



The Challenges Local Government Have in Person Name Identification

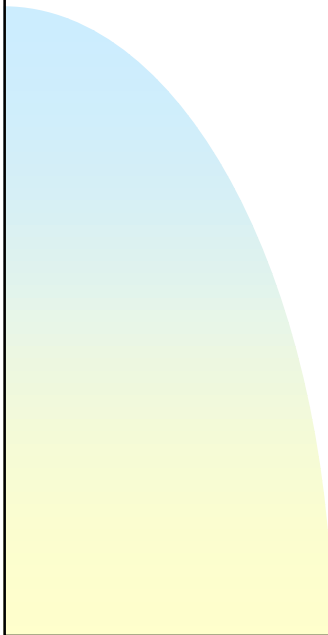
Presented by John Rymell, Core Data Administrator
- People and Organisations

London Borough of Tower Hamlets



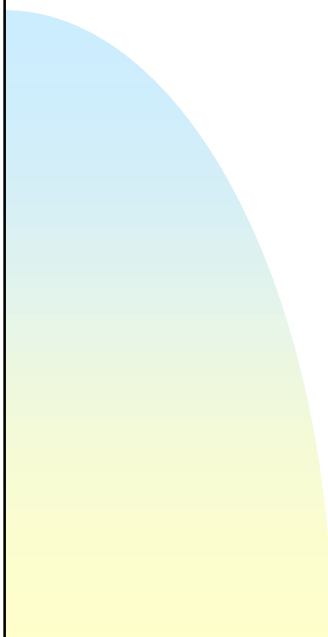
The next half-hour...

- Background History
- Tower Hamlets
- Core Data
- Challenges of Person Name Identification
- Current Solutions
- Future Issues
- Conclusions



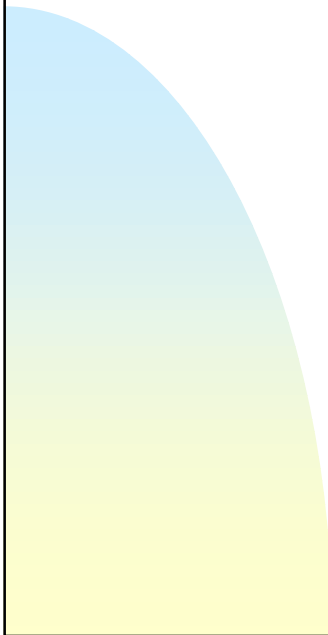
Background History

- Paper-driven systems
- Change to computer-driven systems
- Practicality, trust, cost
- Getting customers to part with information about themselves



What are we doing at Tower Hamlets?

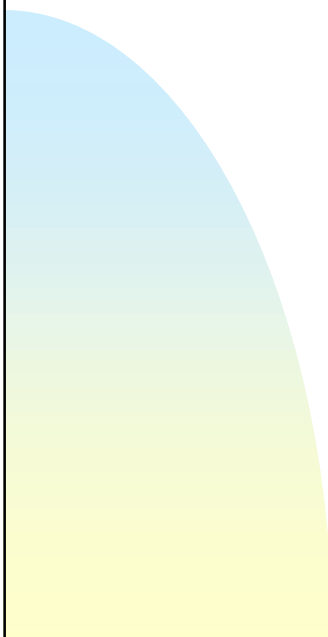
- eGovernment
 - ◆ Local Government Online by 2005
 - ◆ Joined-up Government
- Core Data Project



LASER

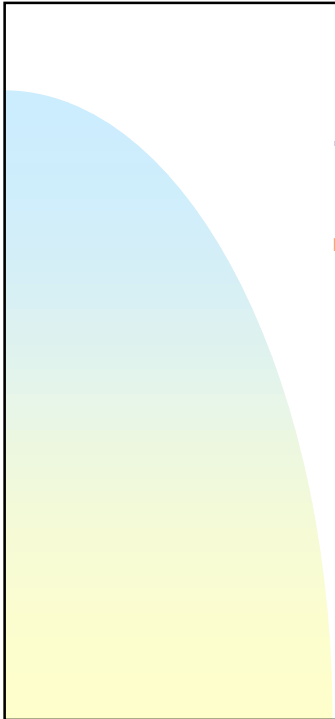
Local Authorities' Secure Electoral Register

- National access to quality controlled electoral register information
- Nationally linked - locally maintained
- Local knowledge
- Address information standard - BS7666 - Master Address Database / LPG
- Developing a standard for the non-address information - Person Names - BS8766
- Matching and maintenance



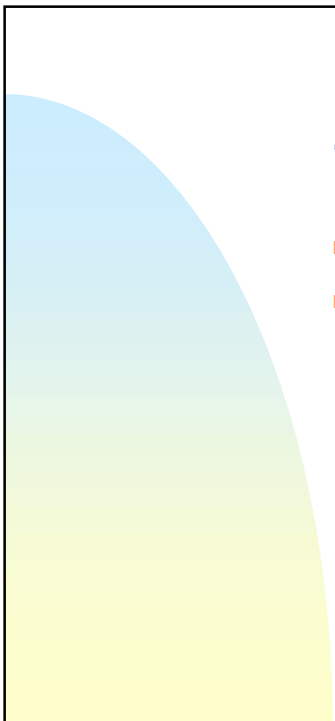
What is Core Data?

- Centralised shared data
- Common data only
 - ◆ Names
 - ◆ Addresses



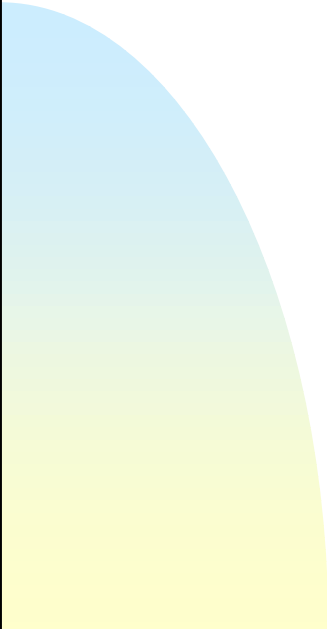
The Core Databases

- Land & Property Gazetteer (LPG)



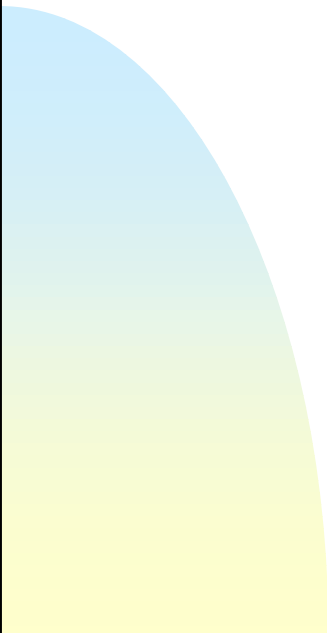
The Core Databases

- Land & Property Gazetteer (LPG)
- Core Organisations Database (COD)

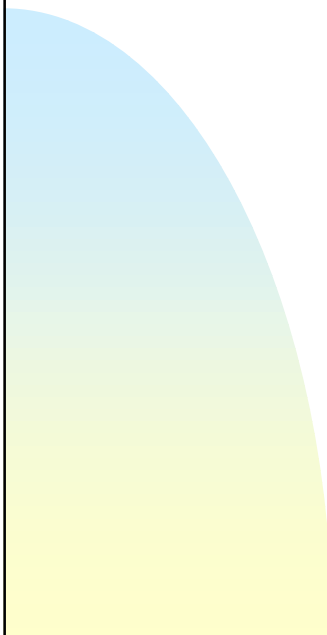


The Core Databases

- Land & Property Gazetteer (LPG)
- Core Organisations Database (COD)
- Core People Database (CPD)

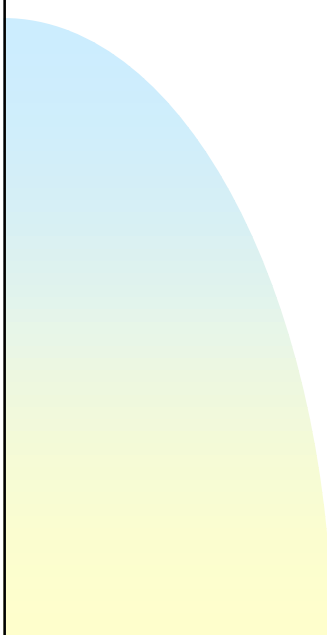


Challenges of Person Name Identification



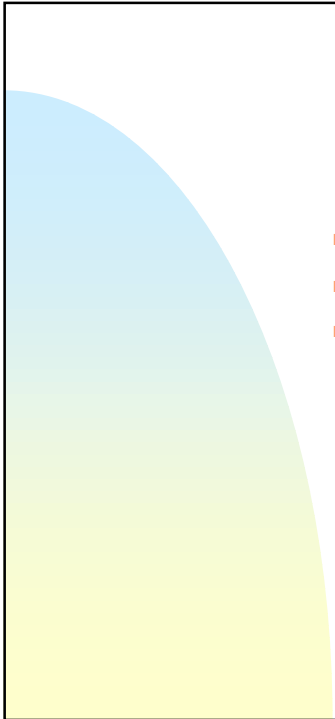
Challenges of Person Name Identification

- Data Collection
- Data Sources
- Completeness of Data
- Referencing



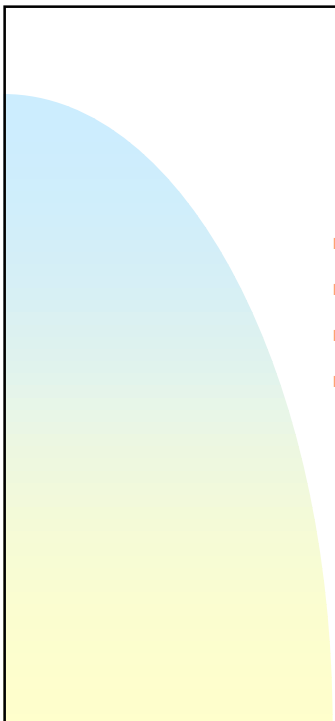
Data Collection

- Parting with information - suspicion
- Spelling
- Pronunciation
- Hearing
- Understanding
- Third-hand information
- Assumptions
- Laziness
- Lack of knowledge
- General attitudes



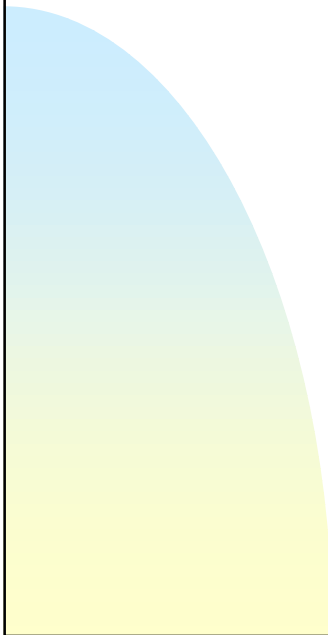
Data Sources

- Inability to verify data first hand
- Trusting received data second hand
- Matching new source data to CPD



Completeness of Data

- Full names and initials
- Abbreviations
- International name formats
- Different scripts



Referencing

- Identical names
- Date of Birth
- References
- Duplication
- Errors



Solutions



Solutions

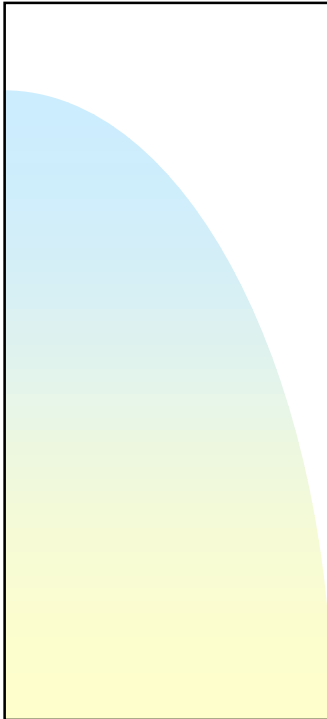
- Training and Design
- Data Sources
- Data Management Steps
- Standards



Training and Design

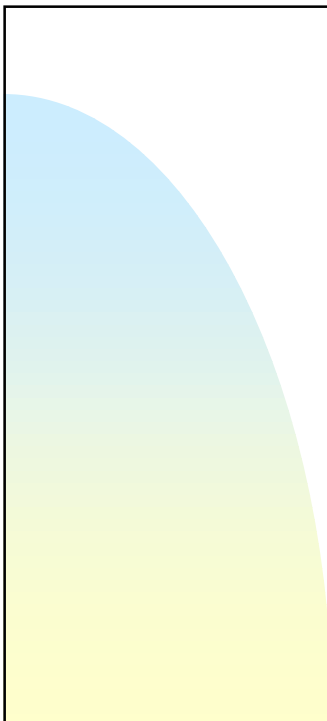
- Change of culture
- Understanding of new needs
- Authentication responsibilities
- Greater precision and accuracy
- Asking the right questions
- Searching methods and reasons

- Standard paper and electronic forms
- Simple Plain English forms
- Sufficient space for responses



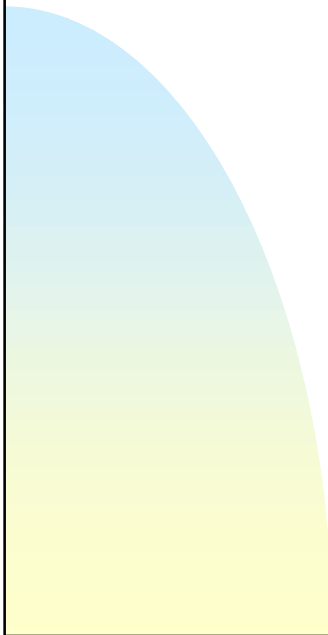
Data Sources

- Primary Data
- Secondary Data
- Matching
- Responsibility



Data Management Steps

- Requirements
- Structure
- Form Design
- Collection
- Completed Form Checking
- Input
- Input Matching
- Input Confirmation



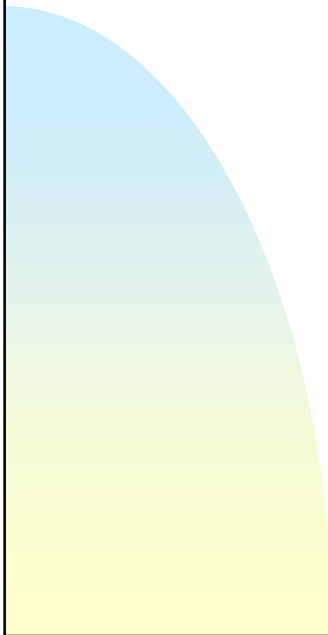
Data Management Steps

- Requirements
- Structure
- Form Design
- Collection
- Completed Form Checking
- Input
- Input Matching
- Input Confirmation
- Searching
- Retrieval
- Retrieval Matching
- Retrieval Confirmation
- Maintenance Prompts
- Maintenance
- Correction



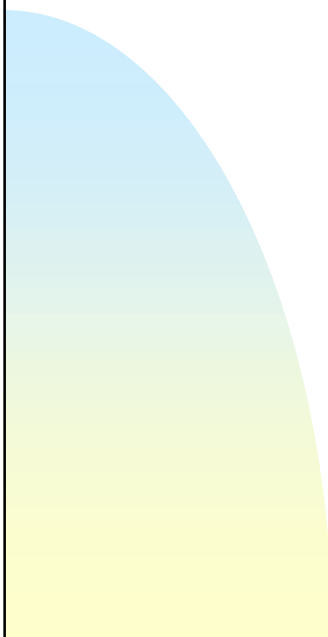
How we choose to store the data

- Available technology
- Application software
- MS Access - flexibility
 - ◆ Development
- Proprietary Database - power
 - ◆ Council-wide use

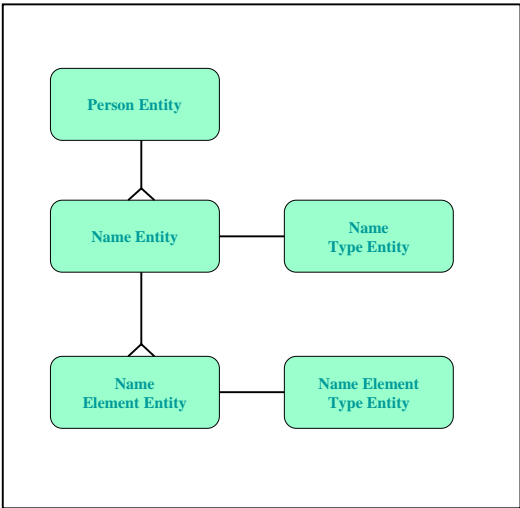


Standards

- BS7666 Data-Sets for Geographic Referencing
- Towards BS8766 Person Names
- Towards a Person Referencing standard



Towards BS8766... Person Names data model



```
classDiagram
    class PersonEntity[Person Entity]
    class NameEntity[Name Entity]
    class NameElementEntity[Name Element Entity]
    class NameTypeEntity[Name Type Entity]
    class NameElementTypeEntity[Name Element Type Entity]
    PersonEntity --|> NameEntity
    NameEntity --|> NameElementEntity
    NameEntity -- NameTypeEntity
    NameElementEntity -- NameElementTypeEntity
```

Person Names

■ Person Entity

- ◆ The identifier for each person in the local data repository
- ◆ e.g. Local Key 1234567
- ◆ It is a local identifier
- ◆ The person may have a different identifier in a different local repository
- ◆ It is NOT an national ID number

Person Names

■ Name Entity

A specific and ordered combination of words by which a person may be known

- e.g.
- 1 - Mr Abdullah Ahmed Muhammed Iqbal
1 - Registered Name
 - 2 - Mr Iqbal Muhammed Abdullah Ahmed
5 - Electoral Registration Name
 - 3 - Mr Ahmed Abdullah Iqbal
2 - Person Requested Name
 - 4 - Mr Iqbul
6 - Other

Person Names

■ Name Element Entity

Each word in a person's name assigned element position and types

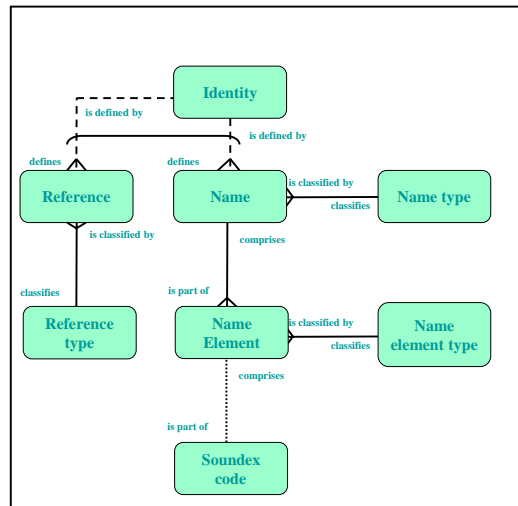
e.g.	Element	Position	Type
	Mr	1	2 - Title
	Iqbal	2	3 - Forename
	Muhammed	3	3 - Forename
	Abdullah	4	3 - Forename
	Ahmed	5	4 - Family Name

Person Referencing

■ John Smith (LocalKey 123456)

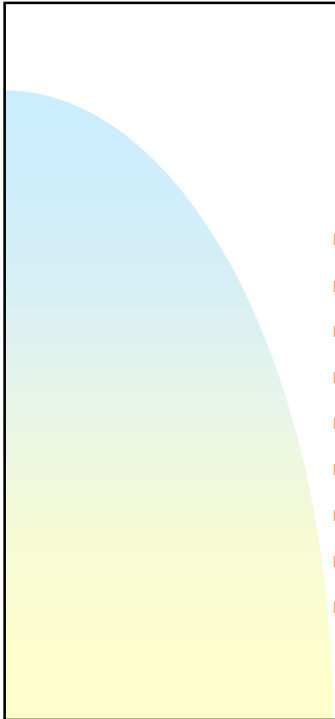
■ John Smith (LocalKey 234567)

Basic model for Person Referencing



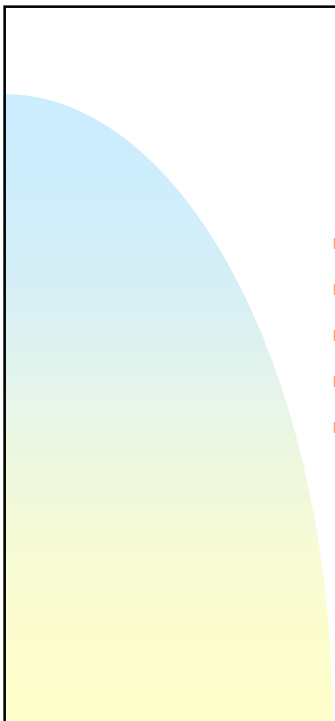
Structure of People's Names and Matching

- Soundex®
- First 4 characters
- Alias tables
- Punctuation
- Standardising & Stripping



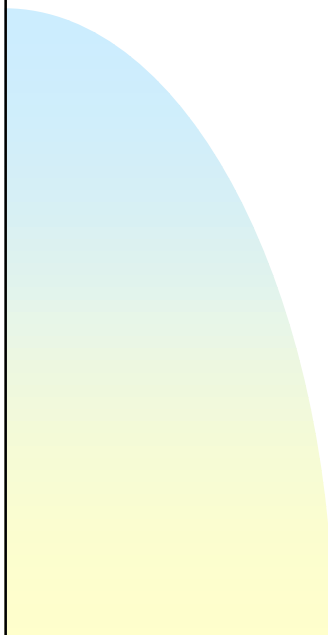
Soundex®

■ A E I O U Y H W	not coded
■ B F P V	coded as 1
■ C G J K Q S X Z	coded as 2
■ D T	coded as 3
■ L	coded as 4
■ M N	coded as 5
■ R	coded as 6
■ John or Johnnie	J500
■ Rymell or Ranleigh	R540



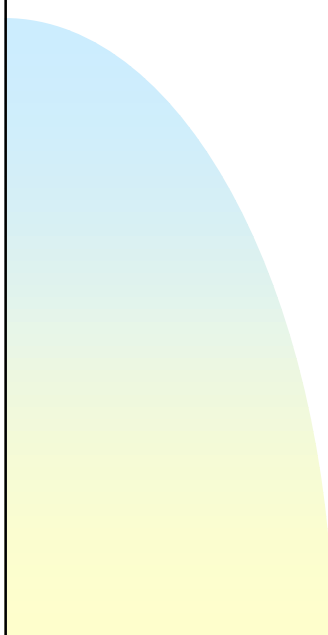
First 4 characters

■ Margaret	Marget
■ Johan	Johannes
■ Bertrand	Bertand
■ Jeffreys	Jeffries
■ Bart	Bartholomew



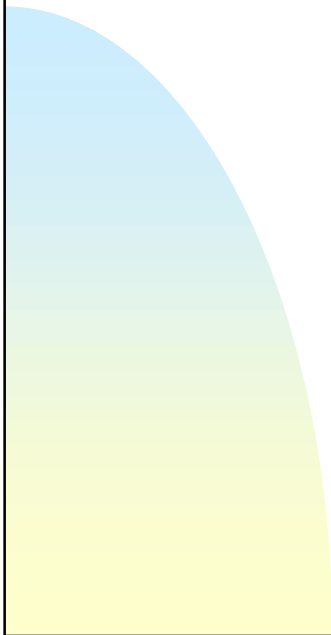
Alias tables

Name	Alias
■ Margaret	Meg Maggie Margret
■ John	Jonathan Jonathon Johnnie Johnny
■ Featherstonehaugh	Fanshaw



Alias tables

Name	Alias
■ Margaret	Meg Maggie Margret
■ John	Jonathan Jonathon Johnnie Johnny
■ Featherstonehaugh	Fanshaw
■ Cholmondley	Chumley



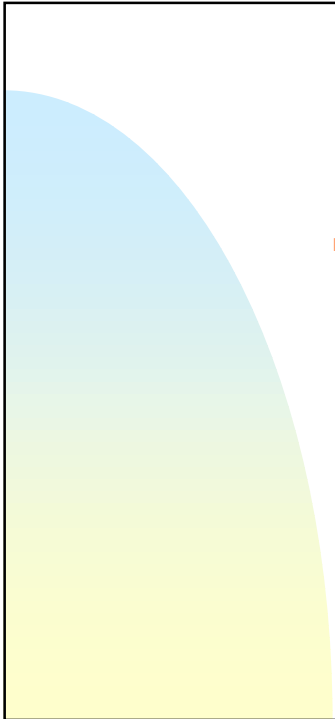
Punctuation & Abbreviations

- St. John
- Mohd
- Md
- O'Neill
- McArthur
- MacArthur
- Dixon-Symes
- Dixon Symes



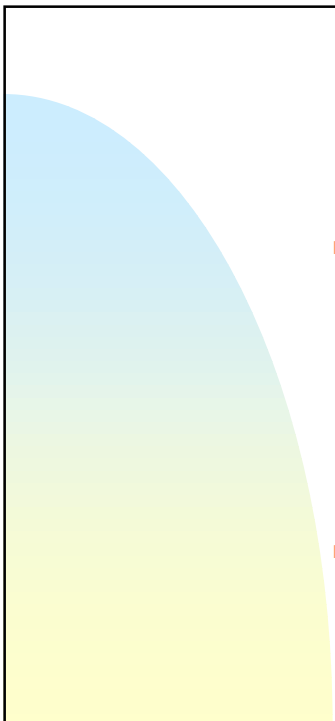
Standardising & Stripping

From	To
■ St. John	STJOHN
■ Dixon-Symes	DIXONSYMES
■ Étienne	ETIENNE
■ Graça	GRACA
■ Ahmed Al-Fayed	AHMEDALFAYED



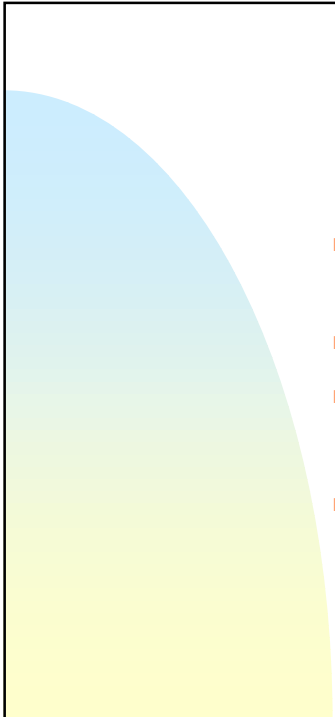
Future issues

- Biometrics
 - ◆ Photographs
 - ◆ Fingerprints
 - ◆ Retinal Scan
 - ◆ DNA
 - ◆ Voice recognition



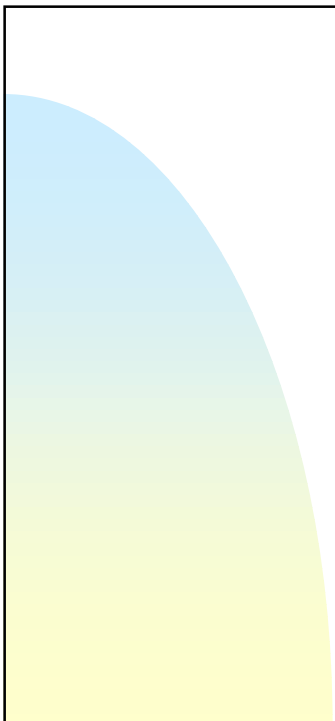
Future issues

- Biometrics
 - ◆ Photographs
 - ◆ Fingerprints
 - ◆ Retinal Scan
 - ◆ DNA
 - ◆ Voice recognition
- Electronic
 - ◆ Digital Identity



Conclusions

- Issues with Person Name Identification
- Importance of Data Collection
- Maintaining data
- **Secure, meaningful and correct**



Questions?



Contact details

John Rymell
Core Data Administrator
London Borough of Tower Hamlets
Mulberry Place
5 Clove Crescent
London
E14 2BG

john.rymell@towerhamlets.gov.uk
020 7364 3150
0701 744 7039