

SSA Richmond and Wandsworth Data Standards

Document Control

Title of document	SSA Richmond and Wandsworth Data Standards
Produced by	Data Quality and Standards Subcommittee
Document type	Internal and external
Subject	Data Standards
Review frequency	Ongoing
Next review due	Ongoing
Target audience	All staff

Contents

Introduction and Purpose	3
Scope.....	3
Template for new standard	4
Minimum Mandatory Standards.....	5
Date.....	6
Time	7
Name (First name, Middle name(s), Last name).....	8
First Name.....	8
Middle Name(s).....	8
Last Name	9
Address	10
UK Postal Address	10
Post code.....	10
LLPG UPRN	11
Ethnic Group or Ethnicity.....	13
E-mail Address	15
UK Phone number (Mobile, Landline).....	16
Mobile number collection.....	16
Landline number collection	16
National Insurance Number.....	17
NHS Number	18

Introduction and Purpose

During consultations for the data and analytics strategy, and Covid response, one of the key challenges highlighted was the divergence in data collection across services that reduces the organisation's ability to get the most out of our data. This situation is not unique to Richmond and Wandsworth and tackling it is a key part of our progress towards better data quality.

This set of data standards has therefore been developed to ensure that data is collected more consistently across the Councils. Standardised data collection is key to enabling data linkages for intelligence, commissioning, emergency response and development and implementation of data and digital solutions. It makes data verification and matching easier and crucially leads to more accurate decision-making.

These standards have been developed in line with the SSA's Data Quality Policy.

The standards have been split into two categories:

1. **Minimum mandatory standards** – this means that where this information is being collected, it **must** be collected in this way. Please note that it does not mean you should always collect this information for every person. Data collection should always follow UK GDPR and be minimised to what is necessary for the specified purpose as allowed under relevant legislation.
2. **Recommended standards** – where this is information is being collected, it is recommended to be collected in this way, but not mandatory.

Recommended standards will be added as they are developed.

The minimum mandatory standards are expected to be implemented across all data collection processes across Richmond and Wandsworth Councils. These should also be used as part of commissioning and procurement of applications and data solutions.

Exceptions for modified collection or maintaining status quo should be backed by clear business, contractual or legislative reasons. When not adopting the data standards, existing systems should provide clear rationales backed up by evidence e.g., of service disruption, suitability for client care, costs, changes to current system not possible etc.

Scope

These data standards should be applied to all areas of data collection. This includes existing and new systems that hold data, surveys (e.g. public consultations), new systems and solutions procured, new and existing commissioned services.

This data standards document will be maintained and reviewed frequently to ensure standards reflect work being done by Government Statistical Service and Government Standards Authority and changing needs of the SSA services and partners.

Due to business needs, sometimes more granular collection of data may be needed. For example, a commissioned service may want to gather more detailed ethnicity data. This can be done as long as it is possible to map the additional categories to the ones listed in the data standards.

Template for new standard

The following template has been adopted from the e-Government Interoperability Framework (e-GIF). This template will be used to document each data standard.

Standard Name	
Definition	
Format	
Validation	
Values	
Default Value	
Based on	
Verification	
Comments	
Version	
Approval Date	

Minimum Mandatory Standards

Minimum mandatory standards – this means that where this information is being collected, it must be collected as prescribed by the standard.

Please note that it does not mean you should always collect this information for every person. Data collection should always follow UK GDPR and be minimised to what is necessary for the specified purpose.

- Date
- Time
- Name (First name, Middle name(s), Last name)
- UK Postal Address
- Post code
- UPRN
- Ethnic group
- E-mail with validation
- Phone number - Mobile, Landline with validation
- National Insurance number with validation
- NHS Number with validation

Date

Standard Name	Date
Definition	The day, month, year and century, or any combinations of these elements, of an event
Format	Max 10 Characters in the format CCYY-MM-DD
Validation	<ol style="list-style-type: none"> 1. Values less than 10 in the day, month or year elements, should be entered with a zero in the first position 2. Days must not be greater than 30 in April, June, September and November. 3. Days must not be greater than 28 in February except when 29 is allowed for a leap year. 4. Where applicable, valid ranges should be provided e.g., no date of birth in the future or no date of birth outside the reasonable ranges.
Values	CCYY should be a valid year number MM in Range 01 – 12. DD in Range 01 - 31
Default Value	N/A
Based on	e-Government Interoperability Framework (e-GIF) ISO 8601 DATE AND TIME FORMAT https://www.iso.org/iso-8601-date-and-time-format.html XML Schema Part 2: Datatypes W3C Recommendation 02 May 2001. See http://www.w3.org/TR/xmlschema-2/
Verification	
Comments	<p>American date formats should especially be avoided as due to lack of use in U.K. they increase chances of inaccurate capture</p> <p>Consult the ISO 8601 where there is need for date and time, inclusion of weeks</p> <p>During data collection provide drop-downs of calendar and if people are allowed to enter dates, validate them</p> <p>Date collection – Service users and frontline staff will think of date starting with DD-MM-CCYY. For paper, online and other forms it is okay to collect the data this way, but it should be transformed to the CCYY-MM-DD format for storage and processing.</p>
Version	
Approval Date	

Time

Standard Name	Time
Definition	The time (using a 24 hour clock) at which an event, or the action in an event, takes place.
Format	Max 8 characters expressed as HH:MM:SS (H=hours, M=minutes, S=seconds)
Validation	<ol style="list-style-type: none"> 1. All times must be expressed in the 24 hour clock format, e.g. one minute past midnight is 00:01:00 2. Values of any element less than 10 should be entered with a zero in the first position
Values	N/A
Default Value	N/A
Based on	<p>XML Schema Part 2: Datatypes W3C Recommendation 02 May 2001. See http://www.w3.org/TR/xmlschema-2/</p> <p>ISO 8601 Conventions</p>
Verification	N/A
Comments	<p>See the W3C standard for details of formatting shortened versions of Time, e.g. hours and minutes only</p> <p>Consult the ISO 8601 where there is need for date and time, inclusion of weeks</p>
Version	
Approval Date	

Name (First name, Middle name(s), Last name)

As best practice all parts of the name (First, Middle, Last) should be collected and collected separately. Consistent name collection is vital for data linkages and cross department service delivery.

You may want to add notes such as 'As it appears on your passport' so people do not use nicknames or shortened names.

Recommended standards are being developed for collection of preferred name and title. Preferred name should not replace full name collection.

Guidance for preferred name collection can be found in recommended standards.

First Name

Standard Name	First name
Definition	First name of the person
Format	Max 70 characters
Validation	None
Values	N/A
Default Value	N/A
Based on	e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf Collecting and using data from surveys and administrative sources https://gss.civilservice.gov.uk/policy-store/full-name/
Verification	N/A
Comments	Important to collect separately from other parts of the name
Version	
Approval Date	

Middle Name(s)

Standard Name	Middle name(s)
Definition	All middle names
Format	Max 99 characters
Validation	None
Values	N/A
Default Value	N/A
Based on	e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf Collecting and using data from surveys and administrative sources https://gss.civilservice.gov.uk/policy-store/full-name/
Verification	N/A
Comments	Important to collect separately from other parts of the name
Version	
Approval Date	

Last Name

Standard Name	Last Name
Definition	That part of a person's name which is used to describe family, clan, tribal group, or marital association. Preferred collection is as last name.
Format	Max 70 characters
Validation	None
Values	N/A
Default Value	N/A
Based on	e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf Collecting and using data from surveys and administrative sources https://gss.civilservice.gov.uk/policy-store/full-name/
Verification	N/A
Comments	Important to collect separately from other parts of the name
Version	
Approval Date	

Address

Most important thing for address collection is that:

- Address should not be manually entered, especially in large systems. You should be providing an address look up feature from a list based on AddressBase. This is the best way to collection correct address information.
- All systems must store a LLPG UPRN along with the address
- Council gets free access to AddressBase and there is an API that lets address lookup features be built for online form and various systems.

UK Postal Address

Standard Name	UK Postal Street Address
Definition	The address of a postal delivery point within the UK. This is usually a building and usually comprises a name/number, street, town, and county.
Format	<ol style="list-style-type: none"> 1. Maximum of 35 characters per line 2. Minimum of 2 lines for address collection, maximum 5 lines 3. Separate line for flat or unit number (e.g. Flat 10) and road/street name and number (10 Downing Street). Title of the line to specify which element is being collected. 4. Add separate and well labelled line for 'Town or City' 5. Add separate and well labelled line for postcode (Standard below)
Validation	Use address lookup rather than manual entry. Justification should be provided when manual entry is used.
Values	N/A
Default Value	N/A
Based on	e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf BSEN 7372:1993 https://data.gov.uk/education-standards/sites/default/files/CL-Address-Line-Type-v3-0.pdf
Verification	N/A
Comments	<u>Address should not be manually entered, especially in large systems.</u> You should be providing an address look up feature from a list based on AddressBase. This is the best way to collect correct address information.
Version	
Approval Date	

Post code

Standard Name	Post Code
Definition	The code allocated by the Post Office to identify a postal delivery point.
Format	Minimum 6 and Maximum 8 characters
Validation	<ol style="list-style-type: none"> 1. The Outward code can be 2, 3 or 4 characters followed by a space and the inward code, which is 3 characters and is always NAA

	<p>2. The outward code has seven valid formats, AN, ANN, AAN, AANN, ANA, AANA, and AAA</p> <p>3. The letters I and Z are not used in the second alpha position (except GIR 0AA)</p> <p>4. The second half of the code never uses the letters C, I, K, M, O, and V</p> <p>A postcode can be validated against a table of all 1.7 million postcodes in Code-Point Open. The full delivery address including postcode can be validated against the Royal Mail Postcode Address File (PAF), which lists 29 million valid delivery addresses, constituting most (but not all) addresses in the UK.</p>
Values	N/A
Default Value	N/A
Based on	Royal Mail
	e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf
Verification	N/A
Comments	<p>The inward and outward code can be broken down into sub parts as follows:</p> <ul style="list-style-type: none"> • the outward (before the space) code consists of a postcode area and a postcode district, • the inward code consists of a postcode sector and a postcode unit. <p>Mandatory for complete address collection</p>
Version	
Approval Date	

LLPG UPRN

Standard Name	LLPG Unique Property Reference Number
Definition	Unique number associated with each property. The UPRN is a unique reference for each and every addressable location in the UK.
Format	Up to 12 characters in length
Validation	The number should not be entered manually. It should be assigned as part of the address lookup process. Once an address is picked the associated LLPG UPRN is filled in.
Values	N/A
Default Value	N/A
Based on	Geoplace and Ordnance survey Unique Property Reference Number Power of UPRN GeoPlace LLP
Verification	Assigned based on the address, not manual entry
Comments	<p>There are many different types of UPRNs. The one that must be used, (services can use additional ones for their need), is the LLPG UPRN.</p> <p>LLPG UPRN should not be manually entered. It should come in from the address lookup feature used.</p> <p>LLPG UPRN should be stored alongside each address</p>

Version	
Approval Date	

Ethnic Group or Ethnicity

Standard Name	Ethnic group
Definition	Self-reported measure of ethnic group asked as 'What is your ethnic group?'
Format	Max 90 characters
Validation	N/A
Values	<p>Full Collection – Specific Groups (18 Categories)</p> <p><u>White</u></p> <ol style="list-style-type: none"> 1. English/Welsh/Scottish/Northern Irish/British 2. Irish 3. Gypsy or Irish Traveller 4. Any other White background <p><u>Mixed/Multiple ethnic groups</u></p> <ol style="list-style-type: none"> 5. White and Black Caribbean 6. White and Black African 7. White and Asian 8. Any other Mixed/Multiple ethnic background <p><u>Asian/Asian British</u></p> <ol style="list-style-type: none"> 9. Indian 10. Pakistani 11. Bangladeshi 12. Chinese 13. Any other Asian background (excluding Arab) <p><u>Black/ African/Caribbean/Black British</u></p> <ol style="list-style-type: none"> 14. African 15. Caribbean 16. Any other Black background <p><u>Other ethnic group</u></p> <ol style="list-style-type: none"> 17. Arab 18. Any other ethnic group, please describe <p>99 Prefer not to say 999 Missing</p> <p>Broad ethnic group categories (5 Categories)</p> <ol style="list-style-type: none"> 1. White (E.g., English, Irish, British, Gypsy, Other white) 2. Black (E.g. African, Caribbean, Black British etc.) 3. Asian (Indian, Pakistani, Bangladeshi, Chinese other Asian) 4. Mixed (Mixed/multiple ethnic background such as White and Black African)

	<p>5. Other 99 Prefer not to say 999 Missing</p> <p>Top level ethnic groups (2 categories)</p> <p>1. Asian, Black and other minority ethnic groups 2. White 99 Prefer not to say 999 Missing</p>
Default Value	N/A
Based on	<p>ONS Measuring equality: A guide for the collection and classification of ethnic group, national identity and religion data in the UK. Government Statistical Service, Harmonised Concepts and Questions for Social Data Sources - Primary Principles – Ethnic Group</p> <p>Ethnicity harmonised standard https://analysisfunction.civilservice.gov.uk/policy-store/ethnicity-harmonised-standard/</p>
Verification	N/A
Comments	<ul style="list-style-type: none"> • Self-reported measure. • Race should not be asked. • NHS or other ethnic groups should be avoided as the population is reported according to ONS groups and using NHS groups makes comparisons inaccurate. • BAME should not be used as a term in collection or reporting. • Default collection and reporting should use full 18 categories. • Top level ethnic group collection should only be restricted to where needed to preserve confidentiality. Should be avoided for routine service provision or where equalities need to be measured and monitored. • In some cases, more specific collection may be needed e.g. Nigerian, Somali, more specific European group etc. In this case add a please specify subcategory under the broader category. Key thing is that groups can roll up into broader categories.
Version	
Approval Date	

E-mail Address

E-mail collection is very important for digital delivery of services and data linkages. Where collecting contact information, always collect e-mail addresses, and ask people to re-enter to ensure correct e-mail is being collected.

Standard Name	E-MAIL ADDRESS
Definition	The string of characters that are needed in order to get an item of e-mail delivered to the correct post-box
Format	As per IETF RFC2822 255 characters
Validation	From IETF RFC2822: An email address is a specific Internet identifier that contains a locally interpreted string followed by the at-sign character ("@", ASCII value 64) followed by an Internet domain. The locally interpreted string is either a quoted-string or a dot-atom. Comments and enfolding white space SHOULD NOT be used around the "@" in the email address. (i.e. no space characters either side of the @ character). The domain portion identifies the point to which the mail is delivered. The local-part portion is a domain dependent string. In addresses, it is simply interpreted on the particular host as a name of a particular mailbox.
Values	N/A
Default Value	N/A
Based on	As per IETF RFC2822 e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf
Verification	N/A
Comments	Especially for online forms, ask clients to enter it twice to ensure correct entry
Version	
Approval Date	

UK Phone number (Mobile, Landline)

International phone number collection not covered at the moment. To enable provision of digital services, always collect mobile phone numbers for clients.

Mobile number collection

Standard Name	Mobile telephone number
Definition	Individual's mobile number (07xxx xxx xxx)
Format	Max 11 characters
Validation	No spaces. No special characters or +. Begins with 07. Google's libphonenumber library can validate telephone numbers from most countries.
Values	N/A
Default Value	N/A
Based on	e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf (OFTEL guidance)
Verification	N/A
Comments	In the field to collect the phone number it should be clear if mobile or landline Ask clients to enter the number twice where the mobile phone is needed for service provision.
Version	
Approval Date	

Landline number collection

Standard Name	UK Telephone Number
Definition	A number, including any exchange or location code, at which a person or organisation can be contacted in the UK by telephonic means.
Format	Max 12 characters
Validation	No spaces. No special characters or + Google's libphonenumber library can validate telephone numbers from most countries.
Values	N/A
Default Value	N/A
Based on	e-GIF - Government Data Standards Catalogue Vol 2 http://xml.coverpages.org/govtalkCat2.pdf (OFTEL guidance)
Verification	N/A
Comments	In the field to collect the phone number it should be clear if mobile or landline
Version	
Approval Date	

National Insurance Number

Standard Name	National Insurance Number (NINO)
Definition	The National Insurance Number is a reference number that is issued to a PERSON by the Department for Work and Pensions (DWP) / HM Revenue and Customs (HMRC) for participants in the National Insurance Scheme.
Format	Max 9 characters
Validation	<ul style="list-style-type: none"> • Must be 9 characters • First 2 characters must be alpha • Next 6 characters must be numeric • Final character can be A, B, C, D • First character must not be D, F, I, Q, U or V • Second characters must not be D, F, I, O, Q, U or V • The letter O is not used as the second letter of a prefix. • Prefixes BG, GB, KN, NK, NT, TN and ZZ are not to be used
Values	Must be 9 characters
Default Value	N/A
Based on	HMRC https://www.gov.uk/hmrc-internal-manuals/national-insurance-manual/nim39110
Verification	None
Comments	<p>Validation is key for fraud prevention, data linkages etc.</p> <p>Ask clients to double enter the NI number.</p>
Version	
Approval Date	

NHS Number

Standard Name	NHS Number
Definition	Unique number used by NHS and often social care in identifying people receiving health and care services, and in locating and communicating their health and care records and other information pertaining to the planning and provision of their care. Also used by Children's services in some places as children lack a NINO.
Format	10 numeric
Validation	No spaces Current system for format is 10 digit number as xxx xxx xxxx. The final digit is error detecting checksum. Checksum validation should be built in electronic collection (details at the link provided) NHS NUMBER (datadictionary.nhs.uk)
Values	N/A
Default Value	N/A
Based on	NHS Data dictionary NHS NUMBER (datadictionary.nhs.uk)
Verification	Use validation steps above
Comments	
Version	
Approval Date	