



**WHITE PAPER**

A detailed black and white photograph of a complex mechanical gear system. The image shows numerous interlocking gears of various sizes, some with teeth that are sharp and others that are more rounded. The gears are set against a dark background, and the lighting highlights the metallic surfaces and the intricate details of the machinery. The overall composition is dense and technical, suggesting precision and complexity.

**GET IT OUT - GET IT RIGHT:  
WHY MACHINES ARE BETTER  
AT HANDLING PUBLIC RECORDS  
REQUESTS**

## Management summary

Government belongs to the people, and federal, state and local agencies have the duty to be as transparent as they can within the law. This task is no longer possible without specialist software to handle public records requests (PRRs) and many, mainly federal agencies, have adopted these technologies to great effect. Our reasons here for urging all agencies to use advanced technologies to process PRRs have nothing to do with cost, or even the cost-effectiveness of specialist software. That argument can wait. Our sole criterion in this white paper is simply how technology best serves the ideal of open government.

# Contents

<b>Management summary</b> .....	<b>2</b>
<b>Contents</b> .....	<b>3</b>
<b>Introduction:</b> .....	<b>4</b>
<b>The Scale of the problem:</b> .....	<b>6</b>
<b>Five reasons why agencies must use specialist software to handle PRRs</b> .....	<b>8</b>
<b>1. Efficiency:</b> .....	<b>8</b>
<b>2. Consistency:</b> .....	<b>9</b>
<b>3. Transparency:</b> .....	<b>10</b>
<b>4. Security:</b> .....	<b>10</b>
<b>5. Control:</b> .....	<b>10</b>
<b>Agencies that are taking control</b> .....	<b>11</b>
<b>Conclusion</b> .....	<b>12</b>

## Introduction

In January 2009, President Obama issued a Memorandum in which he reaffirmed his Administration's commitment to the Freedom of Information Act (FOIA). The President gave a clear legal steer to federal agencies to "adopt a presumption in favor of disclosure" and advocated them to use "new technologies" to administer the act.

These technologies were not new in 2009, much less so three years later when the House of Representatives heard evidence on how transparency in government could be improved by using specialist software. It is worth quoting part of the witness statement from Ms. Melanie Ann Pustay, Director of the Office of Information Policy at the Department of Justice. She writes: "An area that I believe holds great promise in increasing the efficiency of agency FOIA processes is the use of litigation software in the FOIA context. Agencies often have to manually review hundreds, if not thousands, of pages of paper and electronic records for both responsiveness and duplication before disclosure analysis can even be made."

---

*The use of litigation software in the FOIA context holds great promise*

---

Has this "great promise" been fulfilled?

In limited respects, yes. The House of Representatives heard in 2012 how agencies are attempting to contain the number of PRRs by pro-actively releasing information. One of the examples cited is the Department of State adding over 2,000 documents to its online Rwandan Declassification Collection. However, as we observed in our white paper, *The Hidden Costs of Public Records Requests*, data only seems to feed the appetite for more data, and the number of requests for disclosure rises steadily each year. In 2017, federal bureaus processed around 800,000 at a cost of \$500 million.

At the local and state level, the picture is no different: more requests, increasingly complex requests, mounting costs. As we shall see, agencies are sometimes ill prepared for the more complicated PRRs and the costs become prohibitive, both to them and the requesters.

Technology such as that used for eDiscovery is more necessary than ever. However, in April 2018, the FOIA Advisory Committee concluded that adoption of these technologies was still very patchy. It recommended that the Department of Justice collect “detailed information ... regarding the specific methods and technologies agencies are using to search their electronic records, including email”. Of course, many agencies have made great strides in automating their PRR procedures, yet in its tone and content this latest recommendation is very similar to testimony presented to the House of Representatives in 2012.

This white paper explores the question of what technology is needed to fulfil PRRs. For now, considerations of costs or cost savings are outside the scope of this analysis, but it is our belief that technology would still be indispensable even if it was not cost-effective. After surveying the extent and nature of the challenges facing agencies, we describe five areas where specialist technology is a game changer. We end on a positive note with a brief analysis of how some agencies are using specialist software to transform their relationship with the citizens they serve.

For a high-level overview of the costs involved in PRRs see our white paper, [The hidden costs of public records requests](#). A more granular analysis of how specialist software can help bring down these costs is found in our study, [Building the business case to use technology for handling public records requests](#).



## The scale of the problem

A few years ago, a reporter from the Associated Press wanted to determine whether Californians who were poor were more likely to have their driver's licenses suspended than richer inhabitants of The Golden State. A request was submitted with the Department of Motor Vehicles for a breakdown of suspended licenses according to ZIP code. The DMV replied that such a request would take 120 hours of specialist programming at a cost of \$135 an hour. In addition, the DMV would charge a fee of 10 cents per record.

Fortunately for them, the reporter dropped this line of inquiry.

---

*The request would cost \$399,000 and take four years to fulfill*

---

Another example. Jason Parsley, executive editor of the South Florida Gay News, submitted a PRR with the Broward County Sheriff's Office for every email for a one-year period that contained derogatory terms for lesbians and gay men.

The sheriff's office replied that the request would cost \$399,000 and take four years to fulfill. The email system was not capable of searching all accounts simultaneously by typing in key words, Parsley was told, and each employee's account would have to be searched individually. A full-time member of staff would have to be hired to do it.

Broward County Sheriff's Lt. Eric Caldwell commented that the department was not trying to be evasive. He explained that employee emails were stored on a tape and kept at a remote archive facility. The tape or tapes would have to be retrieved physically and then converted into an Outlook file, which could then be searched.

The issue here is not whether these agencies fell short of that "profound national commitment" to open government called for by President Obama in 2009. Without the right technology in place, open government is often not an option, especially for smaller agencies such as the Broward County Sheriff's Office.

Lt Caldwell described the request as “very vague” – but the search capabilities of any good eDiscovery solution would have got some very specific answers in a matter of hours, and for a fraction of the estimated cost of \$399,000. The challenge is the unpredictability of PRRs; most state and probably all local agencies struggle with this. The AP request to the California DMV was simple enough, but because the bureau had not processed such a request before, and lacked the technology to handle any but the most straightforward PRRs, the reporter was frustrated in his inquiry.

---

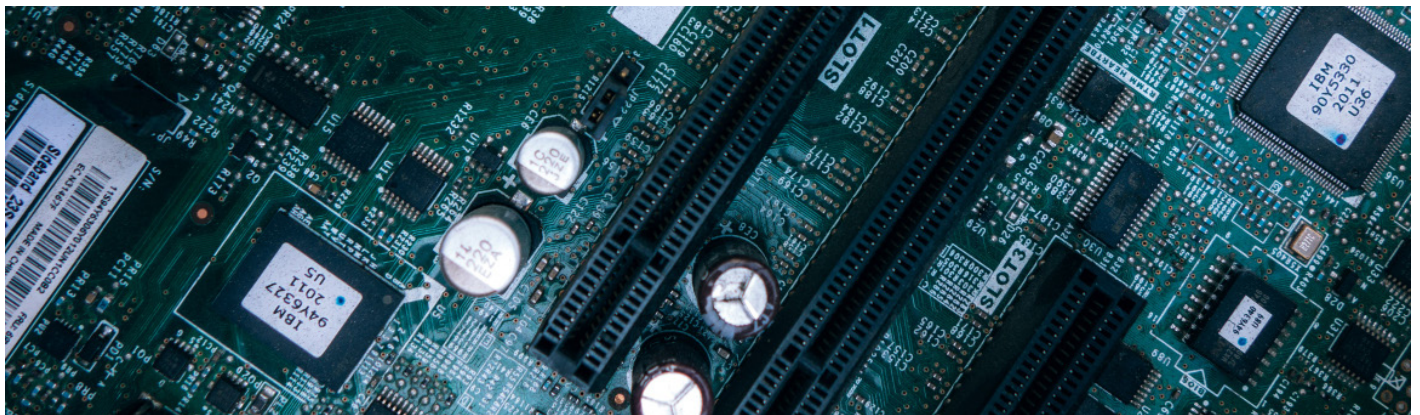
*The challenge is the unpredictability of PRRs*

---

The number of requests is also unpredictable. Whether you get 1 request a month or 100, they all need to be processed according to a strict and legally enforceable timetable, and with the same high standards of accuracy every time. Throwing more man hours at the problem doesn't make it go away, even if this were affordable or practicable. People aren't scalable. Ten times as many people, even highly trained and motivated people, does not mean 10 times faster, 10 times smarter, far from it.

But computers don't need caffeine.

“Never send a human to do a machine's job,” said Agent Smith from The Matrix. In the sections below, we give five reasons why he was right, and look at some agencies that took Agent Smith's advice.



## Five reasons why agencies must use specialist software to handle PRRs

### 1. Efficiency

Technology beats humans every time when it comes to identifying the right documents – and “right” in our context means both the documents to be disclosed, and those that must not be released. Specialist software has been proven to be more accurate as research into technology used for eDiscovery confirms time and again. Lawyers in particular, and human reviewers in general, tend to overestimate their own accuracy; machines also make mistakes, but always the same mistake, for the same reason. This makes machines much easier to correct.

---

*Machines also make mistakes, but always the same mistake. This makes machines much easier to correct.*

---

The efficiencies kick in from the very outset of the PRR process. In most instances, the corpus to be collected from is not stored at an out-of-town facility as was the case with the Broward County Sheriff’s Office, but centrally, possibly even in the cloud. With the right technology, agencies should be able to upload files directly from the usual sources such as Office 365 and SharePoint, as well as from FileNet and other, more esoteric, sources.

Let’s next consider search. Is it possible to search the corpus with standard software such as Adobe Acrobat? It is. IF the PDF contains searchable text. In those instances, keyword searches may work well enough. But how do you know which PDFs are searchable and which are not? Specialist technology will search smarter, much smarter. First of all, it makes everything searchable, even formats that are not natively searchable (TIF, PDF, BMP, PBG), or complex composite formats such as PST and ZIP. Even embedded objects – think of email attachments or Excel spreadsheets in Word or images in PowerPoint– are all made searchable and ready for analysis.

Now that the data has been collected, who will analyze and redact it? What if your corpus contains thousands or tens of thousands of documents? Who is going to manually select and blackline sensitive data such as names, telephone numbers and SSNs and add the code that tells you why something has been redacted or withheld? While Adobe Acrobat allows users to redact information on a per instance basis, it is a very manual, time-consuming and risk-filled process.



Protecting privacy and other sensitive information is enough of a challenge as it is. Text Mining automates redaction, which helps prevent breaches of even the most recent data legislation such as the California Consumer Privacy Act of 2018. The computer identifies email addresses, bank account numbers, social security numbers, telephone numbers, personal names, place names, company names – and recognizes them as such.

Structural analysis detects duplicate or near-duplicate documents such as Word files printed in PDF format, unravels email trails and identifies missing emails. The technology known as Assisted Review in eDiscovery greatly amplifies the power and accuracy of all your searches by (automatically) teaching the computer what to look for. In FOIA this kind of smart fact finding is used to identify for instance a class of sensitive documents on a particular topic that has to be withheld.

Try doing all that by hand.

## 2. Consistency

In its recent report on public records requests, the FOIA Advisory Committee paid tribute to the federal agency workers who between them processed over 800,000 requests in 2017. And rightly so, because the pressure on agency staff to handle PRRs is considerable, and growing. However, it is unfair and poor practice to rely on the individual heroics of agency workers to get the job done. Even the most dedicated FOIA professional will have an off-day: people are not machines.

Arguably, the biggest benefit of specialist technology is the platform: an organized, central location where all your data is standardized and all your PRR processes begin and end. A platform gives you control over who accesses your data, and documents the status of at every stage of fulfilment.

Specialist technology imposes consistency on the process. We know from eDiscovery in litigation that this greatly reduces the risk of legal error. And because you are handling requests in a consistent way, your success becomes repeatable and defensible.

### 3. Transparency

Last year, federal agencies spent \$40 million on FOIA litigation, adding 8% to the bill for processing those 800,000 requests. Delays or inadvertent disclosure of sensitive or privileged information can be grounds for legal action. In a court of law, you want to know why the error was committed or why the delay occurred, which is impossible if you re-invent your PRR procedures every time you get a request. With specialist software, your actions are traceable. This protects you in a law suit and gives you actionable insight to make sure the same mistake is not made again.

---

*With specialist software, your actions are traceable*

---

### 4. Security

The laws around PRRs are difficult to navigate because the legislation around what should not be disclosed is so open to interpretation. Technology itself has nothing to say about this legal no man's land but it can prevent errors that are beyond interpretation, such as data breaches and inadvertent disclosure of privileged information. Broward County Sheriff's Office is by no means the only agency that has to migrate sensitive information from System A to System B; in fact, all agencies that operate a patchwork PRR process expose themselves to unnecessary risk. The more systems, content repositories and connectors you throw at a task, the riskier it becomes.

Limit your processes, control your data – this is the discipline imposed by specialist software.

### 5. Control

If there is no consistent, repeatable process in place, there can be no effective process management. A lack of clear control saps staff morale and undermines the integrity of the process, such as it is. And if you do not control the process, you do not control costs – often, you do not even know what these costs are.

## Agencies that are taking control

Earlier, we looked at two agencies which, because of a lack of technology, were unable to resolve a PRR in a satisfactory manner. We should conclude on a note of optimism and describe how two federal agencies used eDiscovery tools to improve their disclosure processes.

In 2016, the Pipelines and Hazardous Materials Safety Administration (PHMSA) used de-duplication software and an eDiscovery database to review records. The tool allows multiple reviewers to work simultaneously on the same review and enhances collaboration among the reviewers. The product also permits sophisticated searches to target responsive records and has served as a customer-friendly negotiation tool. PHMSA has used the Discovery tool to suggest search terms and immediately report back to requesters on the search results, which has helped narrow the scope of the request and so reduce the number of documents under review.

---

*Technology eliminated a combined 81,820 non-responsive files, which would have required a manual search.*

---

In 2017, Immigration and Customs Enforcement (ICE) created a workflow in its eDiscovery tool that allowed staff to ingest thousands of electronic files into the tool and then use analytics to seek responsive records. The process now takes minutes rather than months. ICE shared this workflow with the Federal Emergency Management Agency (FEMA) and NPPD for use in their FOIA litigation cases. The workflow eliminated a combined 81,820 non-responsive files, which would have previously required a manual search.

## Conclusion

The cliché that data is the new oil or gold is in fact a recent one; nobody was saying this 10 years ago. Citizens, companies, government bodies – Big Data has taken us all by surprise to a greater or lesser extent. Nobody yet knows where this data revolution will lead, but we know technology will be at the heart of it.

Government bodies are by their very nature less agile than companies in how they allocate their scant resources. Public records requests are an administrative duty and burden, but agencies have not always had the time (and money) to self-assess how well they are discharging this duty. Most would admit it is a struggle.

President Obama knew almost 10 years ago that technology had to be brought to bear to protect the principle of open government. Where eDiscovery vendors have made great progress in the past decade is in making their software more accessible and user-friendly. If you can use Microsoft Outlook, you can use specialist technology to handle PRRs. Being “bad at computers” is no longer an excuse.

As we argue in our next white paper, *Building the Business Case to use Technology for Handling Public Records Requests*, the cost of the technology is no longer a deterrent either. The real price being paid by many state and local agencies and the citizens they serve is that of a costly and ineffective PRR process.



Agencies across the nation have been using ZyLAB's FOIA solution for efficient and cost-effective public records requests. With easy to use features such as auto-redaction, auto-identification of sensitive information and workflow management you can significantly save time, resources and costs associated with the burdens of responding to information requests.



an IPRO company

[IPRO.COM](http://IPRO.COM)