IR-01-26-23336

21 July 2026

Marcus
fyi-request-35001-3b0d9338@requests.fyi.org.nz

Dear Marcus

Request for information

Thank you for your Official Information Act 1982 (OIA) request of 30 June 2026, in which you requested copies of the Privacy Impact Assessments (PIAs) performed for Staff Safety Persons of Interest (SSPOI) and Search X.

Police has identified the following documents within the scope of your request:

- Privacy Impact Assessment High Risk Person of Interest
- Privacy Impact Assessment Search X

Copies of these documents are attached. Certain information has been withheld under section 9(2)(a) of the OIA to protect people's privacy.

We trust this response addresses your areas of interest. If not, you have the right to seek an investigation and review by the Ombudsman of the way your request has been handled. Information on how to do so is available at www.ombudsman.parliament.nz or freephone 0800 802 602.

Respectfully

Mike Webb
Chief Assurance Officer
New Zealand Police

Brief Privacy Analysis

SAS HR-POI Dashboard

PROJECT SUMMARY:

Overview

This is a risk-prioritisation model used to rank those who have assaulted police or have presented/discharged a firearm at a police officer. The algorithm is used to identify individuals who have a high chance of assaulting a frontline officer.

A data mining and regression modelling approach has been used to identify common features amongst individuals who in the past have either assaulted a police officer or presented or discharged a firearm at a police officer.

Methodology

As an initial step, the case notes of 34 people who had presented or discharged a firearm at a police officer in the last 2 years were examined in detail.

In total, all people with a person record number in NIA were assessed (2.5 million people) across 2,100 metrics. Metrics included: presence or absence in a table; counts of codes, flags and IDs, date and date-time frequencies; percentiles (including 10^{th} 25^{th}, 75^{th}, 90 percentiles), medians, means and counts of all continuous numeric data.

All individuals were then assigned one of two outcome categories. Those that had assaulted a police officer and those that hadn't.

The regression modelling identified several metrics that were common amongst individuals who had assaulted a police officer or who had presented or discharged a firearm at a police officer.

Key metrics (46) that were highly predictive included cleared offence counts, number of prosecutions, offending severity, family violence offending, recency of offending, frequency of alerts etc. These features were also highlighted via a deep dive into the case notes of these individuals.

Ethnicity, gender, age, gang affiliation, and location identifiers were not predictive or important to the model.

Outputs

The model was used to develop a matrix of scores, and a list of people who may, based on their offending history, present a risk to frontline staff in the future. The list is designed to inform decision making in relation to contact with these individuals.

The model, the matrix of scores, and the list are regularly checked, validated, and updated monthly. The dashboard updates daily.

Primary and secondary functions

The key function of this technology is as an early warning system to alert frontline staff to people who are high risk. This is carried out via a NIA Safety Alert (staff will be notified of the risk via a NIA Safety Alert).

Secondary functions include:

- Identifying high risk individuals with active warrants to arrest.
- Wrapping social services around high risk individuals to prevent further instances of crime and harm.

1.2 PERSONAL INFORMATION THAT THE PROJECT WILL INVOLVE

Type of personal Information	Source of Information	Purpose of information for the project
A wide range of variables associated with an individual NIA record. Examples include: *cleared offence counts, number of prosecutions, offending severity, family violence offending, recency of offending, frequency of alerts etc.*	NIA records	To form the basis of a risk identification and prioritisation model. Individuals are added to a list depending on the algorithm-generated risk ranking.

2. PRIVACY ASSESSMENT

2.1 AREAS THAT ARE RISKY FOR PRIVACY

Does the project involve any of the following?	Yes (tick)	No (tick)	If yes, explain your response
Information management generally			
A substantial change to an existing policy, process or system that involves personal information ***Example**: New legislation or policy that makes it compulsory to collect or disclose information*	✓		The HR-POI Dashboard uses statistical modelling and regression analysis to identify individuals who are likely to present a safety risk to frontline staff. The algorithm has a 90% accuracy rate. The list currently includes around 3,000 names. This approach is distinct from the approach to date, which has been structured around the national Top 5 Most Wanted. This has performed poorly as a predictive risk tool.

Does the project involve any of the following?	Yes (tick)	No (tick)	If yes, explain your response
Any practice or activity that is listed on a risk register kept by your organisation ***Example:*** *Practices or activities listed on our risk register or personal and safety register*		✓	[check re algorithm]
Collection			
A new collection of personal information ***Example:*** *Collecting information about individuals' location*		✓	
A new way of collecting personal information ***Example:*** *Collecting information online rather than on paper forms*		✓	
Storage, security and retention			
A change in the way personal information is stored or secured ***Example:*** *Storing information in the cloud*		✓	
A change to how sensitive information is managed ***Example:*** *Moving personal or financial records to a new database*		✓	
Transferring personal information offshore or using a third-party contractor ***Example:*** *Outsourcing the payroll function or storing information in the cloud*		✓	
Use or disclosure			
A new use or disclosure of personal information that is already held ***Example:*** *Sharing information with other parties in a new way*	✓		Police staff can be expected to rely on and take action based on the information provided in the dashboard. The reliability of the dashboard as a predictive tool depends on the accuracy of the algorithm (90%). Noting the inherent biases in Police - held data, particularly regarding ethnicity and gender. Noting the need to provide for appropriate controls over access to the dashboard. The expectation is that staff use the dashboard in conjunction with a TENR risk assessment, rather than as a predictive tool alone.
Sharing or matching personal information held by different organisations or currently held in different datasets ***Example:*** *Combining information with other information held on public registers, or sharing information to enable organisations to provide services jointly*		✓	
Individuals' access to their information			

Does the project involve any of the following?	Yes (tick)	No (tick)	If yes, explain your response
A change in policy that results in people having less access to information that you hold about them ***Example:*** *Archiving documents into a facility from which they can't be easily retrieved*		✓	
Identifying individuals			
Establishing a new way of identifying individuals ***Example:*** *A unique identifier, a biometric, or an online identity system*	✓		Currently there is no way to exclude a person or remove them from the list. There could be a benefit in reflecting the currency of the offending. Individuals will be flagged on NIA with an alert. The algorithm currently has no process for review at present.
New intrusions on individuals' property, person or activities			
Surveillance, tracking or monitoring of movements, behaviour or communications ***Example:*** *Installing a new CCTV system*		✓	
Changes to our premises that will involve private spaces where clients or customers may disclose their personal information ***Example:*** *Changing the location of the reception desk, meeting rooms where people may discuss personal details*		✓	
New regulatory requirements that could lead to compliance action against individuals on the basis of information about them ***Example:*** *New legislative conditions for a service i.e. vetting*		✓	
List anything else that may impact on privacy, such as bodily searches, or intrusions into physical space		✓	

2.2 INITIAL RISK ASSESSMENT

If you answered "Yes" to any of the questions above, use the table below to give a rating – **Low (L)**, **Medium (M)**, or **High (H)** – for each of the aspects of the project set out in the first column.

For risks that you've identified as Medium or High, indicate (in the right-hand column) how the project plans to mitigate or eliminate the risk (if this is known).

If you answered "No" to all the questions in 2.1 above, move on to section 3 below.

Aspect of the Project	Rating (L, M or H)	Describe any medium and high risks and how to mitigate them

Level of information handling L – Minimal personal information will be handled M – A moderate amount of personal information (or information that could become personal information) will be handled H – A significant amount of personal information (or information that could become personal information) will be handled	H	The Dashboard represents a substantial collation and processing of NIA-held data. Access to NIA and the Dashboard is limited to Police staff with a business need. Access to the list is only on a need to know basis (access to the HRPOI is restricted to approved staff that have been nominated by districts) who need access to the list for their roles (for example – DMIs, AOS Commanders, Intel staff). Access is controlled and must be requested through the BAU support process. Access will be regularly audited and reviewed.
Sensitivity of the information (eg personal, financial, race) L – The information will not be sensitive M – The information may be considered to be sensitive H – The information will be highly sensitive	M	Access to NIA and the Dashboard is limited to Police staff with a business need. Access to the list is only on a need to know basis (access to the HRPOI is restricted to approved staff that have been nominated by districts) who need access to the list for their roles (for example – DMIs, AOS Commanders, Intel staff). Access is controlled and must be requested through the BAU support process. Access will be regularly audited and reviewed.
Significance of the changes L – Only minor change to existing functions/activities M – Substantial change to existing functions/activities; or a new initiative H – Major overhaul of existing functions/activities; or a new initiative that's significantly different	M	The HR POI Dashboard represents a change in the way NZ Police identifies high risk individuals. An annual review of the ongoing suitability and function of the Dashboard is proposed by the business owner. A review of the algorithm and any adjustments could be part of the annual review process. A governance / oversight body is yet to be identified. [check]

Interaction with others L – No interaction with other agencies M – Interaction with one or two other agencies H – Extensive cross-agency (that is, government) interaction or cross-sectional (non-government and government) interaction	L	
Public impact L – Minimal impact on the organisation and clients M – Some impact on clients is likely due to changes to the handling of personal information; or the changes may raise public concern H – High impact on clients and the wider public, and concerns over aspects of project; or negative media is likely	M	There may be a public perception risk about Police using a algorithmic tool to targeting individuals who are a risk to staff safety, as opposed to the community It would be appropriate to refer the HR POI algorithm to the external panel for its consideration and comment. Messaging around this tool should reflect the objectives to protect staff and the wider community, with better informed frontline staff who can act to accurately prevent and deflect high risk interactions.

3. Summary of privacy impact

The privacy impact for this project has been assessed as:	Tick
Low – There is little or no personal information involved; or the use of personal information is uncontroversial; or the risk of harm eventuating is negligible; or the change is minor and something that the individuals concerned would expect; or risks are fully mitigated	
Medium – Some personal information is involved, but any risks can be mitigated satisfactorily	✓
High – Sensitive personal information is involved, and several medium to high risks have been identified	✓
Reduced risk – The project will lessen existing privacy risks	
Inadequate information – More information and analysis is needed to fully assess the privacy impact of the project.	

3.1 Reasons for the privacy impact rating

The privacy impact is medium to high. The information is wholly based on NIA-held information sources. The Dashboard and associated list of high risk individuals is sensitive material that would attract a high public interest.

4. Recommendation

A full privacy impact assessment is not required due to the expectation that the business owners are referring the Dashboard and algorithm to an oversight body. In addition, it is recommended that the external panel has the opportunity to review and comment on the algorithm.

5. Sign off

Include

- Project Manager

- Legal or Privacy Officer
- Senior or Executive Leader

Privacy Impact Assessment

Search X – Unstructured Search

Project Sponsor	[redacted] Manager Tactical Intelligence
Author and position	[redacted] Search X Project Manager
Date and version	*12 Aug 22, Version 0.2* *23 Aug 22, Comments,* [redacted] *Chief Privacy Officer* *25 Aug 22, Version 0.3,* [redacted]

NEW ZEALAND POLICE
Ngā Pirihimana o Aotearoa
Assurance Group

Project Summary

Search X has been established to better inform frontline safety by assisting intelligence analysts query and link together information found in the intelligence holdings.

Search X scope includes delivery of new unstructured search capability targeting the intelligence holdings, the development of a business case for improved visual intelligence analysis, and depending upon the outcome of the business case, the future deployment of software solutions.

This Privacy Impact Assessment specifically targets the implementation of improved unstructured text search operating across NIA attachments and narratives.

Project Detail

1. Purpose - collection, use, disclosure or sharing

Using currently held information for a new purpose

Search X provides intelligence analysts involved with the Tactical Response Model with a new capability to search unstructured text in NIA Noting narratives and NIA attachments. These analysts operate as a subset of the National and District intelligence teams (numbering less than a hundred).

Currently intelligence analysts can search NIA narrative content using the NIA application, Business Objects and SAS but each has limitations compared to the desired capabilities sought by the intelligence analysts. The NIA application has support for phonetic and near spelling search, Business Objects and SAS have support for wildcard characters but query syntax is complicated and it becomes exceedingly difficult to represent complex search patterns.

To deliver a text search capability across NIA, Search X proposes utilising an extended implementation of the existing IMT Search engine built on IBM Watson Explorer.

IBM Watson Explorer allows analysts to specify queries with a richer syntax that can make finding information simpler, examples of these capabilities are:

- Searching within ranges in a document defined by starting text and end text
- Searching for terms within a given number of words of each other (using either a within or near syntax)
- Regular expressions – a very technical but powerful method of identifying text in a document
- Hyphenation support
- Precedence by grouping search expressions to check that given text occurs before other text and
- Wildcard character support for multiple letters (using an asterisk) or any variation on a single letter (using a question mark).

The use of the IBM Watson Explorer search engine will enable faster discovery giving time efficiency benefits to the intelligence analysts. In general, the NIA application, Business Objects and SAS can all enable an analyst to find the same information as found via IBM Watson Explorer, but analysts would otherwise have to become proficient in defining complex queries or accept manually analysing many more irrelevant hits before getting to the final desired result.

Collaborating with others to gain insights

Collaboration between intelligence analysts and their colleagues across Police particularly including the front line teams does not alter based upon the introduction of new search capability based on IBM Watson Explorer.

2. Transparency, technology and social expectations

Notifying individuals and communities

The use of improved text search capability from Watson Explorer is not expected to require notification to interested parties. The search capability is deterministic. This means for a given input we would always get the same output. It would not use any artificial intelligence algorithms that could be subject to bias or would have to be treated as black boxes because we can't explain the outcome.

The narrative text searches on NIA are all technically possible to construct using existing BO/SAS/SQL query syntax but Watson Explorer makes the definition of the query significantly simpler.

Watson Explorer does enable searching NIA attachments. These can currently be searched via NIA but not via BO or SAS/SQL.

Use of technology

Search is enabled using IBM's Watson Explorer search engine.

Watson Explorer is already installed in Police and in use as the search engine in IMT Search.

Watson Explorer is also currently used by ICT as a search index across the Crystal Drives.

Watson Explorer can crawl across file content stored in folders or receive content for indexing. The IMT Search solution pushes content for indexing to Watson Explorer and this pattern is the proposed solution for NIA narrative content.

Involving service users and meeting social expectations

Search requirements have been sourced from the Tactical Intelligence analyst community.

3. Fairness and quality

Collecting, using, disclosing or sharing information respectfully and fairly

Search X Unstructured Search does not result in any new information being collected but it does aid in the extraction of insight from narrative text by making the query process more efficient (faster) and more effective (through better targeted search queries).

Search X enabled by Watson Explorer does provide a mechanism to search NIA attachments. This capability does not currently exist.

Information sources, data quality and purpose

Search X Unstructured Search initially targets NIA narrative content being the narrative field on a Noting or any attached document (which could be a Microsoft Office document or a PDF).

Future phases may extend to include other sources potentially being CHIS or the NatSecPortal but it's recognised that due to the higher threat profile associated with this information that a new and extended Privacy Impact Assessment would be required.

4. Systems and security

Describe the security and privacy features of the platform used to store information

The IBM Watson Explorer search engine currently operates in the Police network hosted in the CCL Trentham data centre. The proposed Statement of Work from IBM describing how it could be extended to index NIA narrative content proposes adding to the currently deployed infrastructure the following components.

1. A new generic web load balancer.
2. A new Watson Explorer Front End server.
3. A new Watson Explorer Engine server.

The Watson Explorer technology stack is a Java stack based on commonly used Java technology components (including Spring and IBM WebSphere). This is a common technology stack used across the world in critical applications.

As a commonly deployed technology it is also subject to constant analysis by information security researchers and threat actors. As a result, multiple historical exploits have been documented for which IBM have responded with patches. In the last two years these have included exploits variously allowing information disclosure or denial of service across both the embedded Log4J library (CVE-2021-44832, CVE-2021-45046, CVE-2021-45105 all notified in Dec 2021), and Spring Framework (CVE-2022-22970, CVE-2022-22971, CVE-2022-22968 all notified in Jan 2022).

Controls covering these known vulnerabilities include patching, restricting network access via firewalls and proxy servers, and limiting access to authenticated and authorised users.

The Police ICT team manage the underlying infrastructure and operating system directly and outsource management of the Watson Explorer search platform to IBM.

The existing deployment of Watson Explorer to IMT Search is Certified and Accredited and the project expects to meet Certification and Accreditation requirements for the indexing of NIA.

5. Identifying individuals

There are no changes to the way individuals are identified.

Privacy Risk Analysis Record any new or elevated privacy risks arising from the project. Consult with relevant subject matter experts as needed.		
Risk description What could happen and how will this affect individuals or Police	***Pre-existing Controls*** What pre-existing controls are in place to reduce the impact and/or likelihood of the risk (briefly describe any existing practice, policy, or technical control)	***Recommended controls*** To reduce risk to an acceptable level
Purpose - collection, use, disclosure or sharing		
Improved search could enable easier access to sensitive personal information for those wanting to access it without valid reason.	• Cyber training. • Terms of access and use. • Authenticated authorised access based off groups representing relevant intelligence analysts (those involved in assessing frontline risk). • NIA activity monitoring with alerts on sensitive records (VIPs and cover identities).	• Existing controls and • Search X activity monitoring with log data sent to SIEM for execution of rules against sensitive record access.
Transparency, participation, and social expectations		
No change to existing risks identified in this category due to the implementation of Search X.		
Fairness and quality		

No change to existing risks identified in this category due to the implementation of Search X.		
Systems and security		
Information security risk due to exploitation of technical vulnerabilities.	• Patching. • SIEM monitoring. • Reduce the information surface area: only expose to Watson Explorer that information for which a valid search requirement exists.	• Continuation of existing controls.
Identifying individuals		
By improving search efficiency and effectiveness it may be possible for a malicious actor to uncover individual identities faster.	• NIA monitoring (in SIEM) of sensitive individuals (VIPs/cover identities).	• Detailed monitoring sent to SIEM to uncover inappropriate activity.

Summary of findings and recommendations

The extension of IBM Watson Explorer from indexing IMT to include indexing NIA narrative content enables more efficient and effective access to intelligence embedded in NIA. This also raises the possibility that either human error or a malicious actor may disclose PII. In response the following controls have been proposed.

1. Reduce the information surface area – only expose information that has a valid requirement to be searched.
2. Ensure users are aware of safe access processes through cyber training and clear terms of use.
3. Authenticated authorised access control. Only those that should be seeing search query results can see it.
4. Monitoring of all activity to support investigation after a disclosure, or more usefully, execution of sensitive information monitoring rules. Sensitive information monitoring rules already exist within the Police Cyber team Security Information Event Monitoring (SIEM) system. By providing the detailed monitoring data to the Cyber team these rules can be enabled to identify when users view narratives connected to VIPs or Police undercover operatives.

It's noted that Investigations and Intelligence staff have begun moving some information content with sensitive aspects from NIA narratives and attachments into IMT cases. This enables them to manage access control groups to the sensitive information. IBM Watson Explorer is currently the underlying technology used to find content in IMT (commonly known as IMT Search or IMS).

Finally, ICT have proposed making the existing Crystal Drive indexing (already on Watson Explorer) available to intelligence and investigations staff. This allows people with access to documents on Crystal Drive to see relevant search results for those documents in Watson Explorer. If a user doesn't have access, then they see nothing. This provides some efficiencies over the existing File and Folder search already available under Windows and is noted here for information.

PIA Sign off

Project Manager

Privacy Officer

Senior or Executive Leader

Appendix A: Guiding principles

Purpose - collection, use, disclosure or sharing

- Any collection, use or sharing of information is for a reasonable and well-defined purpose aligned to Police strategic priorities and values, or those of the agency we share with.
- Consider whether use of identifiable personal information is necessary, and only collect, use, or share the minimum information needed to achieve the outcome.
- Risks and potential negative outcomes should be assessed so it's clear how these balance against positive outcomes.

Transparency, participation, and social expectations

- Tell people in a way that makes sense to them what information is collected about them, how it will be used, who it will be shared with and why, even if it will be used or shared in a way that does not and cannot be used to identify them.
- Even when there is no legal requirement to tell people, being open about what we do with information is important to maintain public trust and respect in Police.
- Consult with a diverse and representative group of community stakeholders to reflect the different perspectives, values, concerns and needs of individuals and communities. For example, when considering use of new technologies or processes such as artificial intelligence and automated decision making.

Fairness and quality

- Collect and use information in ways that are proportionate in scope and scale, respectful of individuals and communities and does not unreasonably intrude on the privacy interests of individuals or groups.
- Ensure the accuracy and quality of information is fit for purpose for its intended use or to share with others.
- Information is a strategic asset and maintaining its accuracy through its lifecycle is a key component in delivering policing services.

Systems and security

- Apply information management practices that align with security standards. Consider the sensitivity of the information being collected, used, shared, and reported.
- Store and maintain information so that it is accessible and reliable, and only keep it for as long as it is necessary and relevant.
- If a privacy breach occurs, address it quickly, and be accountable for it.

Identifying individuals

- Unique identifiers must only be assigned or used where it is necessary for Police to carry out its functions efficiently.
- Use of biometric technologies for identity-related purposes must be carefully evaluated to ensure that it is the most appropriate solution (and there is no suitable equivalent alternative to meet the business need).