



Intellectual
Property
Office


UK IPO Applicant Identifiers

Julie Daltrey, Data Architect & Gareth Jones, Informatics Analyst


Applicant Name Standardization Workshop, Geneva 3rd May 2019




History of Applicant Identifiers at the UK IPO

- Mid 1980s, the UK IPO developed mainframe systems for Patents, Trade Marks & Designs.
 - Each system had its own Name & Address database.
 - Each system used Automated Data Processing (ADP) numbers to uniquely identify each party associated with an IP Right.
- 

Patent ADP numbers

- Every patent/application had to have at least 1 person or company associated with it.
 - Each person/company was allocated an ADP number and assigned 1 or more roles in relation to the patent e.g. a person may be an applicant, inventor and also represent themselves at the Office, such a person would have 1 ADP number and 3 roles associated with their patent.
 - The ADP number was split into 3 parts – the ADP Name Number (7 digits), a check digit and an additional 3 digits for each address, 001, 002 etc.
- 


Trade Marks

- Trade Marks used ADP numbers until the mainframe system was replaced in 2013.
 - The ADP number structure was similar to Patents but did not include the check digit i.e. there was a 7 digit name number and a 3 digit address suffix.
 - On migration from the mainframe system, new owner identifiers were assigned starting from 1, and incremented by 1, for each new identifier.
- 

Registered Designs

- Designs was migrated from its mainframe system in 2016 and the use of ADP numbers was discontinued.
- The migrated data still has the 5 digit name number but with no address element.
- However, from the start of the new Designs system business identifiers were discontinued with the result that for every repeat filing by an applicant, the name and address data is re-entered and no business identifier allocated.

Patent - ADP Name Number

- The “ADP Name Number” consisted of 8 digits, which was intended to uniquely identify the name to which it applied.
 - The first 7 digits are a serial number with the 1st digit identifying whether the name and address originated from GB patent data or EPO data.
 - The 8th digit was a check digit using the modulus 10 algorithm which is a simple checksum formula used to validate that the number had been correctly entered.
- 

Patent - ADP Number

- Each ADP Name Number can have any number of addresses associated with it; this is to allow, for example:
 - A person changing their address – this will result with the 3 digit address number being incremented from 001 to 002 and the original 001 address being “closed”.
 - A company operating from a number of offices with each one using the same ADP number but with a different address suffix – 001, 002, 003 etc.


Patent - Name Data Entry

- Names are entered utilising delimiters to identify the “significant part” of the name to create an index of names in the correct format for easier searching.
- For a person, the significant part is considered to be the surname (family name). These are entered as:
 - Surname^Forename^Prefix^Suffix – the 1st delimiter indicates the text to the left is the surname, the 2nd indicates the forename &/or initials and the 3rd that the text to the right is a suffix such as jnr, snr etc.
- For a company (or other legal entity), 2 delimiters are used and these are placed either side of the significant part of the name – this is the unique element and excludes company descriptors such as Ltd, PLC, GmbH etc.
 - {Marks & Clerk{
 - A A Thornton{& Co


Failed Concept?

- No, 35+ years later the UK IPO's name & address data is not in a good state but that is not because the ADP number concept failed, it is a victim of circumstances!


How have we ended up here?

- Initial data entry – agency staff paid to enter the name & address data by the “ADP Number”, the more they created, the more they earned.
 - Strict desk notes requiring the selection of an existing identifier only if it was an exact match to the information written on the form in front of them.
 - EPO name and address data is mixed with UK data but is supplied as an entity rather than as people and organisations.
- 

Issues

- The “new” patents system still retains all of the constraints of the mainframe system; we could even “run out” of ADP numbers by 2022.
 - The data supplied by customers does not match the fields in the database so data gets entered “somewhere” – result is that the real customer can become unclear and misleading by the time the data is displayed in Esp@cenet.
 - Electronic Filing – 16 years after going live with the EPO’s eOLF system, data is still typed into the backend system.
 - Search capability in the mainframe system is limited by today’s standards – it is often easier to create a new customer than to find the correct one.
- 

Issues cont.

- De-duplication projects mark duplicate ADP numbers as “not to be used” but this does not prevent these ADP numbers from being selected for a new case.
 - A lack of appreciation for the use of this data beyond the IP Rights granting areas.
- 

Data linking at the UK IPO

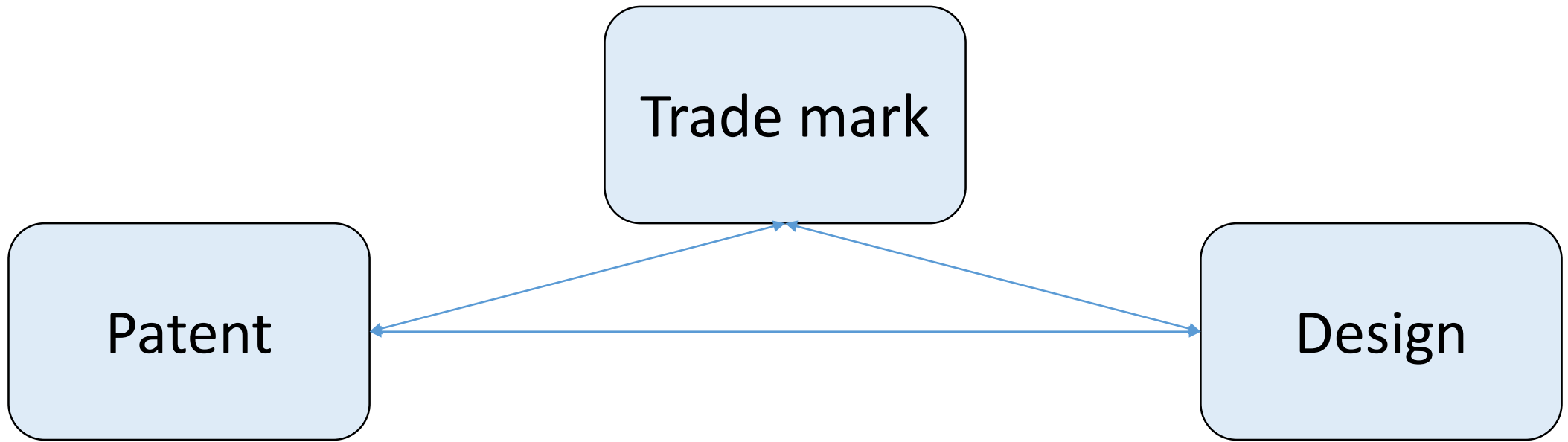
Gareth Jones

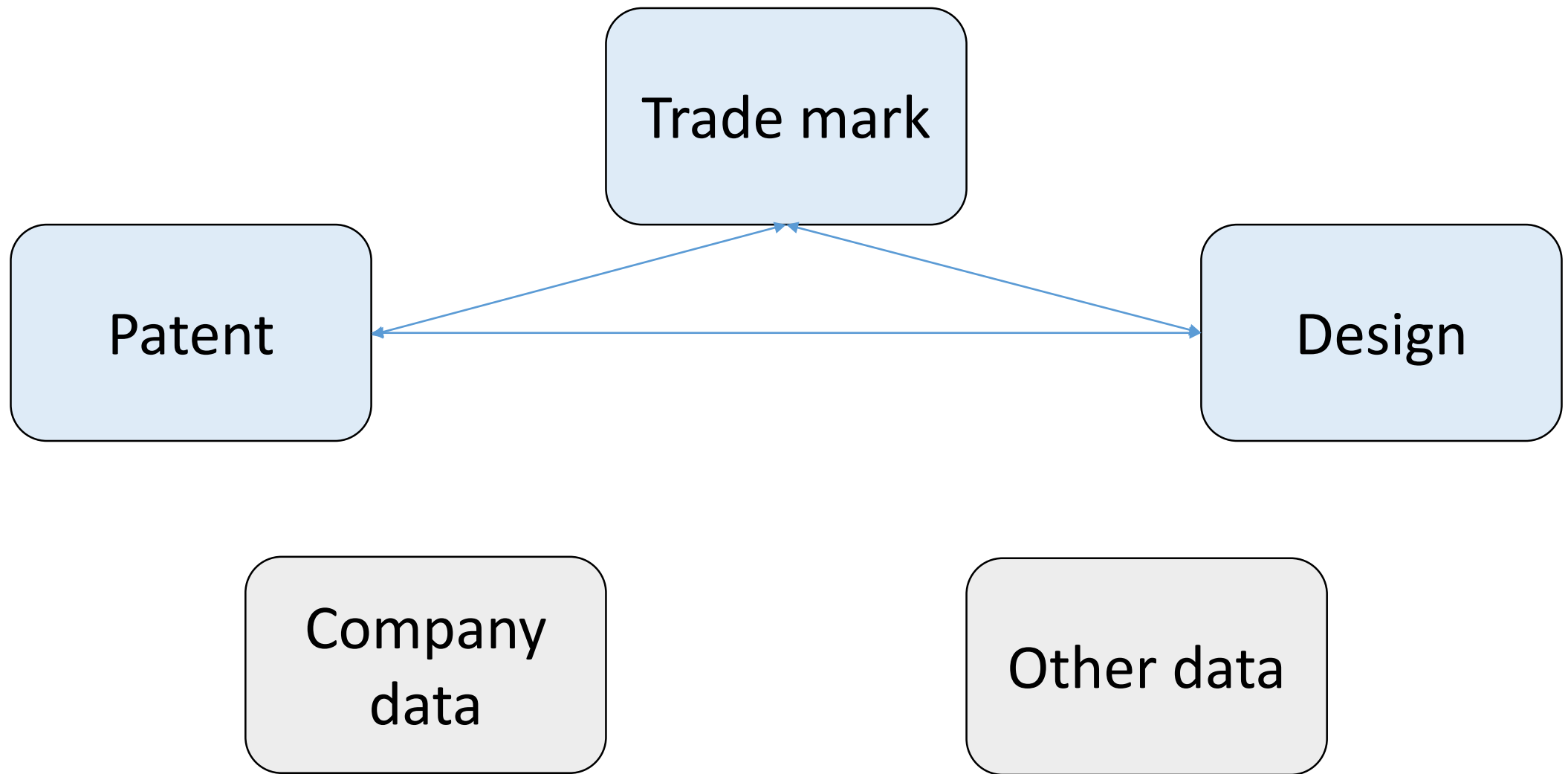
Informatics analyst

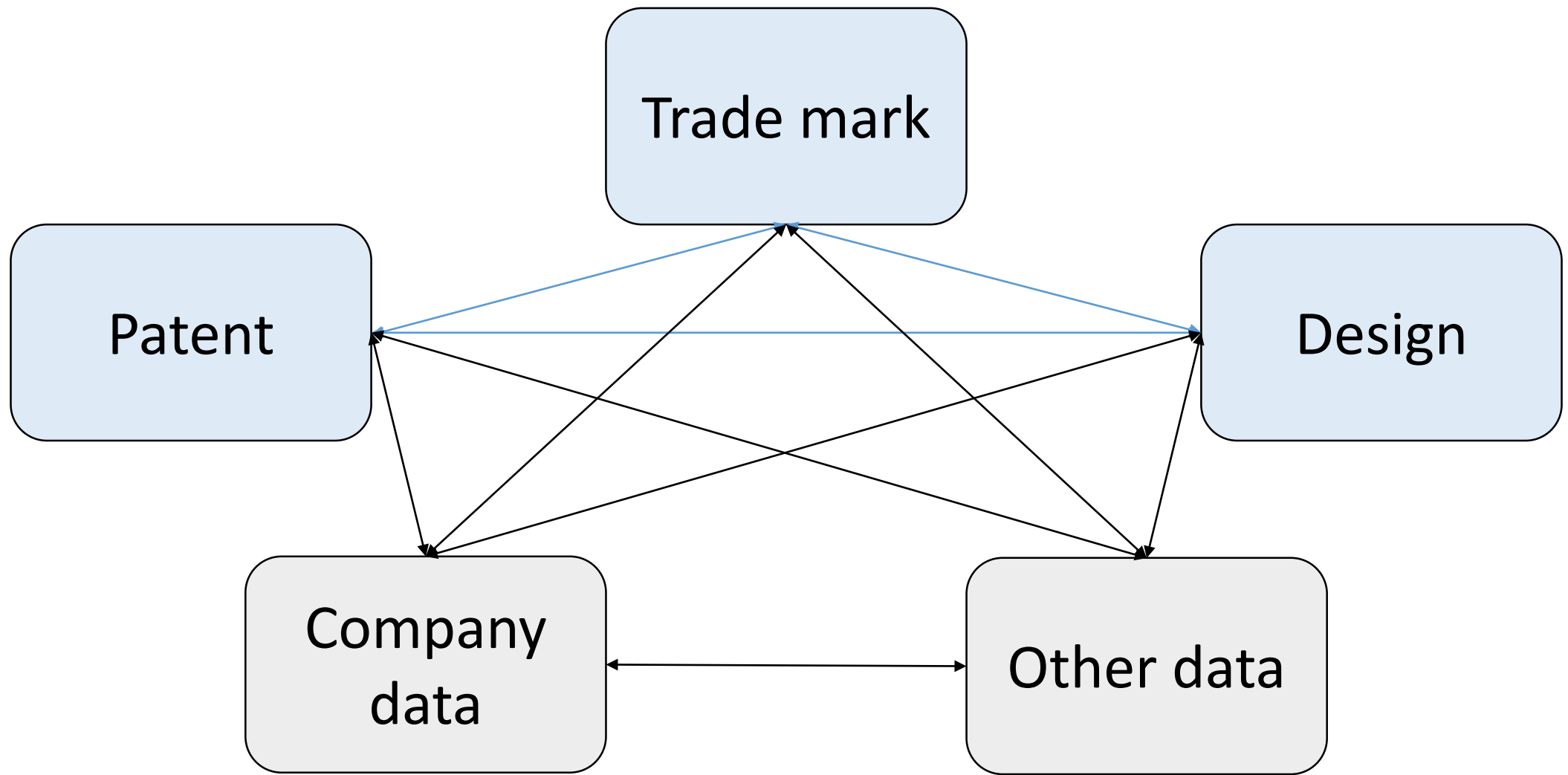
- Clean customer data

- Clean customer data

- Analysis







The process



Cleaning

Categorisation

Matching

Cleaning

Categorisation

Matching

ADP	Name
5593900	{Cross Manufacturing Company (1938){
7528847	{Cross Manufacturing (1938) Company{Limited
392571	{Cross{Manufacturing Company (1938) Limited
4723706	{CROSS MANUFACTURING COMPANY (1938) LTD{
85369049	{CROSS MANUFACTURING COMPANY _(1938) LIMITED{

Cleaning

Categorisation

Matching

ADP	Name	CleanName
5593900	{Cross Manufacturing Company (1938){	CROSS MANUFACTURING COMPANY (1938)
7528847	{Cross Manufacturing (1938) Company{Limited	CROSS MANUFACTURING (1938) COMPANY LIMITED
392571	{Cross{Manufacturing Company (1938) Limited	CROSS MANUFACTURING COMPANY (1938) LIMITED
4723706	{CROSS MANUFACTURING COMPANY (1938) LTD{	CROSS MANUFACTURING COMPANY (1938) LIMITED
85369049	{CROSS MANUFACTURING COMPANY _(1938) LIMITED{	CROSS MANUFACTURING COMPANY (1938) LIMITED

Cleaning



Categorisation

Matching

ADP	CleanName	Category
5593900	CROSS MANUFACTURING COMPANY (1938)	Company
7528847	CROSS MANUFACTURING (1938) COMPANY LIMITED	Company
392571	CROSS MANUFACTURING COMPANY (1938) LIMITED	Company
4723706	CROSS MANUFACTURING COMPANY (1938) LIMITED	Company
85369049	CROSS MANUFACTURING COMPANY (1938) LIMITED	Company




ADP	CleanName	Company name	Company number
5593900	CROSS MANUFACTURING COMPANY (1938)	CROSS MANUFACTURING COMPANY (1938) LIMITED	342798
7528847	CROSS MANUFACTURING (1938) COMPANY LIMITED	CROSS MANUFACTURING COMPANY (1938) LIMITED	342798
392571	CROSS MANUFACTURING COMPANY (1938) LIMITED	CROSS MANUFACTURING COMPANY (1938) LIMITED	342798
4723706	CROSS MANUFACTURING COMPANY (1938) LIMITED	CROSS MANUFACTURING COMPANY (1938) LIMITED	342798
85369049	CROSS MANUFACTURING COMPANY (1938) LIMITED	CROSS MANUFACTURING COMPANY (1938) LIMITED	342798


Next steps...



The Future

- The UK IPO is carrying out a transformational program to move from individual IP-centric systems to a customer-centric environment.
 - This requires the IPO to be able to identify all the IP Rights held by a customer – currently not possible because of:
 - the level of duplication of name & address data in individual systems
 - the lack of a unique identifier for customers both within a legacy system and across the rights
- 

The Solution?

- Some of the issues can be resolved by:
 - Improving our data capture
 - Validating against postal address files
 - Using properly laid out templates for the input/capture of non UK addresses
 - Checking for duplicates before data is loaded into backend systems
 - Collecting additional data such as company numbers and personal identifiers e.g. tax identifiers, passport numbers – if our legislation can be amended to allow this.
- 

Any Questions

?



