

enigma

WHITEPAPER

Using Analytics to Transform the Regulatory Compliance Ecosystem for a more Innovative U.S. Economy





Outgrowing our Form Based Regulatory Compliance System

Government regulations place a large burden on the private sector, which bears the cost of complying with an ever-changing and growing set of rules. The public sector in turn faces the enormous costs of processing documentation to inform compliance decisions. The relentless addition of over two thousand federal, state, and local rules each year necessitates a new way to meet and monitor compliance. Government and the private sector should carefully explore collaborative, ecosystem-wide approaches to minimize the effort required to demonstrate and evaluate regulatory compliance, thereby freeing up scarce resources for more innovative, growth oriented initiatives.

How we think about regulatory compliance

Opportunities to minimize the burden of regulatory compliance can be found along three continuums. The first continuum is data transfer which refers to the format and structure of the compliance data which is transmitted to the government. Rather than complete time-consuming standard forms, technology today allows for more direct and secure access to source documents and records from which compliance-relevant information can be scraped for analysis. The second continuum involves data use which refers to the manner in which the transferred data is analyzed and applied to a compliance decision. Where today regulatory compliance data informs a binary “compliant / non-compliant” determination, more predictive analytics and trend identification could be generated with existing data, or when complimented more creatively with open source information. At minimum, predictive analytics could be used to improve the accuracy of audit selection, or to identify needed or redundant rules and regulations. The third continuum is data sharing, concerning the ability of agencies to exchange information to enable richer analysis and recognizes the incomplete picture any one regulator might have on an organization as it strives to comply with the rules of multiple agencies. A failed environmental audit could, for example, indicate a broader governance concern which impacts other areas of business, such as security.

Improving the way we regulate, and comply

To keep pace with or even surpass the rising tide of rules, regulatory compliance ecosystem partners must evolve from (1) a current state of form filling for only explicitly requested information, to (2) managing data exchanges through a clearing house of complete compliance data, and finally take the jump to (3) a model of a wholly automated data system that meshes standardized compliance data with open data sources to both minimize cost and fully use publically available information in order to garner a complete picture beyond just compliance. This will allow for seamless information sharing and produce valuable insights on regulated



populations. By evolving across the three continuums and models, the government can ultimately make compliance and regulatory bodies more effective and efficient.

The Need for a New Approach to Regulatory Compliance

On average, federal regulators impose 2,400 new regulations each year. The combination of new regulations with those already on the books costs the private sector over \$2 trillion in compliance expenses each year, or approximately 10% of U.S. annual GDP.^{i,ii} Over the last 60 years, the volume of federal regulation has grown eighteen-fold.ⁱⁱⁱ This increase in regulatory constraints diverts resources which would have otherwise contributed to productivity, innovation, investment, and growth.^{iv}

Regulations—which range from local to federal—raise the cost of doing business. In the pharmaceutical sector, for example, final drug review processes from the FDA—which may include submitting paperwork and interfacing with the agency multiple times—can take around 28 months, four times the mandated 180 days as required by U.S. law.^v This time-consuming and costly process delays timely delivery of drugs and depresses bottom lines, and reduces incentives for individuals and businesses to innovate.

The large number of legal and regulatory obligations require considerable private sector time and resources to address. Yet the information submitted to agencies often is used in silos by those whose sole job is to issue a verdict on the compliance state of a given organization.

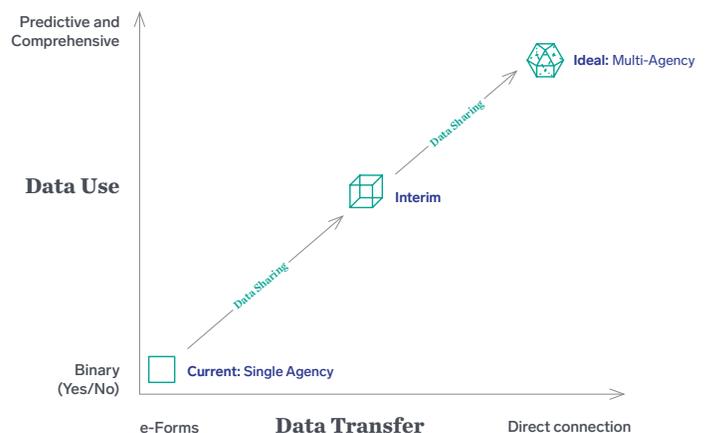
Complicating this dynamic are the lack of government resources to fact-check the details, and the absence of an incentive to look across government and open source data sets to paint a more complete picture of an organization. Regulators—strapped for time and managing significant caseloads—must act within their constraints; thus, they prioritize only certain compliance factors, thereby potentially missing major, albeit obfuscated violations. The future of regulatory compliance will ease the burden on companies, improve the efficacy of audit and inspection functions, reduce redundancies within the federal government, and maximize the value of publicly available and shared data.

Expanding the Regulatory Compliance Information Ecosystem

Credit bureaus of the 20th century evaluated an individual’s credit worthiness by relying on reports written by a network of agents; information that was difficult to share and analyze in bulk. This practice is not dissimilar to the form-based approach used today.

Credit scores were developed in 1989 by FICO as a result of this inefficacy, creating a standard representation of credit worthiness. Standards helped to greatly improve the flow of information, and the industry continues evolving. Credit worthiness is now measured in broader ways, too; for example, using a restaurant’s rating on a social network to evaluate credit worthiness. This represents an expansion of standardized information used in decision-making.

An improved approach to regulatory compliance would similarly enable a more efficient flow of information pertaining government regulation and business reporting. It would require improvement along the three identified continuum within the ecosystem—data transfer, data use, and data sharing. The following framework helps evaluate the trade-offs between today’s form-centric approach to information gathering and a future of seamless, autonomously-managed, and secure data exchanges. Improvement along the three continuums would necessarily require fundamental changes to the way government agencies collect, use, and share regulatory data today.





Data Transfer

This continuum focuses on the format of compliance data submitted to regulating bodies. Traditionally, most agencies receive compliance data through forms completed by companies and individuals, and submitted either electronically or on paper. These forms typically contain only the exact information necessary for attaining a substantive picture of compliance.

A simpler method of submission—one that does not involve filling out a multiple and often duplicative forms—will reduce burdens placed on the private sector and provide access to more comprehensive and useful data for the government. In order to accomplish these two goals, the government should first enable a direct connection with the companies to allow for a seamless, and repeatable collection of data. With a direct connection, major compliance agencies, such as the IRS and SEC, could gather information they normally receive through forms in a highly automated manner. Additionally, this would give agencies the ability to collect data on a time frequency determined by the unique risk profile of each company; the higher the risk of the company, the more often their data would be pulled for evaluation. The risk profile would be derived from both information on past performance and predictive factors driven by sophisticated analytic techniques. This would help both the government through more access to vital data, and the private sector by eliminating billions in compliance costs. High-risk companies get more attention, while low risk companies are pulsed only when absolutely necessary by law.

An even more elegant solution would involve regulators and industry agreeing on data standards to bring currently dissimilar and proprietary data frameworks into a common system. Agencies could then identify trends to link non-linear and/or otherwise unrelated risks in a meaningful, data driven way. Meanwhile, the private sector could gain access to valuable market data that would otherwise only be collected and made public by the government.

Data Use

This continuum revolves around the range of ways to analyze and use collected data. After gathering compliance information, agencies determine each data point's utility. Historically, agencies collected data for a binary decision: compliant or non-compliant.

A more innovative approach requires the government to use compliance data in a predictive manner about current and future compliance. This would allow for more effective allocation of scarce audit resources. A predictive approach would allow regulators to focus on companies that present the most risk. This approach has potentially significant effects across markets and issue areas. For example, information collected for financial compliance may unearth trends that help to further national security efforts.

Data Sharing

This continuum revolves around agencies' efforts to exchange information both internally and between government bodies. Data collected by an agency often has utility that extends beyond a single aspect of government. Health records of active military housed in the Department of Defense, for example, would be extremely helpful to Veterans Affairs, whose programs serve the same individuals when they retire.

Collected data is not shared efficiently—if at all today. Better sharing would increase the value and application of predictive analytics. Even basic improvements in government-wide data standards or inter-agency application program interfaces (APIs) would have enormous benefits. The metadata that could be collected and examined would provide insights into key industry and economic trends, as well as the overall workings of our economy to an extent yet unseen in either the private or public sector. A threshold investment requirement exists on this continuum, as well. While some agencies were built with express capabilities to transfer information across agencies, others would require major infrastructure investments to reach the desired end-state.



Towards a More Innovative Approach to Compliance

Today, most governments employ a model based on form field documents, only using information to answer basic compliance questions, and tend not to share data with other entities unless requested. A structured evolution involves a means for government to securely access companies' data and more dynamically check compliance, while simultaneously providing multiple agencies with the opportunity to use open data to generate insights and share them with government partners.

This would drastically reduce the volume and frequency of form submission requirements; instead, the responsibility for pulling and evaluating data for a compliance decision would fall to relevant government agencies.

What follows is an exploration of two future-state models in comparison with the current-state of regulatory compliance to begin a thoughtful and earnest public discussion on the merits and risks of each.

Three potential regulatory compliance models





EXPLICIT, SUBMITTED DATA

Current State: Form Field Filing

Federal agencies provide forms for the private sector to complete, requesting the bare minimum information necessary to reach compliance. The IRS has been working to simplify and streamline the filing process, so that private citizens and companies can fill out this baseline information as efficiently as possible.

Pros:

- Private information is minimally susceptible to hacking as only absolutely necessary information is conveyed
- Government receives all information in a standard format for analysis

Cons:

- Challenging to combine data with other open source information and conduct analytics
- High level of effort—in terms of financial cost and time—for the private sector to submit information in the exact required format
- Stifles creative use of data and overall innovation in government

CASE STUDY

Enhanced FDA reporting requirements

Life sciences companies are tackling the evolving and increasingly complex reporting requirements of the FDA by enhancing internal data operations and transparency. Form fields filing within the life sciences space is set to become complex as siloed, locked data—ranging from manufacturing to on-market drug use—is requested in cleaner and more timely formats.

Pros:

- Enterprises continue to feel comfortable working with data behind their own firewall
- Enterprises see opportunity to streamline their own data operations for efficiency and revenue-generating possibilities
- Artificial solutions can be applied to old data infrastructure to comply with enhanced reporting requirements

Cons:

- Enterprises lack visibility and access to their own data
- Internal data is locked in old tech infrastructure
- Data management is fragmented by department and geography, limiting transparency into all data required for reporting and minimizing the opportunity for a comprehensive picture of compliance across entities and other functional and geographic lines

The vast majority of regulators currently use some form of this model. Current funding levels need to be maintained or slightly expanded in order to keep up with the growing number of regulations, specifically in terms of keeping forms current and ensuring processing capabilities.



EXPLICIT, PULLED DATA

Managed Connection: Clearing House Distribution model

The government either pulls information through a direct connection or uses private citizen- and company-submitted data to a federal agency for a specific process (e.g., passport creation).

CASE STUDY

USPS

The U.S. Postal Service is sharing the information it collects for identification verification during the passport creation process with other agencies, helping citizens avoid the time and resource costs of receiving verification from multiple agencies. In the end state, government agencies like the DMV for driver's licenses, IRS for taxes, and OPM for clearances will no longer waste resources repeatedly collecting and verifying citizen's identities. Instead, they will receive accurate and timely data from a verified data clearing house.

Pros:

- Greatly reduces the redundancies citizens and companies face when interacting with multiple agencies
- Reduces the resources employed by the government to complete actions already taken by other agencies
- Provides a single, up-to-date source of verification

Cons:

- No fail safes exist to compensate for errors, potentially compounding problems if incorrect data is distributed to multiple agencies
- Requires a high level of coordination and technical infrastructure between agencies that is currently not in place

CASE STUDY

Regulatory Data Model Standardization

Life science organizations submit a wide variety of regulatory documents to bodies such as the FDA and EMA (the European Medicines Agency). While there are considerable overlaps in the content of these submissions (from manufacturing audits to drug applications), form fields and data standards vary drastically. To counteract such inefficiencies, both FDA and EMA have worked towards harmonizing data standards for regulatory submissions, thereby reducing duplicate efforts to submit and validate data about drug names, chemical composition, packaging, and more. This is intended to free up resources and improve the accuracy of submissions for both companies and regulatory agencies. Enigma is currently working with such enterprises to harmonize data standards to improve submission processes and operations.

Pros:

- Greatly reduces the redundancies companies face when interacting with multiple agencies and departments when there is one data model
- Enables central governance and oversight of compliance operations via one standard data model
- Reduces the resources needed to manage regulatory submissions across different sub-agencies.
- Provides a single source of verified information across all enterprise data

Cons:

- Requires a high level of coordination and technical infrastructure internally between departments that is currently not in place
- New data models have potential to upset workflow and other applications without coordination across enterprise software and training

Some agencies and regulators are already exploring this model. The basic infrastructure of receiving and processing the data exists, but significant investments would have to be made to facilitate the information sharing among a more regulators.



UNBOUNDED, PULLED DATA

Automated, direct connection: Direct Connect to Metadata Approach

Government agencies use a secure direct connection with private companies' records and books, pulling requisite compliance information and additional data points. This data pull is combined with open source information for further compliance verification. In addition, metadata will be produced primarily for agencies to conduct mission related analytics. Metadata will also be partially released for public consumption.

CASE STUDY

TSA

TSA is working to improve its compliance capabilities by introducing open source, unstructured data analytics. This should help verify compliance status through a custom-built, persistent monitoring platform. We envision a model wherein all compliance data is made available to TSA directly by companies, with information not relating to compliance interpreted through Natural Language Processing and text analytics. TSA would then combine that information with open source data to not only make a binary compliance determination, but also identify and prioritize across high risk populations and geographies. Through risk-based analytics, TSA could then prioritize grant allocation with a much clearer understanding of the most vulnerable companies and locations.

Pros:

- Minimal to no effort required on part of the private sector in terms of providing requisite data, saving significant costs in filing and compliance work and contracting
- Government has access to several data points that can be used for advanced and predictive analytics and persistent company monitoring
- Metadata could provide valuable insights for other agencies

Cons:

- Privacy and competitive concerns around sharing mandatory company information
- Involves high costs to the government in setting up the process to pull the correct information outside of just the standardized data points and channel it all to the appropriate entities
- Presents a big data problem with exposure to the governing "5 Vs": Velocity (frequency of data generation), Volume (exponential growth of company data), Variety (combining structured and unstructured), Veracity (trust in the data), and Viability (relevance and feasibility of use).^{vi}

CASE STUDY

Collaborative and Proactive Compliance

Several financial services and life sciences companies are deploying monitoring platforms to proactively track and predict compliance events in the field of Pharmacovigilance and Anti-Money Laundering. Developing effective monitoring platforms requires live access to public and private data. Monitoring patient safety reports requires accessing safety reports received by the company and cross-referencing against reports by the FDA or other global agencies. Similarly, anti-money laundering activity requires streamlined access to data across any account-type, cross-referenced against agency-level disciplinary or complaints data. Only by aggregating public and private data can companies begin to effectively leverage monitoring platforms to detect signals in drug quality or anti-money laundering activity to prevent further harm.

Pros:

- Enterprise access to public and private data allows for self-correcting mechanisms to prevent or contain the escalation of illicit activities
- Development of live data access of public and private sources allows enterprises to stay ahead of regulatory sanctions before issues arise



Cons:

- Requires investment in a standardized data model and transparent data-sharing within an enterprise to mesh with government data
- Technical challenges involve working with enterprises to set up the process to pull the correct information from many different types of unstructured data sets and channel to appropriate entities

Conclusion

An optimal model will take time and money to realize, particularly building the technological infrastructure for a seamless direct connection to appropriate data for government access. An interim step of moving to a managed exchange focused on pulling only necessary data may be more politically viable. In addition, workforce retraining is required to teach employees how to translate data into actionable insights.

The private sector is looking for a way to lower compliance costs, while the public is looking for ways to better utilize data—making this a meaningful discussion for both the regulator and the regulated. Deloitte’s and Enigma’s expertise and experience at the intersection of compliance, public private partnerships, and analytics makes for an effective partner on this journey. We support the consideration of a whole government push to move all compliance and regulatory bodies toward a more effective and efficient model.

- <http://dailysignal.com/2015/05/27/were-paying-more-than-ever-for-government-to-regulate-us/>
- <http://www.heritage.org/research/reports/2016/05/red-tape-rising-2016-obama-regs-top-100-billion-annually>
- Ibid.
- <https://www.aei.org/publication/regulations-are-a-really-big-drag-on-us-growth/>
- <http://www.cato.org/pubs/pas/pa-303.html>
- http://www2.deloitte.com/content/dam/Deloitte/it/Documents/deloitte-analytics/bigdata_challenges_success_factors.pdf