk. With high frequencies of self-represented litigants, errors in selecting document types are often made and court resources must be used to correct them. Data extraction technology can be used to reliably and automatically assign document types.

---

[17] This appendix is based on work done by Alan Carlson for the project focus group.

Obviously, the same approach can be used to assign case types and characterize case parties and their relationships. Thus, data about the case can also be extracted to drive subsequent workflows, especially those needed to perform initial triage and place a case into a case processing track. With powerful data extraction capabilities, such automated triage and case management can support business rules of arbitrary complexity, enabling courts to control cases in a much more fine-grained manner than was historically possible using manual resources. This enables courts to much better follow the dictum of allocating the right resources and attention to each case.

In a similar manner, a court can extract data from filings to help judicial officials make case decisions and issue court orders. Examples include "feeding" parents' financial data into child support calculators and populating draft court orders with extracted case data.

Data extraction software can do all these tasks more consistently and reliably than humans can, once it is possible at all. Data extraction technology could ultimately eliminate the need for both e-filing envelopes and case cover sheets.

Appendix D
Definitions

Administrative Record – Court records that pertain to management, supervision, or administration of the court and are not part of a case record.

Automated Case Triage – A method of differentiating cases by assigning them to a track early based on issues and corresponding processing requirements, rather than case type. This method also provides litigants with alternate choices from traditional litigation that might offer a more rapid resolution at lower costs.

Automated Workflows – A well-defined set of business processes where information is exchanged and automated actions take place based on a set of procedural rules.

Bulk Distribution - The distribution of all, or a significant subset, of the information in court case records without modification or compilation.

Case Record - Any document, action or information that is collected, received, or maintained by a court or clerk of court connected to a judicial proceeding. It may include an index, calendar, docket, register of actions, official record of the proceedings, order, decree, judgment, minute order. These may have been collected in a case management system that is used to track information. Case records may contain both public and confidential information.

Court Records – The sum of all administrative and case records in the judicial branch.

Compiled Information - Information that is derived from the selection, aggregation or reformulation of some specified subset of data from more than one individual case record.

Data Extraction – An automated means of taking data out of structured forms or using machine learning and other mechanisms to take data out of unstructured text for use.

Machine Learning – A type of artificial intelligence (AI) that uses patterns and predictive analysis to draw inferences and act without the need for precise programming. Inferences become more precise with greater use.

Metadata – Data that provide additional information about another data source to put the information into context, such as title, author, subject, creation date.

Practical Obscurity – A concept based in a paper record environment where an individual's information in government files enjoys some level of privacy because access is limited to an on-site review of a paper file.

Predictive Analytics – An advanced analytics technique using statistical analysis that utilizes new and historical data to forecast the probability of future activity, behavior and trends.

Redaction – The process of obscuring confidential information contained within a public record from view. Redacted portions of the record are blacked out or masked. Redaction may be accomplished manually or through use of technology such as data identification software.

Remote Account Access – Electronic access to records based on role that is defined by rule or statute, and authentication of that role. This access may include greater view of the redacted or un-redacted information in a case file that one may be a party to or that is required as part of an agency service or function.

Remote Public Access – The ability to electronically search, inspect, or copy information in a court case record without the need to physically visit the court facility where the case record is maintained. This generally does not require any type of login or the need to provide identifying information about the member of the public accessing the case record.

Structured Data – Information contained in a database or structure where the information may be readily identified and used. In the context of data extraction software, structured data are identified based upon their unique patterns. Examples include United States Postal Service zip codes, Social Security numbers, and phone numbers.

Unstructured Data – Information not contained in a data structure or database, such as text in documents or multimedia files such as digital recordings of audio or video without XML markup.