



Workforce Data Project 2022-23

Data Collection

Privacy Impact Assessment Report

Contents

1. Project summary	3
2. Scope of the PIA.....	4
3. Personal information	6
4. Privacy assessment	8
5. Risk Assessment.....	17
6. Recommendations and action plan to minimise impact on privacy	27
Appendix 1: Data sets collected	28
Appendix 2: Collection frequency	31
Appendix 3: Information Management Protocols.....	33
Appendix 4: Related Documentation	35
Appendix 5: Cabinet Documentation	36

1. Project summary

1.1 Background

Te Kawa Mataaho Public Service Commission (the Commission) collects, in the context of the Public Service Act 2020 and a Cabinet directive (see appendix 5), a range of identifiable administrative payroll employee-level data on an annual basis from Public Service departments. This collection, known as Workforce Data, has operated in its current form since 2000.

This consolidated collection seeks to provide the same information benefits that would be possible with a centralised Public Service payroll, which was the case in New Zealand before 1987, and is the case in some overseas jurisdictions.

The data is collected to assist the Commissioner and the Public Service Leadership Team to “develop a highly capable workforce that reflects the diversity of the society it serves and to ensure fair and equitable employment”¹; “promote integrity, accountability, and transparency throughout agencies in the State services”²; and “establish and lead a Public Service leadership team so that Public Service agencies work as a system to deliver better services to, and achieve better outcomes for, the public”³.

Historically the data was collected using Treasury’s CFISNet tool. In 2022 the Commission began a project to migrate the data collection to a different technical platform controlled by the Commission. This technical change was subject to its own PIA. This tool is scheduled to be used for live data collection from July 2023.

1.3 PIA Purpose

While the technical change for data collection has a PIA, the ongoing collection itself never went through a PIA process due to it having started before PIAs were commonplace. Due to the type of data being collected, a review against the requirements of the Privacy Act 2020 is appropriate.

The purpose of this PIA is to:

1. review the privacy impacts of the ongoing data collection against current privacy law and expectations in the Public Service
2. review any impact on the annual collection becoming quarterly⁴.

1.4 Related PIAs

As context it should be noted that a separate PIA was developed in July 2022 regarding the technical change noted above. This focused on the changes to the technology used to collect the

¹ [Section 44\(c\), Public Service Act 2020](#)

² Section 44 (b), Public Service Act 2020

³ Section 44 (a), Public Service Act 2020

⁴ Note some fields will be optional in three of the four quarters. See appendix 2.

data and an intention to collect the data on a “more frequent basis”. Since that document was written “more frequent” has been clarified through consultation and engagement to mean a quarterly collection.

1.5 Current privacy management practices

The Commission has processes and practices in place to ensure privacy is maintained. These include:

- Having a privacy officer role who supports the business upholding our privacy requirements.
- Having an information security officer role to ensure that our information systems and processes meet industry standards.
- Completing an Annual Privacy Self-Assessment Assurance Report to ensure the continued maintenance (and where appropriate improvement) of our privacy practices throughout the organisation.
- Maintaining a risk register to ensure any potential issues are identified and mitigated.
- Having an Information Strategy which supports decision making when managing our strategic information.
- Having information management protocols and resources in place. A summary of this can be found in appendix 3.

2. Scope of the PIA

2.1 Scope

In Scope:

- The personal information that is collected from agencies in the Workforce Data collection
- The reasons for the collection
- How the data is used
- How the data is stored and protected
- Retention and disposal policies
- The impact of moving to quarterly rather than annual collection

Out of Scope:

- The technology used to collect the data (this is covered in a separate PIA)
- Any new collections and uses of data that were not in place for the 2022 and previous collections, as none are currently proposed other than in initial exploration phase. If they proceed, they will be subject to their own PIA or this PIA updated.

2.2 Process

This PIA was conducted as a desktop exercise based primarily on documentation for the process and discussions with Subject Matter Experts in the Strategic Information Team who own the process and data. This document was written by the Strategic Information team with consultation on privacy aspects with the legal team.

Given the collection involves multiple government agencies, several of these agencies⁵ were also consulted for their feedback as this PIA was developed.

2.3 Historic assessment

Aggregated data has been collected since 1988 (see appendix 5). In 1999 an occasional paper⁶ was written to outline the need for more detailed data and as a result the collection moved to individual level. This was accompanied by the “Background paper on Access and Security Protocols for HR capability survey”⁷. It outlines the statutory background for collecting the data, based on the State Sector Act. It referred to “statistical and research purposes” from the Privacy Act but did not do a detailed privacy review.

2.4 Rationale

This scope and process was chosen due to the desire to review the current “as-is” process against current privacy expectations.

⁵ Department of Conservation, Land Information NZ, Oranga Tamariki

⁶ <https://www.publicservice.govt.nz/assets/DirectoryFile/Archived-Occasional-Paper-Measuring-Human-Resource-Capability-in-the-Public-Service.pdf>

⁷ [Background paper on Access and Security protocols for HR Capability survey.DOC](https://sscncz.sharepoint.com/:w/s/sscdms/65098/EQFZXv8I72FNpRuViZnVuaMBTxLMFULRLVp9vAvVDD9UJQ?e=dLVmp0)
(<https://sscncz.sharepoint.com/:w/s/sscdms/65098/EQFZXv8I72FNpRuViZnVuaMBTxLMFULRLVp9vAvVDD9UJQ?e=dLVmp0>)

3. Personal information

The purpose of the collection is to provide information to the Commissioner, the Public Service Leadership Team, ministers, and the public about the composition of the workforce. This helps to ensure that there is the right mix of knowledge and skills to meet government priorities, provide trusted and responsive services, and deliver the best outcomes possible. It also helps officials to understand the extent that the Public Service workforce's diversity reflects the communities we serve.

This purpose is directly aligned with the Commission's statutory functions, noted in section 1.1. and with the purpose of the Public Service Leadership Team under the Public Service Act.

The Workforce Data Inter Agency Protocol⁸ was developed to outline how the collection operates. This includes the information collected, how it will be used by the Commission and how the Commission will manage and secure the information.

3.1 Data Collected

The collection relies on IPP2(2)(g) from the Privacy Act 2020 to allow the Commission to collect the data from agencies, rather than individuals, and IPP11(1)(h) to allow agencies to disclose that information to the Commission.

A full list of the information currently being collected from agencies is outlined in appendix 1. Personal information includes date of birth, salary, gender, ethnicity and job title.

Some fields will not be required every quarter. Appendix 2 notes which fields are collected annually, and which ones will be collected quarterly.

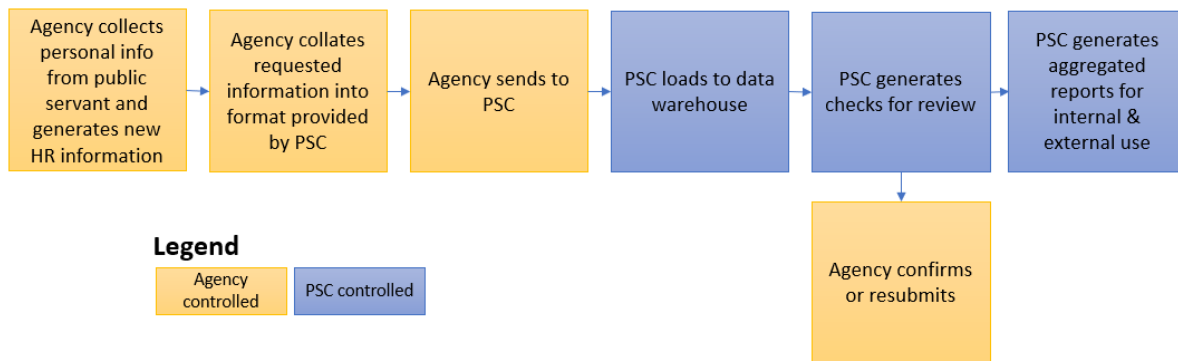
The data collected is a combination of personal demographic attributes (e.g. gender and date of birth) and information related to employment (e.g. salary, full time equivalent and annual leave taken).

A staff reference (Record ID) is collected. This is used for the purposes of communicating with the agency about rows containing data validation issues and for identifying line management. Other than these two uses, it is not used as a unique identifier by the Commission.

There are no individual fields that will consistently identify all individuals – for example a driver's licence ID or full name. However, in a large number of instances, a combination of fields (e.g. title and agency name) could be used to identify individuals with a high degree of accuracy.

⁸ https://sscnz.sharepoint.com/:b:/s/sscdms/92504/EQm9yVjBZ0NIjioAVWKxNxQB2n2zG-aOQKw06t_OIO4Vlw?e=YScGtN

3.2 Information flow – collection of information



The data is collected from agencies rather than individuals for several reasons:

- 1) Role based information (e.g. title, salary, business unit, pay band) is mastered by agencies and individuals are likely to be less accurate in their knowledge of this information.
- 2) Collecting the information directly from individuals would be a substantial burden on individuals and the Commission. It is unlikely it would be possible to get comprehensive information for all affected employees.
- 3) Agencies already collect the relevant personal information for their own purposes and their aggregated reporting should align with the Commission's.

3.3 Data uses and disclosures

As noted in the Workforce Data Inter Agency Protocol, Te Kawa Mataaho will use and disclose the Workforce Data for statistical or research purposes in connection with the Public Service Commissioner's statutory functions.

It will not publish the Workforce Data in a form that could reasonably be expected to identify the individual concerned.

Te Kawa Mataaho will confirm key summary metrics with each Public Service agency before analysis is published to Ministers or the public.

Te Kawa Mataaho may:

- provide the Workforce Data to Statistics New Zealand for inclusion in the Integrated Data Infrastructure; and
- allow researchers external to Te Kawa Mataaho to access the Workforce Data unit record (employee level) data provided that:
 - Te Kawa Mataaho is satisfied that the external researcher has the necessary research experience, knowledge, and skills to access and use the Workforce Data for the purposes of undertaking research;
 - Te Kawa Mataaho has approved the research proposal;

- potentially identifying fields (such as date of birth and job title) are excluded so the data becomes de-identified before being provided to the external researcher⁹;
- the external researcher agrees not to publish research in a form that could reasonably be expected to identify the individual concerned; and
- the data is only accessed onsite in the secure data environment at Te Kawa Mataaho and is not downloaded or removed from that environment.

3.4 Changes / Analysis

This PIA does not propose or analyse any changes to the data collected or its uses. It does analyse a planned shift to quarterly collection and reporting. Any further collection or use that departs from current activity will be the subject of a further PIA if deemed necessary.

4. Privacy assessment

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
1	<p>Principle 1 – Purpose of the collection of personal information</p> <p>Only collect personal information if you really need it</p>	<p>Information is collected for the purpose of informing decision-makers (Commissioner, PSLT, agencies and Ministers), and making aggregated data public, on the composition of the Public Sector workforce and to ensure the Public Service is reflective of the communities it serves.</p> <p>This information is essential to government transparency and to inform workforce policy. This enables to the Commissioner to work with PSLT to fulfil statutory responsibilities as outlined in the Public Service Act.</p> <p>More detailed analysis on the purpose of the collection can be found in appendix 1.</p>	Meets

⁹ As defined here: <https://www.stats.govt.nz/assets/Uploads/Integrated-data-infrastructure/de-identified-data-supporting-analytical-insights-while-maintaining-privacy-and-confidentiality.pdf> At the Commission this involves abstracting Job Title

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
2	<p>Principle 2 – Source of personal information</p> <p>Get it directly from the people concerned wherever possible</p>	<p>The Commission’s access to agencies’ data is supported by the exceptions in principle 2 that the information:</p> <ul style="list-style-type: none"> • “will not be used in a form in which the individual concerned is identified” • “will be used for statistical or research purposes and will not be published in a form that could reasonably be expected to identify the individual concerned” <p>Workforce data is sourced from each agency’s HR and payroll systems as they maintain the employment relationship.</p> <p>Some personal information was originally supplied by the individual to their agency (date of birth, ethnicity, gender) and other information is mastered by the payroll system (salary, job title).</p> <p>Gathering all data directly from individuals is not practical. This is because it is unlikely there would be a high enough level of compliance to make the data fit for purpose.</p> <p>Further, the data is collected by agencies for their own reporting purposes and there is a general expectation that the Commission’s reporting should align with agencies’ reporting.</p> <p>In alignment with the IPP2(2)(g) exceptions, published data is aggregated with suppression protocols used to ensure individuals cannot be identified</p>	<p>Meets, on the basis of reliance on the exception in IPP2(2)(g)(i) & (ii)</p>

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
3	<p>Principle 3 – Collection of information from subject</p> <p>Tell them what information you are collecting, what you’re going to do with it, whether it’s voluntary, and the consequences if they don’t provide it.</p>	<p>This data is provided by the Agency upon request from the Commission under the “statistical and research purposes” exception in the Privacy Principles.</p> <p>Visibility of the collection is maintained via the Commission’s website.</p>	<p>Not applicable as we are not collecting information from the individuals. However, visibility to individuals could be improved, see R-04 and A-01.</p>
4	<p>Principle 4 – Manner of collection of personal information</p> <p>Be fair and not overly intrusive in how you collect the information</p>	<p>Data is collected from agencies to avoid a burden on employees.</p> <p>Data is collected for analysis relating to the Commission’s statutory purposes, as outlined in the background section.</p> <p>The shift to a quarterly collection is required to provide more timely information on the size and makeup on the Public Service workshop.</p>	<p>Meets with some risks around visibility to be mitigated (R-04)</p>

5	<p>Principle 5 – Storage and security of personal information</p> <p>Take care of it once you’ve got it and protect it against loss, unauthorised access, use, modification or disclosure and other misuse.</p>	<p>Technology</p> <p>The data is stored in tools (both historically and with the recently changed technology) that have completed security certification and accreditation.</p> <p>Access to data is controlled via standard IT controls (e.g. username and password with appropriate permissions).</p> <p>The new tools also introduce audit functionality that automatically record which users access reports and data.</p> <p>Additional information on technical tools can be found in the July 2022 PIA¹⁰ that focuses on the technology.</p> <p>Access - staff</p> <p>Permissions are set at the role level and based on group membership.</p> <p>Currently four users have write access to the database. These users are all in Strategic Information. All updates are audited.</p> <p>Approximately 12 users have read access to identifiable row level (employee level) data. Most are analysts in Strategic Information. In addition, there is one analyst each from HR, Employment Relations and Diversity & Inclusion.</p> <p>In order to access raw row level workforce data, internal Commission employees are required to review the Access, Security and Privacy Protocols for the Public Service Workforce Data and sign the confidentiality agreement form. Users need to justify to the SIT manager why this level of detail is required.</p> <p>All other users can only interact with aggregated (non-identifiable) data.</p>	<p>Meets (with some opportunities to improve, see A-04)</p>
---	--	--	---

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
		<p>More generally, during the staff hiring process, the Commission undertakes reference and criminal checks.</p> <p>Access – non staff</p> <p>Approved researchers must agree to the same protocols as Commission staff and sign the confidentiality form. In addition they have to satisfy the Commission of having appropriate experience and an approved research proposal. They will only receive de-identified data (e.g. date of birth and job title removed) and unit-level data will not be able to be taken offsite.</p> <p>De-identified data means information that cannot be identified just using the information by itself, however it may be re-identifiable if the user has additional contextual information¹¹. Re-identification is prohibited in the protocols the researcher agrees to.</p> <p>The Manager Strategic Information, the Deputy Commissioner Strategy and Policy and the Public Service Commissioner are the only people who can grant access to the de-identified unit record data, for both staff and non-staff.</p> <p>Additional detail can be found in the document Workforce Data Access and Security protocols May2021.docx</p>	

¹⁰ [PIA as data collector.docx](#) <https://sscnz.sharepoint.com/:w/s/sscdms/65098/ETViHj-gcdhFrhiflIWc8SwB4f6ei5p2zMaiyYMAw0FcyA?e=sQ5gam>

¹¹ <https://www.stats.govt.nz/assets/Uploads/Integrated-data-infrastructure/de-identified-data-supporting-analytical-insights-while-maintaining-privacy-and-confidentiality.pdf>

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
6	<p>Principle 6 – Access to personal information</p> <p>People can see their personal information if they want to</p>	<p>As the workforce data is mastered by supplying agencies and only provided to Te Kawa Mataaho for statistical and research purposes it is likely queries about the data would be directed to the source agency. As such, specific protocols have not been put in place to support individual queries. However, if such a query did occur it and the specific individual was able to be identified (e.g. using the Record ID), then it would be managed through our standard privacy policy and practices and the data provided to the individual.</p>	Meets
7	<p>Principle 7 – Correction of personal information</p> <p>They can correct it if it's wrong, or have a statement of correction attached</p>	<p>As the workforce data is collected is for statistical and research purposes and it is mastered by supplying agencies then the Commission would recommend any corrections would be directed to the source agency.</p> <p>If the individual prefers that our records are updated too, the Commission will note that the individual has provided corrected information in accordance with our privacy policy.</p> <p>If aggregated statistics for the period have already been published, the Commission is unlikely to reprocess the published statistics.</p>	Meets

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
8	<p>Principle 8 – Accuracy etc. of personal information to be checked before use</p> <p>Make sure personal information is correct, relevant and up to date before you use it</p>	<p>The key accuracy metric for the Commission is aligning with agencies’ own reporting. To achieve this, after the data is validated, summary information is sent back to the owning agency for review before the results are published.</p> <p>The Commission expects agencies have a robust process in place to ensure the information they collect on their staff is accurate and appropriate from a privacy perspective, in line with their statutory responsibilities.</p> <p>Each iteration of the collection is recorded separately and when results are published the date the data is collected as-at is noted. This separation will continue as the collection is expanded to quarterly collections.</p>	Meets
9	<p>Principle 9 – Not to keep personal information for longer than necessary</p> <p>Get rid of it once you’re done with it</p>	<p>Business processes adhere to the disposal and retention policies and rules set out in our Information and Records Management policy and the Disposal Schedule (8.7 and 8.8).</p> <p>Transfer files (“returns”) are securely disposed of within three months.</p> <p>Confirmed data sets will be retained for five years after their last use. In practice this could mean the data sets are retained for a significant period of time, due to their potential research value in addressing policy questions.</p> <p>This retention is subject to security, confidentiality and statutory obligations (from protocol 3, see appendix 4). This policy is consistent with the Public Records Act.</p>	Meets with some risks to be examined (R-13 & A-04)

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
10	<p>Principle 10 – Limits on use of personal information</p> <p>Use it for the purpose you collected it for, unless one of the exceptions applies</p>	<p>Previously collected and future collections of workforce data will only be used for the purpose it was originally collected for, i.e. for statistical analysis and research on the Public Service workforce in order to monitor the workforce and provide informed policy advice.</p> <p>Any currently unanticipated future uses would rely on IPP(10)(1)(b)(ii) and/or would go through its own Privacy Impact Assessment.</p>	Meets
11	<p>Principle 11 – Limits on disclosure of personal information</p> <p>Only disclose it if you’ve got a good reason, unless one of the exceptions applies</p>	<p>The Commission will only disclose de-identified data for a “statistical or research purposes” exceptions (such as approved external researchers). The process for doing so is outlined in principle 5.</p>	Meets
12	<p>Principle 12 – Disclosure of personal information outside New Zealand</p>	<p>Personal information may be stored on offshore third-party systems but where this occurs it is subject to appropriate contractual controls to prevent the third party vendors accessing or using this data. The privacy commissioner has advised that such an arrangement is not treated as a disclosure under IPP12¹².</p> <p>This is also noted in the PIA on the technology change¹³.</p>	Meets

¹² <https://www.privacy.org.nz/publications/statements-media-releases/new-principle-for-disclosing-personal-information-overseas/>

¹³ [PIA as data collector.docx](https://sscncz.sharepoint.com/:w:/s/sscdms/65098/ETViHj-gcdhFrhiflIWc8SwB4f6ei5p2zMaiyYMAw0FcyA?e=NmihAL) (https://sscncz.sharepoint.com/:w:/s/sscdms/65098/ETViHj-gcdhFrhiflIWc8SwB4f6ei5p2zMaiyYMAw0FcyA?e=NmihAL)

#	Description of the privacy principle	Summary of personal information involved, use and process to manage	Assessment of compliance
13	<p>Principle 13 – Unique identifiers</p> <p>Only assign unique identifiers where permitted</p>	<p>Unique identifiers are not allocated to individuals.</p> <p>Agencies provide their unique identifiers for the purpose of communicating with them about that row and for identifying management relationships.</p>	Meets

5. Risk Assessment

Ref. no.	Purpose of collecting the information	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-01	Reason for collection	Stakeholders such as staff (both in their personal and professional capacity) are not clear on clear lawful purpose for collection	The collection has been going on in the background for a long time and some stakeholders may not be aware why.	New workforce inter agency protocols and the recent PIAs outline the section 44 PSA 2020 reason for collection	Low	Ask agencies to check their contracts and/or staff policies to ensure staff are advised of the collection, including its purpose. The Commission to provide common messaging to support this.	Low
R-02	Scope of collection	Scope of collection may collect unnecessary information if not carefully considered.		All fields carefully considered internally with clear use cases which demonstrate why collection is necessary.	Medium	Publish information on website analysis outlining each field collected and why.	Low

Ref. no.	Source of personal information	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-03	Data source	Personal information may be less accurate because it is sourced from the employing agency instead of direct from the individual.	Direct collection from individuals is impractical. In addition collecting from agencies mean the Commission statistics will align with theirs	<p>Agencies collect this data directly from the individual. They have high interest in accurate data for their own reporting purposes, including meeting statutory requirements and Commission expectations.</p> <p>To check accuracy, the Commission provides aggregated responses back to the agency for their review (twice) and confirmation (once).</p>	Low		Low

Ref. no.	Telling the individual what you're doing	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-04	Individual knowledge	Public servants may not be aware of the collection.	Direct advice to all public servants would require coordination across 37 agencies and may not make a material difference to knowledge levels	Website explains collection and purpose for it. Results are publicized via media each year.	Low	Ask agencies to check and update their contracts and/or staff policies to ensure staff are advised of the collection.	Low
R-05		Proposed legislation change will require that individuals are advised of when their data will be passed on to third parties.		As above		Ask agencies to check and update their contracts and/or staff policies to ensure staff are advised of the collection.	

Ref. no.	How you are collecting personal information	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-06	Frequency	Some may see the frequency moving to quarterly is over collecting	Key stakeholders including the Commissioner and Ministers have requested more frequent collection as they are unable to make timely decisions with the current availability.	Clear use cases/rationale	Low		Low

Ref. no.	How you are storing and securing personal information	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-07	Data storage	Data repository is hacked	Data / privacy breach	Integrated login with the Commission's Active Directory New solution completed penetration testing and risk analysis/mitigation. Followed NZISM	Low		Low
R-08	Unauthorised or untrained access	Users without proper training or without clear purpose are able to access the data or use it for unauthorised purposes	Data is used for unauthorised purposes	IT access to systems managed centrally by data custodians. Documented process for new access exists including declaration from user on adhering to authorized uses only. Documented process on acceptable use, including for researchers (Workforce Data_Access and Security protocols).	Low		Low
R-09	Unauthorised modification of data	Users without proper training or without clear purpose are able to edit the data	Data is edited for unauthorised purposes	Only necessary system administrators have write access and all actions are logged in the audit system.	Low		Low

Ref. no.	Responding to people's requests for information about themselves, or requests to correct information about themselves	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-10	Access to information	Individual wishes to view/update information about themselves but no clear processes exist	Data is used at aggregated level for stats and research purposes so having detailed processes and capacity for individual level data is not generally necessary	Users in the first instance would be referred back to the agency as the data source. If the individual still wishes us to access our data, organisation wide policies for managing people's requests for personal data are in place and the data will be made available.	Low		Low
R-11	Correcting information	Users information is incorrect and they wish to update it, but changes may take us out of alignment with agency reporting	Alignment with agency published data is a key aspect. Data is not used at individual level so limited value to the	Capability to note a correction will fulfil obligation. Also advise individual to update the source of information for next collection.	Low		Low

		(depending on timing)	individual in correcting data.	Where appropriate, the Commission would also work with agency on case by case basis to ensure correction and reporting are aligned.			
Ref. no.	What steps do you take to check the accuracy, relevance etc of personal information before you use it?	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-12	Accuracy	Data does not align to the source agency's view.	Commission reports differ from the agency's reports and/or incorrect information is published.	Summary reports are sent to agency for sign off before data use	Low		Low

Ref. no.	How long do you keep personal information and why?	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-13	Retention	Data is retained for longer than necessary	<p>Because of the statistical value of the information, the data is likely to be retained indefinitely at unit record level.</p> <p>From an information management perspective, information and records must be identifiable, retrievable, accessible and usable for as long as they are required (from Archives' Information and records Management Standard).</p>	Most users only have access to aggregated data.	Medium	Review if specific personal fields such as DOB and job title can be removed from older data sets and replaced with more generic values such as month or year of birth and ANZSCO or standardised job titles .	Low

Ref. no.	Who are you going to disclose the personal information to (if anyone) and why?	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-14	Sharing with IDI	Stats NZ may request the data to include in the IDI	Stats collects a large range of personal information in the IDI and has specific legislation and processes to enable this. Risk is that this may not be followed.	<p>Policies are in place at both Stats NZ and the Commission to manage this with business processes to ensure key staff are aware of them.</p> <p>Agencies have been advised of this potential use via the Protocol</p>	Low		Low
R-15	Sharing with researchers	Researchers may gain access to and use the data for inappropriate purposes	Researchers may have bona fide uses for the data so it should be available to them but this could be misused if not closely managed	<p>Research use case must be approved by the Commission before granting access.</p> <p>Researcher must show they have suitable experience to use the information in a safe way</p> <p>Researcher will have to read and agree to various controls in using the data.</p> <p>Researcher will only receive de-identified</p>	Low		Low

				data and cannot take it offsite. Aggregated results will be reviewed by the Commission before it being removed			
Ref. no.	Who are you going to disclose the personal information to offshore (if anyone) and why?	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-16	N/A	No overseas disclosure expected. Overseas infrastructure is used, this is covered in PP5.					
Ref. no.	Why do you need a unique identifier, and are you allowed to use this one?	Description of the risk	Rationale and consequences for the agency or individual	Existing controls that contribute to manage risks identified	Assessment of residual current risk	Recommended mitigations or privacy enhancements	Residual risk remaining despite new safeguards
R-17	N/A	No unique identifier is used.					

6. Recommendations and action plan to minimise impact on privacy

Ref	Agreed action	Who is responsible	Completion Date
A-01	Ask agencies to check and update their contracts and/or staff policies to ensure staff are advised of the collection. Provide common messaging to support agencies in doing so.	Strategic Information Team	
A-02	Publish on the website a table showing exactly what data is collected and reason for its collection	Strategic Information Team	
A-03	Review whether the manager to employee relationship field can be removed or made optional.	Strategic Information Team	
A-04	<p>Review if specific personal information can be removed in some scenarios:</p> <ul style="list-style-type: none"> • for detailed internal analysis use cases, can de-identified data be used instead of identifiable data. This may reduce the number of analysts who have access to identifiable data • for long term retention of data can the raw data have personal information replaced with more generic info (e.g. DOB or title) 	Strategic Information Team	

Appendix 1: Data sets collected

Table 1 contains the fields being captured from agencies for each employee and provided to the Commission.

“Public analysis” notes whether the field is visible in some way to the public. For example Gender is used in the publicly available gender pay gap analysis. Some fields are not used in analysis available to the public, for example pay bands, but are used internally by the Commission.

Table 1:

Field	Description	Purpose
Record ID	A code for each employee as assigned by the organisation.	Identifying rows with data validation issues and defining management hierarchies. Public analysis: No
Gender	The employee’s gender	Demographic analysis (including gender pay gap) Public analysis: Yes
Ethnicity 1	The employee’s ethnicity	Demographic analysis Public analysis: Yes
Ethnicity 2	The employee’s 2nd ethnicity – if specified	Demographic analysis Public analysis: Yes
Ethnicity 3	The employee’s 3rd ethnicity – if specified	Demographic analysis Public analysis: Yes
Date of Birth	The employee’s date of birth	Demographic analysis (used to calculate age) Public analysis: Yes (aggregated age)
Occupation	The employee’s current job, as classified according to the Australian and New Zealand Standard Classification of Occupations (ANZSCO)	Public analysis by occupation group, internal analysis at more detailed levels. Public analysis: Yes (aggregated)
Job Title	Current job title for the position held	Provide advice on workforce capability and how it is changing over time. Public analysis: No
Business Unit	Up to three Business Unit fields can be provided to allow for classification at several different levels in the organisation’s hierarchy, e.g. division, branch and team. This should be the highest in the hierarchy.	Provide advice on workforce capability and how it changing over time. Public analysis: No

Business Unit 2	Next level in the organisation hierarchy below Business Unit	Provide advice on workforce capability and how it is changing over time. Public analysis: Yes
Business Unit 3	Next level in the organisation hierarchy below Business Unit 2	Provide advice on workforce capability and how it is changing over time. Public analysis: Yes
Status	The employee's current status within the organisation	Calculate workforce size (only active staff are counted) and turnover Public analysis: Yes
Salary	Annual base salary	Calculate average salaries Public analysis: Yes
FTE	Proportion of full-time hours worked	Calculate count of FTEs and determine full-time/part-time status. Public analysis: Yes
Contract Term	The term of employment: open term (permanent), or fixed term (temporary)	Calculate total and percentage for permanent/fixed term conditions. Public analysis: Yes
Start Date	The date the employee started in the organisation	Calculate new starts and tenure. Public analysis: Yes
End Date	Last day of duty for departing employees	Calculate terminations and turnover. Public analysis: Yes
Termination Reason	The reason category why an employment relationship ended for those employees who have	Calculate unplanned turnover and count redundancies Public analysis: Yes (redundancies and unplanned turnover only)
Region	The region where an employee's workplace is located based on regional council boundaries	Location analysis Public analysis: Yes
Management level	Employees in the top three tiers of management	For analysis at senior management level Public analysis: Yes
Sick and domestic leave TOTAL	The total number of sick and domestic leave days taken due to sickness or domestic leave related reasons, relates to part-time and full-time permanent staff that are eligible (usually those that	Calculate average sick and domestic leave taken Public analysis: Yes

	have been employed for more than six months)	
Domestic leave ONLY	The total number of days of absence due to domestic leave relates to part-time and full-time permanent staff that are eligible for domestic leave (usually those that have been employed for more than six months)	Calculate average domestic leave taken (e.g., for gender analysis) Public analysis: No
Pay Band Name	The name used to describe the pay band internally	Contributes to understanding of pay gaps (e.g., allows comparison of salary to pay gaps between different groups). Public analysis: No
Pay Band Minimum	The minimum dollar figure of the pay band	Contributes to understanding of pay gaps (e.g., allows comparison of salary to pay gaps between different groups). Public analysis: No
Pay Band Maximum	The maximum dollar figure of the pay band	Contributes to understanding of pay gaps (e.g., allows comparison of salary to pay gaps between different groups). Public analysis: No
Pay Band Midpoint	The midpoint dollar figure of the pay band	Contributes to understanding of pay gaps (e.g., allows comparison of salary to pay gaps between different groups). Public analysis: No
Manager ID	The RecordID of the employee's manager.	Allows org structures to be understood. Public analysis: No
Super Organisation Contribution	The percentage of salary that the employer is contributing to superannuation	Allows analysis of non-salary remuneration. Public analysis: No
Performance Pay	The dollar value of performance pay received during the data collection period	Allows analysis of non-salary remuneration. Public analysis: Yes
Annual Leave Entitlement	An employee's annual leave entitlement in days	Allows analysis of non-salary remuneration. Public analysis: No

Appendix 2: Collection frequency

All fields are required on an annual basis. Some fields are optional for quarterly submissions. As noted in the Workforce Data Inter Agency Protocols, these fields may be subject to change based on information needs from stakeholders. Any changes will be socialised with agencies, who can feedback via appropriate channels.

	Required	
Field name	Annually	Quarterly
RecordID (Employee ID)	Yes	Yes
Gender	Yes	Yes
Ethnicity1	Yes	Yes
Ethnicity2	Yes	Yes
Ethnicity3	Yes	Yes
DateOfBirth	Yes	Yes
ANZSCO ¹	Yes	Optional
JobTitle	Yes	Yes
BusinessUnit	Yes	Optional
BusinessUnit2	Yes	Optional
BusinessUnit3	Yes	Optional
Status	Yes	Yes
FTESalary	Yes	Yes
FullTimeEquivalent	Yes	Yes
ContractTerm	Yes	Yes
StartDate	Yes	Yes
EndDate	Yes	Yes
TerminationReason	Yes	Yes
Region	Yes	Yes
ManagementProfile	Yes	Yes
SDLeaveTaken (sick leave)	Yes	Optional
DomLeaveTaken (domestic leave)	Yes	Optional

PayBandName	Yes	Optional
PayBandMin	Yes	Optional
PayBandMax	Yes	Optional
PayBandMidpoint	Yes	Optional
ManagerID	Yes	Optional
SuperOrgContrib	Yes	Optional
PerformancePay	Yes	Optional
AnnualLeaveEntitlement	Yes	Optional

Appendix 3: Information Management Protocols

This table is from the Commission’s Information Management protocols_updated.pdf document (https://sscnz.sharepoint.com/:b:/s/sscdms/92504/Ed4aau7OL5tGhPt7c5E2W_cBfO2DUNyEDWR0zVVjw-vdLg?e=AhLmO3). Additional information on each of these protocols can be found in that document.

Protocols	Practices	Resources
We embed our information roles and responsibilities	<p>Information is an important asset with appropriate governance structures</p> <p>Information sources each have a steward responsible for life cycle management</p> <p>All staff know their responsibilities with information</p> <p>These responsibilities also apply to external researchers given access to the Commission data</p>	Information roles and responsibilities
We secure the privacy and confidentiality of our information	<p>Legal and ethical obligations around managing information are adhered to</p> <p>The permissions we have to use data are understood</p> <p>Security practices are built into our processes and infrastructure</p> <p>Privacy or security breaches are managed openly and quickly, recognising the seriousness of maintaining confidentiality</p>	<p>Privacy Policy</p> <p>Confidentiality Guidelines</p>
We invest wisely in our information resource	<p>All data that is collected has a clear use and its value is understood</p> <p>Information is actively used, and then archived, to get full value from it</p> <p>Data is appropriately and efficiently sourced</p> <p>Information infrastructure is invested in to enhance the value of our data</p>	<p>Information register</p> <p>Collection guidelines</p>
We assure our information quality	<p>There is a culture of professionalism and good practice</p> <p>Information meets the needs of users, within available resources •Information is accurate</p> <p>Information is timely enough to be of value to users</p> <p>Information is consistent</p> <p>Methods used to produce information are understood and documented</p>	

<p>We share an information language</p>	<p>Common information standards are used to manage the Commission information</p> <p>The Commission uses national and international information standards where possible</p> <p>The Commission promotes common information standards across the system</p>	<p>Information register</p>
<p>We use our information wisely</p>	<p>Information infrastructure assists with turning data into insight</p> <p>Processes get the right information to the right people at the right time</p> <p>Published information is presented clearly and supported by analysis</p> <p>Published information is open and accessible</p> <p>Published information is understandable</p> <p>Significant errors in published information are corrected quickly</p>	<p>Release guidelines</p> <p>Declaration on open and transparent government</p>

Appendix 4: Related Documentation

- (1) Access, Security and Privacy Protocols for Human Resources Capability Information Project – Background Paper:
https://sscnz.sharepoint.com/:w:/s/sscdms/65098/EQFZXv8I72FNpRuViZnVuaMBTxLMFU_LRLVp9vAvVDD9UJQ?e=dLVmp0&CID=1042E836-5E1C-4B77-B673-F11FB52DB0E9&wdLOR=cC5B3AA2B-4E65-418B-8167-0961CCE278C5
- (2) Occasional Paper from 1999 that explained the need for the Workforce Data collection (known then as Human Resources Collection):
https://sscnz.sharepoint.com/:b:/s/sscdms/65098/EVtyJDJmNnxAlmzHckEtxa0Bi0_8GKdJQup1rx4oivK4gg?e=4nbOXu
- (3) Privacy Impact Assessment for technology change, July 2022:
<https://sscnz.sharepoint.com/:b:/s/sscdms/65098/ET03oVJSuSNDqQol00MXNDgBCfzTr-ggya2HpxQ8m5zABA?e=YmpCTZ>

Appendix 5: Cabinet Documentation

Below are short summaries of documentation provided to and from Cabinet and the Cabinet State Sector Committee.¹⁴

CO (88) 11 – Cabinet Office Circular 2 August 1988

Directs departments to provide staffing levels, including redeployment of staff resources within departments, to the Minister of State Services every four months. The State Services Commission will collate the reports from each department and prepare a combined report for the Minister to submit to the Cabinet Management and State Employment Committee.

STA (90) 6 – State Sector Committee 10 December 1990

The Minister of State Services recommends to Cabinet State Sector Committee that the data is also collected from a number of Crown Agencies, based on a prior request from the Committee (Cabinet Minute SAS(90)M 21/1). Examples include NZDF civilian staff, NZQA, Police civilian staff, Transit New Zealand.

Refers to the collection being quarterly rather than every four months (no reference).

STA (90) M 4/1 - State Sector Committee 12 December 1990

Minute noting the agreement of the Committee to the information sought in the paper STA (90) 6 (above).

STA (94) 38 - State Sector Committee 1 August 1994

Notes a review on the information collected by the State Services Commission has been completed (as advised via STA (93) M 25/2). Recommends the data be collected every six months instead of every quarter. Notes that additional information on age, salaries and recruitment and retention is being collected by the Commission on an annual basis.

STA (94) M 12 - State Sector Committee 3 August 1994

Minute noting the agreement to collection of statistical information on a 6 monthly basis instead of quarterly. Noted the additional information being collected annually.

STA (96) 8 – State Sector Committee 26 February 1996

Report containing information provided by Public Service departments and selected crown entities on their staffing levels as at 31 December 1995. The Committee is asked to note it.

STA (96) M 2-2 – State Sector Committee 28 February 1996

Noted the above report and directed officials to provide more information on staff turnover, skills shortages and other related matters in future six-monthly reports, in order to provide a more dynamic picture of the Public Service staffing situation.

¹⁴ Full copies of the documents referenced can be found here:

[https://sscnz.sharepoint.com/:f:/r/sites/sscdms/65098/Shared%20Documents/HRC%20\(Restricted\)/Governance%20and%20Management/Cabinet%20Directives?csf=1&web=1&e=W3teUj](https://sscnz.sharepoint.com/:f:/r/sites/sscdms/65098/Shared%20Documents/HRC%20(Restricted)/Governance%20and%20Management/Cabinet%20Directives?csf=1&web=1&e=W3teUj)