The Importance of Technology in Classification and Declassification
A White Paper of the Public Interest Declassification Board
June 2016

Introduction to the PIDB Declassification Technology Working Group

At the direction of the President, the Public Interest Declassification Board (PIDB) continues to investigate technologies and related policy solutions to transform the security classification system to one capable of functioning more effectively in an increasingly complex information age[i]. Core to our democratic ideals is the ability for the public to access its government's records. The responsibility lies with senior government leaders to develop sound policies and implement technological capabilities that will ensure long-term preservation and accessibility to the nation's historical records. Nearly all users of the security classification system agree that it is no longer able to handle the current volume and forms of information, especially given the exponential growth of digital information that is only exacerbating the many challenges facing the system. **As the PIDB has previously noted in all of our reports, we reaffirm that our most important recommendation for developing and ensuring such a system is the adoption of a government-wide technology investment strategy for the management of classified information.**

In support of this recommendation and those commitments found in the President's *Open Government National Action Plans*, the PIDB began an in-depth study of agency declassification technology initiatives last year. In May 2015, we established an informal Declassification Technology Working Group (Working Group) at the National Archives and asked for agency participation in a high-level questionnaire concerning agency preparedness for declassification in the digital age. We sought support from agency Chief Information Officers (CIOs) when setting up the Working Group in order to highlight declassification technology development as a need for agencies. We believe the support of agency CIOs is critical to modernizing declassification and making the management of classified information at agencies a priority in planning their information technology programs now and in the years ahead.

The Working Group has representation from technologists at 14 agencies and departments in the Executive Branch. The PIDB hosted four Working Group meetings in the past year. These meetings are an opportunity for agencies to share their successes, challenges, best practices, requirements and declassification program needs. Agenda items covered at these meetings included agencies briefings on their efforts at declassification technology planning, discussions of best practices concerning the management of classified records (including email), the sharing of metadata standards and transfer guidance, and more. We have received positive feedback from agencies about the usefulness of meeting in this informal Working Group environment; agency technologists are able to work collaboratively, share best practices and discuss new ideas with their inter-agency counterparts on these often overlooked technology challenges.

Now, at the one-year anniversary of the beginning of our Working Group exercise, we have collected some observations and lessons-learned to share from these meetings with the public. Our

PIDB

goal is to reflect on the progress of the Working Group and plan next steps and potential areas in need of further study.

<u>Finding the Baseline: Where Agencies Stand</u>

Overall, agencies lack appropriate technological investment to support the activities of their declassification and related records management programs. Most agencies do not possess basic workflow applications to assist human review of records, applications that are readily available in the commercial world. While one or two agencies are exploring advanced content understanding and analytics as technical capabilities to assist review, the vast majority of agencies lack the most basic technological infrastructure to support simple automation or search technologies to assist in the management of records through the review process.

By policy design, declassification largely operates in an information environment twenty-five years in the past, making paper the dominant review format agencies must prioritize. Solutions that can assist in managing the large volumes of paper found at agencies and the National Archives already exist in the commercial world. But implementing these known solutions within government remains elusive and problematic. Funding for declassification and records management in most agencies is minimal, at best. What little funding is available supports outdated processes designed in the 1990s in response to the mandates afforded with the onset of automatic declassification. Prior to the notion of automatic declassification, declassification review occurred ad hoc and inconsistently across agencies. When adopted and implemented, these 1990s processes elevated declassification review to the program level. They have served their intended purpose - to institutionalize declassification at agencies – and presently are largely outmoded for managing electronic records. These 1990s processes will remain in place for the foreseeable future, barring resources for the development of new processes and the adoption of automated workflow tools.

In addition to the challenges of outdated paper-based processes, agencies also lack capabilities to manage the review of special media formats and legacy electronic records, including first generation born-digital records. As prioritization of records for declassification review largely depends upon records' age, the coming of "age" of electronic records review is now of serious consequence for agencies, with the added complication that no relief from paper records review appears to be in sight. Common challenges exist among agencies in managing legacy electronic records, yet there is no serious effort underway to acknowledge or describe these challenges, let alone develop a universal approach or solution.

Other common problems exist concerning electronic records beyond the issue of exponential growth and volume in need of review. Connectivity, integration and communication of systems that support declassification and records management within and between agencies is fragmented and sparse. Agencies lack universal metadata requirements and standards for managing declassification. Requirements and standards are of the utmost importance as declassification is increasingly dependent on the ability of agencies to refer their records to other agencies for equity review. Agencies must adopt and implement common solutions to these challenges across government; progress of any one agency in building a technological framework to modernize its declassification program is dependent on its ability to interact and share information with its counterparts.

Sharing information among agencies also exposes cultural challenges found in the declassification world. A common understanding and agreement for how agencies should mitigate risk does not exist. Agency practices are intolerant of risk and the consequences of not striking a balance between openness and continued secrecy in declassification review are too high for the

system to sustain indefinitely, both in resources and credibility. Today's information world, including the national security structure, is increasingly dependent on transparency and open source informational content. Risk management and mitigation must be key elements of forethought in designing technical declassification capabilities, not an afterthought in response to disclosure events.

Next Steps: What Agencies Need

Technological modernization of declassification and its related functional counterpart, records management, will require leadership and resources. Agencies require both simple workflow tools and advanced content processing, analytic tools and storage/access means. Agencies should integrate declassification reviewers and records managers, organizing for success, to share best practices, manage metadata and efficiently harvest all the capabilities of information age technologies for the benefit of all system users, including policymakers and historians. Additionally, special media and first generation born-digital records demand serious consideration. A government-wide investment strategy should consider and build upon those tools in use at agencies with more modernized declassification capabilities, such as the intelligence community.

A phased adoption of sophisticated content analytic solutions should occur, beginning with an increase in the number of pilots used to test these capabilities within declassification programs. Capabilities, like those developed at the Center for Content Understanding at the Applied Research Laboratory at the University of Texas at Austin, should be implemented to a greater extent at agencies[ii]. For most agencies, there is an immediate need to implement automated workflow solutions and basic search capabilities, solutions that largely exist in the commercial world that are readily available for adoption. Even while grappling with basic workflow challenges, agencies must also seriously invest in advanced content analytic tools. The sustainability of the system is dependent on agencies exploring advanced content analytic solutions while also solving immediate workflow automation challenges.

Even more importantly, the long-term transformation of the declassification system will require leadership from the White House and a commitment to funding a government-wide technology investment strategy. The PIDB will continue studying declassification technology investment at agencies with the recommendation that agencies receive the resources they need to make the records of our government accessible to future generations. Our desire is to support policymakers, while maintaining our principle responsibility of responding to the public interest in having an open and transparent government. We believe the government will only be able to achieve this goal with the adoption of technological capabilities that will modernize the security classification system to function effectively in the current digital information environment.

---

[i] Memorandum for Implementation of the Executive Order 13526, "Classified National Security Information," December 29, 2009, 75 FR 733, Document Number E9-31424.

[ii] At the request of the CIA and the National Archives, the Center for Content Understanding at the Applied Research Laboratories at the University of Texas at Austin piloted decision- support technology for records declassification review and release. The pilots successfully yielded a Sensitive Content Identification and Marking (SCIM) tool that uses a combination of natural language processing, expert systems, machine learning and semantic knowledge representation to identify sensitive content in textual information found in classified email records. The SCIM tool is the only tool of this level of sophistication being explored for the sole purpose of aiding decision-support in classification and declassification.