Open Contracting in the U.S. Public Sector

Converting Procurements from Documents to Data

Omid Ghaffari-Tabrizi[*,†]

* Originally submitted to the faculty of the George Washington University Law School in partial satisfaction of the requirements for the degree of Master of Laws in Government Procurement Law

† Correspondence: *see* https://www.ghaffaritabrizi.com/make-dad-proud/

Turning the public sector procurement process from a manual paper-based approach to an automated AI-powered system can happen – and is happening – today. Converting the approach and artifacts governments use for buying anything from products to services into an open-standard based machine-readable and human-readable format will allow a procurement system to remain modern in nature and approach, and in the IT sector at least, also influence the architecture of the very infrastructure the procurement system runs on. Where wide-scale changes are required for a procurement system, such as those efforts underway in the United States and Europe for example, a data-based procurement system will also support change management efforts as lessons can be individualized in content and teaching tactics. Implementing this move towards machine-readable data and documents via the procurement process in the United States specifically will provide immediate and significant results and use cases from around the world demonstrate exactly how.

Introduction

Without use of an open data standard to convert procurements from digital paper to actual data, the procurement system can only be digitized, not modernized. Even if documents are digital, when a process remains paper-based, the advances of technologies like artificial intelligence ("AI") and the collaboration between stakeholders available via cloud-based solutions will fail to achieve optimal outcomes. When considering this in the context of Federal Government procurement, this could mean tens of billions of dollars of efficiencies lost considering trillions are spent every year (BFS, 2024).

Machine-readable data and documents, the next step after paper, have been used for decades for a number of critically important and complex systems. With those decades of experience, in some sectors at least, the importance of interoperability has been built into a solution, allowing the underlying technology to advance with time. In the context of machine-readability and procurements, this means moving from simply being machine- readable to machine-executable, machine- interpretable, and ultimately machine- controllable (Eichstädt et al., 2022, 1).

The effort to make this work in the United States via the open government data movement has been slow but steady. Critically still, it has been bi-partisan and consistently advanced across each administration since the earliest days of the digital age. However, with the progress of the past and the bold reforms underway today, there is an opportunity to cross the threshold and achieve a procurement system that will objectively achieve the desiderata of a public sector procurement system and do so in a measurable way that can allow for improvements over time.

In exploring how the United States is positioned to become a world leader in AI-powered procurement, this paper will cover the following: relevant historical background that outlines how paper-based procurements are a natural fit for automation via machine-readability (and beyond), the current state of open contracting in the public sector based on use cases from around the world, and the current state of open contracting in the United States. Ultimately, this paper will show that the United States is ready – and should – implement an open contracting data standard, whether it is based on one that exists today or not, in order to make the most of technology currently available while achieving the defining goals of a procurement system, whatever they may be for this or future administrations.

Historical Background

The promise of AI is often based on what it can automate, "chang[ing] how businesses and the Federal Government provide goods and services" by "perform[ing] tasks that have traditionally required human intelligence." Essentially, the algorithms that power AI are able to take on more complex problems as they become more complex themselves, generating creative solutions and insight from incomprehensible volumes of data (CBO, 2024, 1, 3, 9-10). Whether being used by the Premier League team Liverpool F.C. to develop tactics (Wang et al., 2024, 1) that make it the most successful club in England (Stats Perform, 2025) and put it among the best teams in Europe (if not the world) (UEFA, 2025), or by meteorologists making weather predictions (Lam et al., 2023, 1421), it is structured data that proves to be the most

effective at providing the type of results that achieve the promise of AI and the automation that comes with it (Radhakrishnan et al., 2024, 4).

How Documents led to Automation

It is no accident that today's most complex AI systems are built on automation that relies on standardized approaches – "[t]he first complex machines produced by man were automata" as developed by the Greeks. In fact, much of modern-day automation is based on the work of Heron of Alexandria, who lived during the mid-first century AD and had his manuscripts translated into Latin in 1501 (Bedini, 1964, 24-25). Much of the automata for the next 80 years was heavily influenced by his designs (Papadopoulos, 2016, 217). Heron himself owed a "debt to Philo of Byzantium," who lived around 280 to 220 BC and from whom Heron seemingly borrowed liberally, if not verbatim in several notable instances, including in his works titled Pneumatica and Automata. As noted in those writings and from automata's earliest representations as mythical devices to Philo and Heron's time to the Renaissance period when their seminal works were translated, automated devices were "[m]ainly used for entertainment" (Grillo, 2019, xxviii, XCV).

The connection between entertainment and automation is rather critical in the history of machine readability – much of the inspiration for Basile Bouchon's semi-automated loom, introduced in 1725 and credited with creating "the concept of a *stored program*" (emphasis in original) (Park & Jayaraman, 2017, 605), came from the work of his father, who "was an accomplished organ builder" and responsible for Bouchon becoming "familiar with automated carillons." With automated carillons, all that was required to play a desired tune were "paper maps" that relied on "the insertion of pins in an appropriation pattern." By repurposing the carillion's pin and barrel with a perforated paper cylinder, two people (hence why it was semi-automated) were able to "control the patterns woven by looms."

Improving upon this concept, in 1728, Jean Falcon, an associate of Bouchon's, replaced the perforated paper cylinder with "a chain of sturdy punched cards" that allowed for "editing of the programmed patterns [...] without having to replace the whole set." In 1745, Jacques Vaucanson, a countryman of Bouchon's, introduced an "automatic advance mechanism" that allowed for the paper to be fed into the loom, meaning it was no longer necessary to have a person involved in the process at all. Ultimately, though, it was the work of Joseph Marie Jacquard in 1801 who combined the concepts to create the Jacquard head, commonly referred to as the Jacquard loom, creating a scalable automatic loom that revolutionized the textile industry (Lee, 2009, 70) – and provided inspiration to Charles Babbage and his pursuit of what he

called the Analytical Engine, aimed at performing ever more complex mathematical computations (Park & Jayaraman, 2017, 607).

Early Days of Using Paper to Store Data

With most of the concepts released in 1838, an analysis of the Analytical Engine and its predecessor, Babbage's own Difference Engine, shows that they were completely original but for "the adoption of the barrel, from automata and music boxes, and punched cards, from the Jacquard loom." In adopting the punch cards, Babbage was able to create "a program [...] of unlimited extent" by allowing the machine "to read and punch numbers on other cards." Essentially, Babbage created two types of punch cards: "operation cards" that provided instructions to the system and "variable cards" that contained the data the system was intended to digest and utilize (Bromley, 1982, 197, 215). The significance of storing data on paper in this fashion was summed up by Lady Ada Lovelace, who stated that "the Analytical Engine weaves algebraic patterns just as the Jacquard-loom weaves flowers and leaves" (Hofstadter, 1999, 33).

While Babbage never saw his Analytical Engine come to fruition, the idea of using punch cards to automate the processing and tabulation of vast amounts of data via punch cards was implemented for the 1890 Census due to the work of Herman Hollerith and his "electric counting machine" (Census Bureau, 2025). Having been influenced by Dr. John Shaw Billings (Census Bureau, 2025) who originally wanted to create a device "which would operate on the principle of the Jacquard loom," Dr. Billings credits Hollerith with applying "the power of electricity" to "a machine for doing the purely mechanical work of tabulating population and similar statistics" (Truesdell, 1965, 30). With the hope of obtaining that government contract associated with supporting the 1890 Census, Hollerith created a machine that recorded data on punch cards (Hollerith, 1889).

The system in place at the time was manual and time-consuming, resulting in Census tabulation taking up to a decade to complete (Census Bureau, 2021). As part of a competition in 1888, Hollerith's machine proved to be nearly ten times faster in tabulation and as much as twice as fast in capturing data as its competitors (Census Bureau, 2024). The system that went live in the 1890 Census was revolutionary in that it used an individual punch card per family or household, formatted such that the required details on each individual and the family or household as a whole would be collected on a punch card with a standard format. This approach allowed for automated tabulation, taking the time to complete the Census down from over a decade in 1880 to a little more than six weeks in 1890 (Truesdell, 1965, 57, 62).

Machine Readability Generally

The advent of a "machine-readable punch card" as a means to store and manipulate data showed its value when the Social Security Board, formed as part of the Social Security Act of 1935, chose the "electric accounting method" being proposed by International Business Machines Corporation (the company that evolved out of the merger of four firms that included Hollerith's firm that won the 1890 Census contract) to manage social security numbers over an otherwise manual approach being proposed by two others (Cronin, 1985). By using two different cards, one categorized by social security number and another by surname, staff could search for one particular individual, using either method, among the 35 million records they started with when launching this effort in 1937 (Puckett, 2009). This allowed the punch card to change from a discardable "processing tool" to the most critical component of a "large register" being used for "bookkeeping" (Heide, 2009).

The benefits were immeasurable to the private sector as well – "the insurance industry depended on the ability to run statistical analyses on large populations; telephone and utilities companies used tabulators to bill large numbers of customers for small amounts; and, manufacturing firms developed real-time accounting methods for tracking fluctuating costs and sales across large multidivisional corporate structures" (Driscoll, 2012, 10). However these punch cards lacked storage capacity and the ability to quickly access the information "stored" in them, which led to the transition to magnetic tape and disk drives in the 1950s (Heide, 2009).

One of the great benefits of that transition away from machine-readable punch cards to these non-paper based machine-readable formats is that it made it cheaper and easier to begin moving "textual material in machine-readable form" (Schneider, 1971, 39). Even in the 1970s, the amount of the information that could be converted into a machine-readable medium made a significant impact on "library and information services," including those databases focused on providing research material (Williams, 1977, 101). Any system intended to make use of data structures and algorithms will be dependent to a significant degree on the manner of implementation, as "an appealing model or formatting language may need to be rejected if it is impossible to build efficiently" (Furuta et al., 1982, 466). This quest for scaling how much information we can store and ingest in this format has led us to standardize everything from how we display tables in journals (AAS, 2000) to how we keep libraries produced for genomics assays (Booeshaghi et al., 2024), ideally aiming for that data to be searchable and easy to find (Rowe, 1982, 351), in the appropriate format for the device being used to view it (Hurst et al., 2009, 99), while also preparing it for technology that doesn't yet exist (Williams, 2024).

Machine-Readable Data versus Documents

Machine-readable information consists of "records created for processing by a computer" (Dollar, 1978, 423). When viewed through the lens of government records in particular, these records can be generally categorized as two types: "'documents' (which are typically static and frozen in their format) and 'data' (which may be dynamic and can be open to further processing)" (Hendler & Pardo, 2012). With the National Archives storing documents and data on magnetic tapes and appraising those records from as early as 1969, the line between documents and data has blurred. The content of the documents is essentially data represented through a "variable card" and the format, schema, or framework used to identify what that content actually represents and how it can be utilized as a data source is essentially an "operations card" (Dollar, 1978, 423-24).

Using the internet as an example of how content can be converted into data in this fashion, Schema.org represents a standard set of schemas that search engines and other applications can use to make the information represented on a page machine-readable and thus manipulable (Schema.org, 2025). Using open-standards such as RDFa (W3C, 2015), Microdata (WHATWG, 2025), and JSON-LD (W3C, 2020), "45 million web domains markup their web pages with over 450 billion Schema.org objects," showing the extensibility of this concept to develop and adopt with ease community led standards (Schema.org, 2025). As an issue recognized for decades by researchers in this space, failure to implement a standard, especially as it relates to "library and information services" will make things harder and eventually lead to failure (Williams, 1977, 98). Related specifically to machine-readable documents, the importance of a standard that treats such documents differently from the machine-readable data universe they are intended to join was recognized from the very beginning when Richard C. Roistacher and Barbara Noble introduced their Style Manual.

In the Style Manual, which was itself a machine-readable document, it recognized that what is considered "machine-readable data" in the format of a document was (and arguably still is) a sector ripe for expansion and refinement (Roistacher, 1977, 32-34). In a circular fashion, as machine-readable documents convert content into data, machine-readable data can in turn be cited as a source for a document (Census Bureau, 2025), and where it has achieved a high degree of machine readability (i.e., it has followed appropriate structuring and formatting), that data could also automatically end up being available for use in follow-on research and documentation (NIH, 2023).

An example of this in practice is the interactive map provided by the Digital Art History Journal ("DAHJ"), which "represents all DAHJ articles and their associated keywords as a network" (Int'l Journal for Digital Art History, 2023), allowing a different kind of approach to publishing that provides for organization of content in "thematic tracks" rather than issues and in a way that can be "expanded upon and modified over time" (Klinke, 2023), An additional example worth noting is the "Arlington PDF Model" (PDF Association, 2023), a result of research funded by the Defense Advanced Research Projects Agency ("DARPA") to convert the ISO-32000 standard into a machine-readable definition "for the entire PDF document object model," a program that not only allowed for the creation of "the first open access, vendor neutral, comprehensive, specific-derived machine-readable definition of all formally defined PDF objects and their intra- and inter-object relationships" but it also helped those machines, and in turn the humans using those machines, make sense of the more than 1,000 pages that make up the standard itself (Wyatt, 2021, 13). Accommodating that sharing of data and knowledge via standards that support "interoperability and flexibility" is recognized as one of the eight "Data Citation Principles" (Martone, 2014) and shown to be necessary for advancing the use of machine-readability in these and similar examples.

Machine-Readable Formats

Building on those data citation principles and the variety of other efforts to standardize the use of content in documents as data, the International Electrotechnical Commission ("IEC") and the International Organization for Standardization ("ISO") are working on an effort to turn standards themselves into a "standards machine-applicable, readable, and transferable" format called "Smart Standards" (IEC, 2023). Beyond the benefits this work arguably can provide, which is worthy of a paper of its own, the effort has resulted in a "classification scheme for Smart Standards" that developed around the concept of "utility to machines" and outlines five levels that increase in level "the more useful it becomes." These levels are:

- Level 0 – Paper format.
- Level 1 – Open digital format, like PDF, that allows you to read and search on screen.
- Level 2 – Machine-readable document, like XML, that has structured content of standard documents and content that can be processed by software.

- Level 3 – Machine-readable and executable content, the first set of Smart Standards, it has semantic enrichment of content for selective access and can receive content of multiple standards for a given purpose.
- Level 4 – Machine interpretable content, the second set of Smart Standards, it has information modeling capabilities that express content and the relationship between elements, has interruption-free data flow within the value chain, and has automatic answering or predictive content supply capabilities (IEC, 2023, 12-13).

The reason why the latter levels are considered Smart Standards is because at Level 3, a machine can "identify requirements and recommendations" while at Level 4 a machine "can understand the meaning and relate to the content" (IEC-ISO, 2022), The higher the level a document can attain, the better the opportunity will be for the content within the document to be preserved and also made accessible and available to researchers (NARA, 2022, 13). When it comes to documents, government data, and other information that are intended to be used for research, a combination of machine-readable structured data presented in a way "that can [also] be conveniently read by a human" is, unsurprisingly, referred to as a human readable format (Open Knowledge Foundation, 2015), and serves to enhance the accessibility and availability provided by machine-readable formats.

Machine and Human Readable Formats

Examples of machine-readable formats that are also human readable and worth noting include Level 2 formats like Extensible Markup Language ("XML") (W3C, 2008), JavaScript Object Notation ("JSON") (JSON.org, 2017), and Internet JSON ("I-JSON") (IETF, 2015). Also worth noting is Markdown (Chen et al., 2025, 1), an open file format intended to be an "easy-to-read, easy-to-write plain text format [that can] then be convert[ed] to structurally valid [Extensible HyperText Markup Language ("XHTML")]" (Gruber, 2004).

XML was started in 1996 (W3C, 2003) and while "[o]riginally designed to meet the challenges of large-scale electronic publishing," it has proven to be flexible enough to support a number of offshoots such as Really Simple Syndication ("RSS") (RSS Advisory Board, 2009) and OpenDocuments standardized as ISO 26300 (ISO-IEC, 2015). Having branched out of XML, XHTML was considered "the next step in the evolution of the Internet" when the revisions were released in 2002 (W3C, 2018). The need for XHTML arose out of "[d]ocument developers and user agent designers [...] discovering new ways to express their ideas through new markup" (W3C, 2018). This

resulted in "interoperability problems for documents across different platforms" stemming from the original intention of Hypertext Markup Language ("HTML") to serve as "a language for the exchange of scientific and other technical documents" rather than the rich data-rich sources of documents and information today. Considering the move towards AI-powered solutions, it is worth noting that the leading Large Language Models ("LLMs") ingest prompts and deliver responses using Markdown, which is converted to XHTML when displayed to users.

Public sector versions of XML that have branched out and bear mentioning include the eXtensible Business Reporting Language ("XBRL") (XBRL Int'l, 2013), used by the Federal Financial Institutions Examination Council ("FFIEC") (2025), and Inline XBRL ("iXBRL") (XBRL Int'l, 2022), used by the U.S. Securities and Exchange Commission ("SEC"), respectively (SEC, 2025), both of which are also intended to ultimately be displayed and manipulated by users. These formats allow for thousands of pages of content to be shared and analyzed for a variety of use cases, from compliance to business intelligence. Internationally, the Universal Business Language ("UBL") is used in a relatively similar fashion for a number of commerce and legal related data sharing issues (OASIS Open, 2024). It originally started as part of the XML standards community and has now evolved into a standard of its own that includes syntaxes such as JSON (Bausá Peris, 2014).

The first JSON message was sent in 2001 (Crockford, 2009) and originally standardized in 2013 and now widely available as ECMA-404 (ECMA Int'l, 2017) and ISO/IEC 21778:2017 (ISO-IEC, 2017). JSON has been adopted to such an extent that it is one of, if not "the main format for exchanging information over the web," primarily because it "is a language which can be easily understood by both developers and machines" for a variety of reasons, including making calls against other systems via an application programming interface ("API") (Pezoa et al., 2016, 263). JSON and I-JSON, the latter which "is a restricted profile of JSON designed to maximize interoperability and increase confidence" that the output will be successful and predictable (IETF, 2015), differs from XML in several important ways: order of elements is not important in JSON but it is in XML; there are array types in JSON that are missing from XML; subelements cannot have the same name in JSON but they can in XML; and JSON is integrated into scripting languages but XML has a separate query language (Goyal & Dyreson, 2019, 135).

These differences are particularly important for the open government movement, especially as it relates to the unique subelements available via JSON – the use of unique identifiers, or Uniform Resource Identifiers (URIs), are critical to all of the "four basic principles expected to be fulfilled when aiming to publish linked data" (Hoxha &

Brahaj, 2011, 110). It is also important because from the United States in 2009 to Malaysia in 2014, all six of the nations deemed to have implemented open government data in a significant enough fashion were found to have adopted the Comprehensive Knowledge Archive Network ("CKAN") platform (Bachtiar et al., 2020, 330). CKAN describes itself as "an open-source DMS (data management system) [that] makes it easy to publish, share and use data" (Open Knowledge Foundation, 2025). The importance of this is the use of JSON by CKAN's APIs (CKAN, 2025) is demonstrated by the fact that the two institutions that implemented open government data in Indonesia included JSON among the data formats they used, with one using JSON exclusively (Bachtiar et al., 2020, 330, tbl. VI).

Open Contracting in the Public Sector

In the United States, the use of open government data formats for critical financial data by FFIEC and SEC is not an exclusive pursuit – for example, in 2018 Government Accountability Office ("GAO") pushed for more transparency into the trillions of dollars in federal spending (GAO, 2018, 1). While efforts in the United States to modernize procurement data reported in the Federal Procurement Data System ("FPDS") (GSA, 2025) or via USAspending (GSA, 2025) have brought mixed results (Garland, 2023, 4), global efforts to achieve transparency and efficiency have seemingly coalesced around I-JSON based Open Contracting Data Standard ("OCDS") (Open Contracting Partnership, 2025). Arguably, OCDS presents an easily adoptable and currently available approach to address the three main challenges identified as most critical to the procurement process: technical, legal, and political (Adam et al., 2025, 3).

The Recent History of Open Contracting

With a recognition that striking a balance between transparency and efficiency is critically important to a well-functioning, responsive, and fair public procurement process that survives those and other challenges (Beth, 2007, 29), open contracting is an important and growing part of the open government and open data movements (Marchessault, 2013, 79) and touches on all three of those challenges. OCDS, originally developed by the World Wide Web Foundation with support from the Omidyar Network and World Bank, is the most widely adopted form of open contracting (Open Contracting Partnership, 2021). It is important to note that e-procurement is distinct and different from open contracting – the former represents a procurement system that utilizes "information outside their data environment" while the latter involves "machine-to-machine communication," among other advances, using AI-powered

technologies (Mavidis & Folinas, 2022, 2). While the concepts behind open contracting have built on ideas and processes that came before it, like e-procurement, 2015 represents the launch of OCDS (Open Contracting Partnership, 2022) via Open Contracting Partnership ("OCP") (Open Contracting Partnership, 2024) and the early 2020s represent the moment when open contracting arguably could claim enough adherents (Edwards et al., 2020, 25) who had adopted the approach for long enough to generate actionable and high-quality data (Adam et al., 2025, 10).

Desiderata and Open Contracting Principles

With OCDS and OCP driving much of what is considered the most successful implementations of open contracting globally, analyzing the Open Contracting Global Principles (Open Contracting Partnership, 2020) against what is considered to be desiderata of a public procurement system (Schooner, 2011, 103) provides a useful ability to determine how open contracting would fit within a modernized public procurement system in the United States (Exec. Order No. 14275, 2025). OCP breaks their Global Principles up into two categories: Affirmative Disclosure, which has six principles, and Participation, Monitoring, and Oversight, which has five principles (Open Contracting Partnership, 2020), while the identified desiderata number nine in total. Using the desiderata as a base, the two sets of principles align closely.

Alignment with the Three Overarching Principles

Beginning with the "three overarching principles" identified in desiderata, "competition, transparency, and integrity," the Global Principles and desiderata are particularly aligned. Competition is focused on "the power of the marketplace," aiming to give the government the best possible value without compromising industry's drive for profit. Integrity is critical because consistent and clear rules help battle corruption via "[b]ribery, favouritism, or unethical behaviour." Finally, transparency is essential to ensure the participants and the public writ large trust the outcomes are the result of an "impartial and open" process (Schooner, 2011, 2-5). Arguably, when analyzing the OCP's Global Principles, these three desiderata-based principles are baked into all eleven principles.

Beginning with the principles grouped under Affirmative Disclosure, OCP starts with a recognition of "the right of the public to access information" related to the procurement process. Even more to the point, the second principle recognizes the need for public sector procurement to "be conducted in a transparent and equitable manner." The third principle's detailed focus on what is published aims to create a system that includes enough information to satisfy any line of inquiry. While the fourth

principle is focused on adopting OCDS, it is ultimately based on the idea of making information accessible via a consistent and comprehensive format. The fifth principle on making sure information is as complete as possible, especially where recourse is required or disputes are present, is particularly important for a trustworthy procurement system. Sixth, and finally, ensuring information sharing is feasible without overly-restrictive confidentiality clauses will be critical to the openness of open contracting.

Competing Aspirations and Tensions

While OCP recognizes the importance of making information available, there is a natural tension between openness and legitimate needs for confidentiality in certain circumstances, including those established by legislation or regulation. To that end, OCP's five principles categorized under Participation, Monitoring, and Oversight provide the type of principles that guide how governments and parties involved within a procurement process can make the most of OCDS with that tension in mind. First is the right of public participation in the oversight of the entire procurement process. Second is the need to support an environment to protect that right of public monitoring of the procurement process. Third is the need for public private partnerships throughout. Fourth is the need for internal oversight of government functions themselves. Fifth, and finally, is the requirement for individual contracts to include an opportunity for public participation (Open Contracting Partnership, 2025).

Within the desiderata, six individual principles are identified as those that create a natural tension with the three overarching ones. First is efficiency, which is focused on "administrative or transactional efficiency" meaning the act of buying happens with the least amount of resources consumed. With that efficiency, ideally there is no compromise of participants' satisfaction in the process and the government would still end up obtaining the best value, the second and third competing aspirations of an ideal procurement process. Additionally, there would ideally be an ability to distribute wealth in accordance with a government's policy, such as those of recent years focused on small businesses, without compromising trust. This is critical here as well as in the process generally as risk mitigation can manifest in a variety of ways, some of which may inadvertently stifle contracting goals. Finally, a "valuable systemic goal [that] is frequently lost in efforts to achieve some or all of the previously discussed policies" is uniformity, a concept that suggests each buying function within the government buys in the same way, using the same rules, and via the same means (Schooner, 2011, 9-14). Arguably, the use of OCDS in accordance with the principles identified by OCP would align and address the desiderata principles as well.

Global State of Open Contracting

With the Global Principles themselves recognizing they "are to be adapted to sector-specific and local contexts" and with the understanding that they "are complementary to sector-based transparency initiatives" and other open government movements (Open Contracting Partnership, 2025), it is worth identifying particular implementations of OCDS from around the world and comparing them against the desiderata themselves.

Competition

The adoption of OCDS in Ukraine helped combat corruption that amounted to "by [a] very low estimate, 20% of spending in public procurement" while also saving $55 million through improved competition. This 2016 World Procurement Award winning program saw all related information, "from planning through to payments," input into a centralized and searchable database with industry participants providing their solutions via commercial marketplaces. With platforms competing, the procurement system itself incentivized the improvement of products and services available via each marketplace (Open Contracting Partnership, 2016).

Integrity

Use of OCDS in Bogotá, Colombia identified $22 million earmarked for providing high-quality meals to school children that was lost to corruption. With "more than 800,000 hungry students between the age of four and 18" reliant on this program every day, the need to trust the system is important, not just for the children who rely on these meals and the parents who are the taxpayers expecting the appropriate use of their funds, but also by the vendors themselves. The removal of corruption saw the number of vendors participating in the school lunch program jump from 12 to 55, many of whom had never participated in the public sector procurement process prior to this effort (Open Contracting Partnership, 2018).

Transparency

Slovakia utilized the OCDS to implement a "regime of unprecedented openness" by publishing most of their contracts online. Nearly 8% of the Slovakian public checks a contract or receipt every year and media attention has increased by 25% since the start of this effort, leading to a documented decrease in the "perception of corruption." Additionally, the increased transparency led to a drop in single bidder tenders,

dropping from more than 50% to "an all-time low of 34%" in just four years (Sípos et al., 2015).

Efficiency

Tracking physical and financial progress of hundreds of major infrastructure projects in Nepal is relatively harder than other countries due to the decentralized nature of the government. In using OCDS to feed a public portal, the public saw major efficiencies, including moving from holding hearings every four months to holding hearings every day, project amendments decreasing from 13% to 2%, and staff managing the projects seeing an efficiency gain of 75% and saving 3.5 hours on average, among other benefits. The improved processes not only led to greater public satisfaction from the transparency now available but also from the savings seen (Brown, 2020).

Customer Satisfaction

With more than $3.5 billion being spent on public works in Buenos Aires, Argentina, public requests for information were numerous and costly in terms of required resources. Prior to adopting OCDS, the municipal government found itself completing those requests in weeks. After implementing an interactive open contracting data platform, information became available within seconds. Critically, this information being available so openly and so quickly also helped the government programs themselves as they were able to take advantage of the same access (Open Contracting Partnership, 2021).

Best Value

Use of OCDS in Chile allowed the nation's largest pharmaceutical purchaser to reduce costs by $9 million dollars across 60% of the medicines purchased. This was particularly impactful on Chilean society as, prior to this, a significant portion of the population would travel abroad to obtain healthcare elsewhere. These improvements have also been tapped to scale to an upcoming investment of $8 billion dollars for infrastructure projects (Open Contracting Partnership, 2021).

Wealth Distribution

With more than 10% of the Philippine population consisting of indigenous groups and with a disproportionate number of them living in locations that are the most mineral-rich, the use of open contracting data was able to ensure the equitable distribution of royalty payments via withholding taxes. Beyond the financial benefits

that accrued to the local population, industry participants were able to create deeper relationships with local communities built on fairness and trust (van Schalkwyk & Canares, 2021, 136).

Risk Avoidance

Researchers reviewed nearly 50,000 procurement packages from Paraguay that were formatted via OCDS and found that more than 90% of the filed protests could have been identified in the pre-solicitation phase using a model trained on contracts stored in OCDS. Protests against 92% of the same data set ultimately validated the results, demonstrating there is a realistic and feasible approach to detect anomalies in procurements before publication (Niessen et al., 2020, 127).

Uniformity

With Moldova's population disproportionately suffering from HIV and Hepatitis when compared to others in Europe, the adoption of OCDS was critical to improving savings on medicines – 19% on HIV drugs alone – as well as medical procurements generally, which saw savings of 14.5%. Even more critically, the organizations that distributed the drugs that partnered with the government adopted the open contracting approach themselves, improving the delivery of services in the economy generally (Open Contracting Partnership, 2021).

Open Contracting in the United States

Beyond the support for a procurement system that incorporates these desiderata, there has also long been bipartisan support for open government data, with the December 17, 2002 E-Government Act of 2002 (2002), the September 26, 2006 Federal Funding Accountability and Transparency Act of 2006 (2006), and the January 21, 2009 memorandum establishing a policy of transparency via open government (Presidential Memoranda, 2009) being among the most significant early attempts to adopt the approach in the digital age. The May 9, 2013 EO (Exec. Order No. 13642, 2013) and accompanying Office of Management and Budget ("OMB") (M-13-13, 2013) memo that make machine-readable data the default for government to increase openness are some additional examples of Democrat and Republican administrations pursuing open government data, including as it relates to contracting and procurement data (M-19-23, 2019). Together, these and other efforts related to open government provide the necessary foundation for the United States to move towards this data-based procurement approach.

Open Government's Influence Open Contracting

The open government movement's initial start was built on eight principles, the fifth of which was that the data should be "machine processable" and the last three which state that data should be open and available to anyone who wants to use it (Tauberer, 2014). State and local governments in the United States already recognized the power of converting procurements to data – the Arlington Independent School District used a digital procurement system to reduce the time it takes to onboard vendors, from months to a year down to just minutes (Levinson, 2025) while the state of North Carolina has been able to aggregate spending across "state agencies, community colleges, school systems, and local governments" for more than $78 billion in purchases since the start of their "NC eProcurement" system (NC Dep't of Admin. Div. of Purchase & Contract, 2025). Dozens of others of local governments adopted digital procurement systems and those who have centralized procurement can use these systems "as a way to navigate complex procurement policies [and] to maintain control and enhance their information capacity" (Chen et al., 2021, 917).

Digital procurement or e-procurement is not always open procurement – the Arlington Independent School District solution, for example, is built around proprietary technology (Levinson, 2025). On the other hand, Montgomery County, Maryland's legislation mandated the use of open government data (Montgomery County Open Data Act, 2013), resulting in their contracts being updated daily while enjoying administration-level support for mandating the use of open data formats (Montgomery County, MD, 2025). While not OCDS-based, it is part of a broader effort to create an "Open Government Partnership" that is aimed at making "governments more efficient, effective, and responsive by embracing open government innovation" (Montgomery County, MD, 2012). The city of Washington, D.C. does something similar in that they publish their procurement data on a regular basis (Dist. of Columbia, 2018), but not in the same file format as Montgomery County, meaning they would need to create an API to create automated processes between their systems.

That is the problem OCDS promises to solve by serving as "a common structured data model" intended to provide governments with a "deeper analysis of contracting data" (Open Contracting Partnership, 2024). In the United States, the Austin-led effort to adopt OCDS in the contracting effort for their courts led to results that were successful enough to inspire a scaling out of the implementation throughout other offices within the Austin city government (Open Government Partnership, 2019) – but the concept has not spread to the federal level. While the United States is a member of and could be considered among the most important founders of the Open

Government Partnership (Piotrowski, 2017, 159), the United Kingdom was the first G7 country to commit to using OCDS from May 2016 (Crown Commercial Service, 2021).

Legislative Foundations for Open Contracting

As previously mentioned, there have been efforts in the United States to improve procurement data via USAspending.gov – this is the result of the enactment of the Digital Accountability and Transparency Act of 2014 ("DATA Act"), signed into law on May 9, 2014 (DATA Act, 2014), "which required agencies to publish government spending information in open formats" (Begany & Martin, 2017). With Congress recognizing that the American taxpayer needed increased transparency "to know more about how their tax dollars are being spent," one of the most important advances of the DATA Act was the requirements for the creation of "[g]overnment-wide financial data standards" (S. Rept. 113-139, 2014, 3-4) now implemented as the Government Spending Data Model ("GDSM") (BFS, 2025). The GDSM, along with other developments associated with the DATA Act resulted in the launch of the open beta version of USAspending.gov on November 10, 2015 (Lebryk, 2015).

Going further, the Open, Public, Electronic, and Necessary Government Data Act ("OPEN Government Data Act"), which passed as part of the Foundations for Evidence-Based Policymaking Act of 2018 ("Evidence Act") on January 14, 2019, requires that agencies "ensure that any public data asset [...] is machine-readable" (Foundations for Evidence-Based Policymaking Act of 2018, 2019) In other words, it "would direct federal agencies to publish all data they collect in an open format that can be used by any computer" (CBO, 2016), essentially codifying the 2013 EO and the policies focused on implementing it that followed (Begany & Martin, 2017). As part of the OPEN Government Data Act, each agency was required "to establish a comprehensive data inventory" that is intended to include every data asset (Stuessy, 2022).

While the OPEN Government Data Act requires critical data be made available to decision makers, the delineated responsibilities "to make data open by default" include changes to 44 USC 3506(b)(2) to require strategic plans for managing that data (Foundations for Evidence-Based Policymaking Act of 2018, 2019). Influencing that decision and the passage of the law itself, the concept of transparency "and ensuring that evidence produced is made publicly available" was identified among the guiding principles. On the inventory itself, Congress stated that it would be "extremely valuable" as it would "allow [users] to quickly ascertain the scope of data products available and the location of the available data" so they can use it as needed (H.R. Rep. No. 115-411, 2017, 2, 12-13).

Executive Actions Intended to Implement Open Data

At the same time as Congress was working on this issue, so too was the Executive Branch, with a relevant starting point being the May 9, 2013 issuance of M-13-13, a memorandum that required the government "to promote openness and interoperability, and properly safeguard systems and information." It required that information be treated as a strategic asset stating that it will help the government "increase operational efficiencies, reduce costs, improve services, support mission needs, safeguard personal information, and increase public access to valuable government information" while also helping "fuel entrepreneurship, innovation, and scientific discovery" for all Americans (M-13-13, 2013, 1). The memo's supplemental guidance provided additional information on how to implement the inventory, but more critically, established the use of CKAN and a JSON-based schema (OMB, 2013). That schema, now called DCAT-US Schema v1.1 (OMB, 2025), is based on the W3C-maintained Data Catalog Vocabulary ("DCAT") (W3C, 2024). Nearly a year after M-13-13 on February 14, 2014, the Executive Branch issued M-14-06, providing agencies with guidance on how to use open administrative data for statistical purposes. Specifically to "provide the foundation for the research, evaluation, and analysis" the United States needs to undertake in order to better understand the way the American public is changing, how policies are impacting or addressing those needs, and how to improve the way those needs are being addressed (M-14-06, 2014, 1).

Several years later, on June 4, 2019, the Administration issued M-19-18 to establish "A Framework for Consistency" recognizing that "[n]ew tools, technologies, and norms" will allow the U.S. Government to make greater use of data but that success is dependent on "consistency in skills, interoperability, and best practices in how agencies use and manage data" (M-19-18, 2019). Accompanying that memo was an elevation of the importance of managing data at an enterprise-wide scale from the perspective of the Federal Government as a whole, not individual agencies to a Cross-Agency Priority in the President's Management Agenda, requiring the implementation of a Federal Data Strategy (Performance.gov, 2021). Shortly afterwards, on June 28, 2019, the Administration issued M-19-21, requiring agencies to "transition to electronic records" to achieve a "paperless government" by transitioning all of their permanent records to a digital format, among other requirements (M-19-21, 2019). Together, those two memorandums set the stage for the following month on July 10, 2019 when M-19-23 was issued, implementing Phase 1 of the Evidence Act, focused

on "Learning Agendas Personnel, and Planning" (M-19-23, 2019), with Phase 2 being identified as "Open Data Access and Management."

Several more years later, on January 15, 2025, M-25-05 was issued, implementing Phase 2 of the Evidence Act, requiring agencies move to data that is open by default, contained in a comprehensive data inventory managed by that agency, listed in a Federal Data Catalog, have a plan to keep data open within appropriate levels of access, and to engage with the public to discuss usage of public data and to encourage participation in the process while being structured or semi-structured to the extent feasible (M-25-05, 2025). To execute on these ideas, on the same day, the Administration released M-25-06, re-establishing the Chief Data Officer Council (M-25-06, 2025). When combined with the Executive Order issued on April 15, 2025, titled "Restoring Common Sense to Federal Procurement," aimed at rewriting the Federal Acquisition Regulation (Exec. Order No. 14275, 2025), as well as the Executive Order issued the next day on April 16, 2025, titled "Ensuring Commercial, Cost-Effective Solution in Federal Contracts" (Exec. Order No. 14271, 2025), this is among the best opportunities that have ever arisen to make the most of private sector technological advances applied to public sector problems and doing so in an actually modern, digital, and open way of contracting.

Conclusion

A modernized, data-driven procurement system requires more than mere digitization – and the efforts of the last several decades have resulted in a moment where success to create that system is, at long last, achievable. Tracing the historical roots of automation, the importance of standardization is clear, especially as it relates to storing and pulling data from paper or documents. Going further, that standardization will pave the way for advanced forms of machine-readable material without compromising human-readable formatting. Globally, digital procurement and open procurement systems have emerged as ways to capitalize on these advances. While there are several approaches and formats available to implement open contracting, the OCDS has become the globally accepted standard at national government levels equivalent to the United States. With the foundation provided by legislative and executive activities that have taken place in the United States since the launch of OCDS ten years ago, there is a real opportunity for a full U.S. embrace of the benefits of an open contracting data standard. Whether adapting OCDS or developing a new standard, applying a consistent data standard across all public sector procurements will be crucial to unlock the full potential of AI and cloud-based solutions as they are applied to public sector procurement, achieving the cost-savings,

productivity, and economy-sparking boons that a well-functioning procurement system promises.

Dedication

To the memory of Touraj Ghaffari and Parichehr Meshkat.

Acknowledgements

The author wishes to thank: my wife, Bridget Fields; Christopher Yukins; Mathew Blum; Jennifer Kuk; Jeffrey Koses; Michael Garland; Michelle McNellis; Mark Hopson; Alla Seiffert; Joslann Feldpausch; Lee Dunn; Leah Popoff; Dave Standish; Aubrey Milham; Andrew Howell; the Core Four; Ryan Thurlwell; Jonathan Hooper; Amy Shevlin; Peter Huy; BKMM; Samuel Paper and the PM; and my entire family of Leila, Amir, and Piom Ghaffari and Kate, Jimmy, and Jackie Fields.

References

AAS. (2000, November 28). *Machine Readable Tables in the AAS Journals*. https://aastex.aas.org/MR/aasetables.html

Adam, I., Dávid-Barrett, E., & Fazekas, M. (2025). The Political Economy of Open Contracting Reforms in Low- and Middle-Income Countries. *Governance*, *30*(2). https://doi.org/10.1111/gove.12897

Adam, I., Fazekas, M., & Tóth, B. (2020, January 20). *Measuring the benefits of open contracting: Case studies on Mexico, Paraguay, and Slovakia* (GTI-WP/2020:01). Government Transparency Institute. https://www.govtransparency.eu/measuring-the-benefits-of-open-contracting-case-studies-on-mexico-paraguay-and-slovakia/

Bachtiar, A., Suhardi, & Muhamad, W. (2020, October 19-23). *Literature Review of Open Government Data* [Conference presentation]. 2020 International Conference on Information Technology Systems and Innovation (ICITSI), Bandung, Indonesia. https://doi.org/10.1109/ICITSI50517.2020.9264960

Bausá Peris, O. (2014, May 20). History of UBL. *XML.org*. https://ubl.xml.org/wiki/history-of-ubl

Bedini, S. A. (1964, Winter). The Role of Automata in the History of Technology. *Technology and Culture*, *5*(1), 24-42. https://doi.org/10.2307/3101120

Begany, G., & Martin, E. (2017, May 5). The OPEN Government Data Act: What's at Stake? *Rockefeller Institute for Government*. https://rockinst.org/blog/open-government-data-act-whats-stake/

Beth, E. (2007). *Integrity in Public Procurement: Good Practice from A to Z*. Organisation for Economic Co-operation and Development (OECD). https://ssrn.com/abstract=987026

BFS. (2024, September 30). *Federal Spending*. U.S. Department of Treasury. https://fiscaldata.treasury.gov/americas-finance-guide/federal-spending/

BFS. (2025, April 25). *Government Spending Data Model (GSDM)*. U.S. Department of the Treasury. https://tfx.treasury.gov/data-transparency/gsdm

Booeshaghi, A. S., Chen, X., & Pachter, L. (2024, April 5). A Machine-Readable Specification for Genomics Assays. *Bioinformatics*, *40*(4). https://doi.org/10.1093/bioinformatics/btae168

Bromley, A. G. (1982, July-September). Charles Babbage's Analytical Engine, 1838. *Annals of the History of Computing*, *4*(3), 196-217. https://doi.org/10.1109/MAHC.1982.10028

Brown, S. (2020, September 30). Citizens empowered: An open secret to building local infrastructure on time and budget. *Open Contracting Partnership*.

https://www.open-contracting.org/2020/09/30/citizens-empowered-an-open-secret-to-building-local-infrastructure-on-time-and-budget/

CBO. (2016, December 5). *Cost Estimate: S. 2852 OPEN Government Data Act*. https://www.cbo.gov/sites/default/files/114th-congress-2015-2016/costestimate/s2852.pdf

CBO. (2024, December 20). *Artificial Intelligence and Its Potential Effects on the Economy and the Federal Budget*. CBO Publications. https://www.cbo.gov/publication/60774

Census Bureau. (2021, October 8). *From Tally Marks to Modern Computers — The Early Evolution of Census Data Processing*. Resource Library. https://www.census.gov/newsroom/blogs/research-matters/2012/06/from-tally-marks-to-modern-computers-the-early-evolution-of-census-data-processing.html

Census Bureau. (2024, August 14). *The Hollerith Machine*. Resource Library. https://www.census.gov/about/history/bureau-history/census-innovations/technology/hollerith-machine.html

Census Bureau. (2025, February 12). *Notable Alumni Herman Hollerith*. Resource Library. https://www.census.gov/library/photos/1880/herman-hollerith.html

Census Bureau. (2025, February 12). *Notable Alumni John Shaw Billings*. Resource Library. https://www.census.gov/library/photos/1880/john-shaw-billings.html

Census Bureau. (2025, March 5). *Citing our Data, Tools, Technical Documents, and Research*. https://www.census.gov/about/policies/citation.html

Chen, Y., Bretschneider, S., Stritch, J. M., Darnall, N., & Hsueh, L. (2021, February 3). E-procurement system adoption in local governments: the role of procurement complexity and organizational structure. *Public Management Review*, *24*(6), 903-925. https://doi.org/10.1080/14719037.2021.1874497

Chen, Z., Liu, Y., Shi, L., Wang, Z.-J., Chen, X., Zhao, Y., & Ren, F. (2025, January 25). MDEval: Evaluating and Enhancing Markdown Awareness in Large Language Models. *arXiv*. https://arxiv.org/abs/2501.15000

CKAN. (2025, April 16). *API Guide - CKAN 2.11.3 documentation*. https://docs.ckan.org/en/2.11/api/index.html

Crockford, D. (2009, July 27). *The JSON Saga* [Video]. Yahoo User Interface Library. https://www.youtube.com/watch?v=-C-JoyNuQJs&t=122s

Cronin, M. A. (1985, June). Fifty Years of Operations in the Social Security Administration. *Social Security Bulletin*, *48*(6). https://www.ssa.gov/history/cronin.html

Crown Commercial Service. (2021, January 1). *Open Contracting: Transparency data*. https://www.gov.uk/government/publications/open-contracting

DATA Act. (2014, May 9). *Pub. L. No. 113-101*. https://www.congress.gov/bill/113th-congress/senate-bill/994

Dist. of Columbia. (2018, July 13). *Office of Contracting and Procurement (OCP) - Contracts from PASS*. https://opendata.dc.gov/datasets/DCGIS::contracts-from-pass/about

Dollar, C. M. (1978, October). Appraising Machine-Readable Records. *The American Archivist*, *41*(4), 423-430. http://www.jstor.org/stable/40292816

Driscoll, K. (2012). From Punched Cards to 'Big Data': A Social History of Database Populism. *Futures of Communication*, *1*(1). https://doi.org/10.7275/R5B8562P

ECMA Int'l. (2017, December). *ECMA-404: The JSON Data Interchange Standard*. ECMA International. https://ecma-international.org/publications-and-standards/standards/ecma-404/

Edwards, D., Lopez Franco, E., & Shutt, C. (2020, September). *Open Up Contracting*. The TAP Room Consultants Ltd. https://www.hivos.org/assets/2020/10/ETE-Open-Up-Contracting.pdf

E-Government Act of 2002. (2002, December 17). *Pub. L. No. 107-347*. https://www.congress.gov/bill/107th-congress/house-bill/2458

Eichstädt, S., Hutzschenreuter, D., Niederhausen, J., & Neumann, J. (2022, September 19-21). *The Quality Infrastructure in the Digital Age: Beyond Machine-Readable Documents* [Conference presentation]. First International IMEKO TC6 Conference on Metrology and Digital Transformation, Berlin, Germany. https://doi.org/10.21014/tc6-2022.042

Exec. Order No. 13642. (2013, May 14). *Making Open and Machine Readable the New Default for Government Information*. 78 FR 28111. https://www.federalregister.gov/d/2013-11533

Exec. Order No. 14271. (2025, April 15). *Ensuring Commercial, Cost-Effective Solutions in Federal Contracts*. 90 FR 16433. https://www.federalregister.gov/d/2025-06835

Exec. Order No. 14275. (2025, April 15). *Restoring Common Sense to Federal Procurement*. 90 FR 16447. https://www.federalregister.gov/d/2025-06839

Federal Funding Accountability and Transparency Act of 2006. (2006, September 26). *Pub. L. No. 109-282*. Congress.gov. https://www.congress.gov/bill/109th-congress/senate-bill/2590

FFIEC. (2025). *Download Bulk Data*. FFIEC Central Data Repository's Public Data Distribution. https://cdr.ffiec.gov/public/PWS/DownloadBulkData.aspx

Foundations for Evidence-Based Policymaking Act of 2018. (2019, January 14). *Pub. L. No. 115-435*. https://www.congress.gov/bill/115th-congress/house-bill/4174

Furuta, R., Scofield, J., & Shaw, A. (1982, September 1). Document Formatting Systems: Survey, Concepts, and Issues. *ACM Computing Surveys (CSUR)*, *14*(3), 417-472. https://doi.org/10.1145/356887.356891

GAO. (2018, December 13). *Open Data: Treasury Could Better Align USAspending.gov with Key Practices and Search Requirements* (GAO-19-72). https://www.gao.gov/products/gao-19-72

Garland, M. (2023, January). *Vendor-Lock and Lack of Competition in the Government's Software Estate*. NetChoice. https://netchoice.org/wp-content/uploads/2023/01/NetChoice_Garland_The-Pernicious-Consequences-of-Vendor-Lock.pdf

Goyal, A., & Dyreson, C. (2019, December 12-14). *Temporal JSON* [Conference presentation]. 2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC), Los Angeles, CA, USA. https://doi.org/10.1109/CIC48465.2019.00025

Grillo, F. (2019). *Hero of Alexandria's Automata. A Critical Edition and Translation, Including a Commentary on Book One* [PhD thesis, University of Glasgow]. Enlighten Publications. https://doi.org/10.5525/gla.thesis.76774

Gruber, J. (2004, December 17). Markdown. *Daring Fireball*. https://daringfireball.net/projects/markdown/

GSA. (2025). *Federal Procurement Data System - Next Generation*. https://www.fpds.gov/

GSA. (2025, April 15). *USAspending: Government Spending Open Data*. https://www.usaspending.gov/

Heide, L. (2009). *Punched-Card Systems and the Early Information Explosion, 1880–1945*. Johns Hopkins University Press. https://doi.org/10.1353/book.3454

Hendler, J., & Pardo, T. A. (2012, September 24). A Primer on Machine Readability for Online Documents and Data. *Data.gov Developers - Blog*. https://web.archive.org/web/20170502210649/https://data.gov/developers/blog/primer-machine-readability-online-documents-and-data/

Hofstadter, D. R. (1999). *Godel, Escher, Bach: An Eternal Golden Braid*. Basic Books, Inc. https://www.hachettebookgroup.com/titles/douglas-r-hofstadter/godel-escher-bach/9780465026562/?lens=basic-books

Hollerith, H. (1889). *Electric Tabulating Machine* (U.S. Patent no. 395,781). U.S. Patent and Trademark Office. https://patents.google.com/patent/US395781A/en

Hoxha, J., & Brahaj, A. (2011, September 7-9). *Open Government Data on the Web: A Semantic Approach* [Conference presentation]. 2011 International Conference on Emerging Intelligent Data and Web Technologies, Tirana, Albania. https://doi.org/10.1109/EIDWT.2011.24

H.R. Rep. No. 115-411. (2017, November 15). *Foundations for Evidence-Based Policymaking Act*. https://www.congress.gov/congressional-report/115th-congress/house-report/411

Hurst, N., Li, W., & Marriott, K. (2009, September 16-18). *Review of automatic document formatting* [Conference presentation]. DocEng '09: ACM Symposium on Document Engineering, Munich, Germany. https://doi.org/10.1145/1600193.1600217

IEC. (2023, October 17). *Smart: Transforming standards for a digital world*. https://www.iec.ch/basecamp/smart-transforming-standards-digital-world

IEC. (2023, October 20). *Smart Standards – From a Market and Industry Perspective*. https://www.iec.ch/basecamp/smart-standards-market-and-industry-perspective

IEC-ISO. (2022, April). *SMART FAQ*. https://webdesk.jsa.or.jp/pdf/dev/md_5615.pdf

IETF. (2015, March). *The I-JSON Message Format* (RFC 7493). https://datatracker.ietf.org/doc/html/rfc7493

Int'l Journal for Digital Art History. (2023). *Map*. https://dahj.org/map

ISO-IEC. (2015, July). *Information technology — Open Document Format for Office Applications (OpenDocument) v1.2* (ISO/IEC 26300-1:2015). https://www.iso.org/standard/66363.html

ISO-IEC. (2017, November). *Information technology — The JSON data interchange syntax* (ISO/IEC 21778:2017). https://www.iso.org/standard/71616.html

JSON.org. (2017, December). *Introducing JSON*. https://www.json.org/json-en.html

Klinke, H. (2023, May 2). Articles As Data. Machine-Readable Writing. *Medium*. https://medium.com/@HxxxKxxx/articles-as-data-machine-readable-writing-55c4110adac9

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., & Battaglia, P. (2023, November 14). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416-1421. https://doi.org/10.1126/science.adi2336

Lebryk, D. (2015, November 10). Help Treasury Reinvent USAspending.gov. *Obama White House Archives*.

https://obamawhitehouse.archives.gov/blog/2015/11/10/help-treasury-reinvent-usaspendinggov

Lee, T. H. (2009, February 6). From Mechanism to Monolith. *IEEE Solid-State Circuits Magazine*, *1*(1), 69-75. https://doi.org/10.1109/MSSC.2008.930937

Levinson, K. (2025, April 3). *Texas school district sees efficiency gains with e-procurement*. Route Fifty. https://www.route-fifty.com/finance/2025/04/texas-school-district-sees-efficiency-gains-e-procurement/404270/

M-13-13. (2013, May 9). *Open Data Policy - Managing Information as an Asset*. Office of Management and Budget. https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf

M-14-06. (2014, February 14). *Guidance for Providing and Using Administrative Data for Statistical Purposes*. Office of Management and Budget. https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2014/m-14-06.pdf

M-19-18. (2019, June 4). *Federal Data Strategy - A Framework for Consistency*. Office of Management and Budget. https://trumpwhitehouse.archives.gov/wp-content/uploads/2019/06/M-19-18.pdf

M-19-21. (2019, June 28). *Transition to Electronic Records*. Office of Management and Budget. https://www.archives.gov/files/records-mgmt/policy/m-19-21-transition-to-federal-records.pdf

M-19-23. (2019, July 10). *Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Leaming Agendas, Personnel, and Planning Guidance*. Office of Management and Budget. https://www.whitehouse.gov/wp-content/uploads/2019/07/m-19-23.pdf

M-25-05. (2025, January 15). *Phase 2 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Open Government Data Access and Management Guidance*. Office of Management and Budget. https://bidenwhitehouse.archives.gov/wp-content/uploads/2025/01/M-25-05-Phase-2-Implementation-of-the-Foundations-for-Evidence-Based-Policymaking-Act-of-2018-Open-Government-Data-Access-and-Management-Guidance.pdf

M-25-06. (2025, January 15). *Re-establishing the Chief Data Officer Council*. Office of Management and Budget.

https://www.cdo.gov/assets/documents/M-25-06-Re-establishing-the-Chief-Data-Officer-Council.pdf

Marchessault, L. (2013). *Open Contracting: Accountability Through Enhanced Disclosure and Participation*. Commonwealth Governance Handbook 2013/2014. https://web.archive.org/web/20191206081311/http://www.commonwealthgovernance.org/assets/uploads/2013/11/8-Open-contracting.pdf

Martone, M. (2014). *Data Citation Synthesis Group: Joint Declaration of Data Citation Principles*. https://doi.org/10.25490/a97f-egyk

Mavidis, A., & Folinas, D. (2022, September 8). From Public E-Procurement 3.0 to E-Procurement 4.0; A Critical Literature Review. *Sustainability*, *14*(18). https://doi.org/10.3390/su141811252

Montgomery County, MD. (2012, December 17). *Proposed Framework for a Counties Open Government Partnership*. https://www.montgomerycountymd.gov/open/partnership.html

Montgomery County, MD. (2025, May 13). *dataMontgomery - Contracts*. https://data.montgomerycountymd.gov/Government/Contracts/vmu2-pnrc/about_data

Montgomery County Open Data Act. (2013, March 18). *Bill No. 23-12*. https://www.montgomerycountymd.gov/open/Resources/Files/SignedOpenDataBill_23-12.pdf

NARA. (2022, March). *Strategic Plan 2022-2026*. https://www.archives.gov/about/plans-reports/strategic-plan/strategic-plan-2022-2026

NC Dep't of Admin. Div. of Purchase & Contract. (2025, April 13). *NC eProcurement*. https://eprocurement.nc.gov/

Niessen, M. E.K., Paciello, J. M., & Fernandez, J. I.P. (2020, April 22-24). *Anomaly Detection in Public Procurements using the Open Contracting Data Standard* [Conference presentation]. 2020 Seventh International Conference on eDemocracy & eGovernment (ICEDEG), Buenos Aires, Argentina. https://doi.org/10.1109/ICEDEG48599.2020.9096674

NIH. (2023, January 23). *Data Management & Sharing Policy Overview*. U.S. Department of Health and Human Services. https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policies/data-management-and-sharing-policy-overview#after

OASIS Open. (2024, June 20). *Universal Business Language Version 2.4* (UBL-2.4). https://docs.oasis-open.org/ubl/UBL-2.4.html

OMB. (2013, May 9). *Supplemental Guidance on the Implementation of M-13-13 "Open Data Policy – Managing Information as an Asset"*. https://resources.data.gov/resources/m-13-13-guidance/

OMB. (2025, March 21). *DCAT-US Schema v1.1 (Project Open Data Metadata Schema)*. https://resources.data.gov/resources/dcat-us/

Open Contracting Partnership. (2016, November 28). 'Everyone sees everything': Overhauling Ukraine's corrupt contracting sector. *Medium*. https://medium.com/open-contracting-stories/everyone-sees-everything-fa6df0d00335

Open Contracting Partnership. (2018, April 9). The deals behind the meals: How open contracting helped fix Colombia's biggest school meal program. *Medium*. https://medium.com/open-contracting-stories/the-deals-behind-the-meals-c4592e9466a2

Open Contracting Partnership. (2020, April 4). *Global Principles*. https://www.open-contracting.org/what-is-open-contracting/global-principles/

Open Contracting Partnership. (2021, January 29). Diagnosis open: how open contracting is bringing down the cost of medicines in Chile. *Open Contracting Blog*. https://www.open-contracting.org/2021/01/29/diagnosis-open-how-open-contracting-is-bringing-down-the-cost-of-medicines-in-chile/

Open Contracting Partnership. (2021, March 5). *Open Contracting Data Standard*. https://documents1.worldbank.org/curated/en/744551614955316901/pdf/Open-Contracting-Data-Standard.pdf

Open Contracting Partnership. (2021, July 12). From open data to joined-up government: driving efficiency with BA Obras. *Open Contracting Blog*. https://www.open-contracting.org/2021/07/12/from-open-data-to-joined-up-government-driving-efficiency-with-ba-obras/

Open Contracting Partnership. (2021, November 23). Patients, experts, and openness: A powerful combination therapy for unaffordable medicines in Moldova. *Open Contracting Blog*. https://www.open-contracting.org/2021/11/23/patients-experts-and-openness-a-powerful-combination-therapy-for-unaffordable-medicines-in-moldova/

Open Contracting Partnership. (2022, December 19). *Development and Appreciation — Open Contracting Data Standard 1.1.5 documentation*.

https://standard.open-contracting.org/latest/en/history/history_and_development/

Open Contracting Partnership. (2024, January 29). *Governance — Open Contracting Data Standard 1.1.5 documentation*. https://standard.open-contracting.org/latest/en/governance/

Open Contracting Partnership. (2024, December 18). *The Open Contracting Data Standard*. https://www.open-contracting.org/data-standard/

Open Contracting Partnership. (2025, April 8). *Open Contracting Data Standard 1.1.5 documentation*. https://standard.open-contracting.org/latest/en/

Open Government Partnership. (2019). *Inclusion in Court Contracting (AUS0007)*. https://www.opengovpartnership.org/members/austin-united-states/commitments/AUS0007/

Open Knowledge Foundation. (2015, August 24). Human Readable. *Open Data Handbook | Glossary*. https://opendatahandbook.org/glossary/en/terms/human-readable/

Open Knowledge Foundation. (2025, February 5). *CKAN - The open source data management system*. https://ckan.org/

Papadopoulos, E. (2016). Heron of Alexandria (c. 10–85 AD). In *Distinguished Figures in Mechanism and Machine Science: Their Contributions and Legacies* (pp. 217-245). Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6366-4_9

Park, S., & Jayaraman, S. (2017). The wearables revolution and Big Data: the textile lineage. *The Journal of The Textile Institute*, *108*(4), 605-614. https://doi.org/10.1080/00405000.2016.1176632

PDF Association. (2023, June 14). *Arlington PDF Model*. https://pdfa.org/resource/arlington-pdf-model/

Performance.gov. (2021, January 7). *Leveraging Data as a Strategic Asset*. https://trumpadministration.archives.performance.gov/CAP/leveragingdata/

Pezoa, F., Reutter, J. L., Suarez, F., Ugarte, M., & Vrgoč, D. (2016, April 11-15). *Foundations of JSON Schema* [Conference presentation]. In Proceedings of the 25th International Conference on World Wide Web (WWW '16), Montréal, Québec, Canada. https://doi.org/10.1145/2872427.2883029

Piotrowski, S. J. (2017). The “Open Government Reform” Movement: The Case of the Open Government Partnership and U.S. Transparency Policies. *The American Review of Public Administration*, *47*(2), 155-171. https://doi.org/10.1177/0275074016676575

Presidential Memoranda. (2009, January 21). *Transparency and Open Government*. The White House.

https://obamawhitehouse.archives.gov/the-press-office/transparency-and-open-government

Puckett, C. (2009, July). The Story of the Social Security Number. *Social Security Bulletin*, *69*(2). https://www.ssa.gov/policy/docs/ssb/v69n2/v69n2p55.html

Radhakrishnan, P., Chen, J., Xu, B., Ramaswami, P., Pho, H., Olmos, A., Manyika, J., & Guha, R. V. (2024, September 10). Knowing When to Ask -- Bridging Large Language Models and Data. *arXiv*, *2409*(13741). https://doi.org/10.48550/arXiv.2409.13741

Roistacher, R. (1977). A Source Documentation Style Manual. *iassist Newsletter*, *1*(4), 32-36. https://iassistquarterly.com/public/pdfs/iqvol1_4_roistacher.pdf

Rowe, J. S. (1982, Summer). Primary Data for Historical Research: New Machine-Readable Resources. *RQ*, *21*(4), 351-356. https://www.jstor.org/stable/25826792

RSS Advisory Board. (2009, March 30). *RSS 2.0 Specification*. https://www.rssboard.org/rss-specification

Schema.org. (2025, March 24). *Getting Started with Schema.org using Microdata*. https://schema.org/docs/gs.html

Schneider, B. R. (1971, September). The Production of Machine-Readable Text: Some of the Variables. *Computers and the Humanities*, *6*(1), 39-47. http://www.jstor.org/stable/30199450

Schooner, S. L. (2011, May 21). Desiderata: Objectives for a System of Government Contract Law. *Public Procurement Law Review*, *11*, 103-120. https://dx.doi.org/10.2139/ssrn.304620

SEC. (2025, January 17). *Inline XBRL*. https://www.sec.gov/data-research/inline-xbrl

Sípos, G., Spác, S., & Kollárik, M. (2015, December). *Not in Force Until Published Online: What the Radical Transparency Regime of Public Contracts Achieved in Slovakia*. Transparency International Slovakia. http://dx.doi.org/10.13140/RG.2.1.4747.4001

S. Rept. 113-139. (2014, March 27). *Digital Accountability and Transparency Act of 2013*. https://www.congress.gov/congressional-report/113th-congress/senate-report/139/1

Stats Perform. (2025, April 27). *The Most Successful English Clubs*. Opta Analyst. https://theanalyst.com/eu/2025/04/most-successful-english-clubs-football

Stuessy, M. M. (2022, December 29). *The OPEN Government Data Act: A Primer*. Congressional Research Service. https://www.congress.gov/crs-product/IF12299

Tauberer, J. (2014, December 1). *The Annotated 8 Principles of Open Government Data*. https://opengovdata.org/

Truesdell, L. E. (1965). *The Development of Punch Card Tabulation in the Bureau of the Census, 1890-1940: With Outlines of Actual Tabulation Programs*. U.S. Government Printing Office. https://books.google.com/books?id=Krg-AAAAIAAJ&pg=PA29&focus=viewport&dq=related:UOM39015072113361&output=html_text

UEFA. (2025, December 5). *Club coefficients*. UEFA rankings. https://www.uefa.com/nationalassociations/uefarankings/club/

van Schalkwyk, F., & Canares, M. (2021, October 6-8). *Public procurement and social inclusion in the network society* [Conference presentation]. ICEGOV '21: Proceedings of the 14th International Conference on Theory and Practice of Electronic Governance, Athens, Greece. https://doi.org/10.1145/3494193.3494299

W3C. (2003, January 6). *XML Development History*. https://www.w3.org/XML/hist2002

W3C. (2008, November 26). *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. https://www.w3.org/TR/xml/

W3C. (2015, March 17). *RDFa 1.1 Primer - Third Edition*. https://www.w3.org/TR/rdfa-primer/

W3C. (2018, March 27). *XHTML 1.0: The Extensible HyperText Markup Language (Second Edition)* (XHTML 1.0). What is XHTML? https://www.w3.org/TR/xhtml1/#xhtml

W3C. (2018, March 27). *XHTML 1.0: The Extensible HyperText Markup Language (Second Edition)* (XHTML 1.0). Why the need for XHTML? https://www.w3.org/TR/xhtml1/#why

W3C. (2020, July 16). *JSON-LD 1.1*. https://www.w3.org/TR/json-ld/

W3C. (2024, August 22). *Data Catalog Vocabulary (DCAT) - Version 3*. https://www.w3.org/TR/vocab-dcat/

Wang, Z., Veličković, P., Hennes, D., Tomašev, N., Prince, L., Kaisers, M., Bachrach, Y., Elie, R., Wenliang, L. K., Piccinini, F., Spearman, W., Graham, I., Connor, J., Yang, Y., Recasens, A., Khan, M., Beauguerlange, N., Sprechmann, P., Moreno, P., … Tuyls, K. (2024, March 19). TacticAI: an AI assistant for football tactics. *Nature Communications*, *15*(1906), 1-13. https://doi.org/10.1038/s41467-024-45965-x

WHATWG. (2025, April 30). *Microdata*. https://html.spec.whatwg.org/multipage/microdata.html

Williams, M. E. (1977). The impact of machine-readable data bases on library and information services. *Information Processing & Management*, *13*(2), 95-107. https://doi.org/10.1016/0306-4573(77)90043-7

Williams, S. C.P. (2024, October 2). *Quantum research paves the way toward efficient, ultra-high-density optical memory storage* [Press Release]. Argonne National Laboratory. https://www.anl.gov/article/quantum-research-paves-the-way-toward-efficient-ultrahighdensity-optical-memory-storage

Wyatt, P. (2021, May 27-28). *Work in progress: Demystifying PDF through a machine-readable definition* [Conference presentation]. Seventh LangSec IEEE S&P Workshop at the IEEE Security & Privacy Symposium 2021, Virtual. https://langsec.org/spw21/papers.html#pdfReadable

XBRL Int'l. (2013, February 20). *The XBRL Standard* (XBRL 2.1). https://www.xbrl.org/Specification/XBRL-2.1/REC-2003-12-31/XBRL-2.1-REC-2003-12-31+corrected-errata-2013-02-20.html

XBRL Int'l. (2022, December 4). *What is iXBRL?* https://www.xbrl.org/the-standard/what/ixbrl/