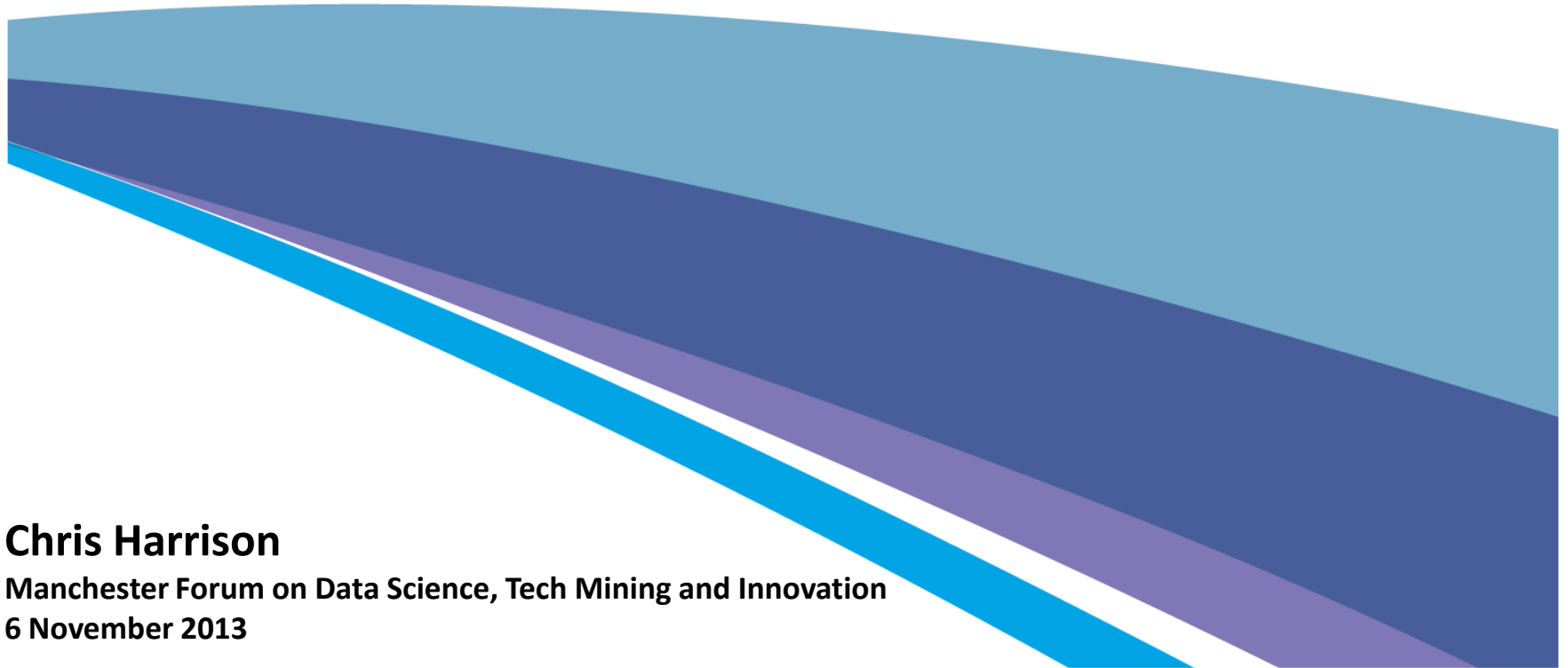




Intellectual
Property
Office

Limitations: Handling big data and other issues



Chris Harrison

Manchester Forum on Data Science, Tech Mining and Innovation

6 November 2013

Intellectual Property Office is an operating name of the Patent Office

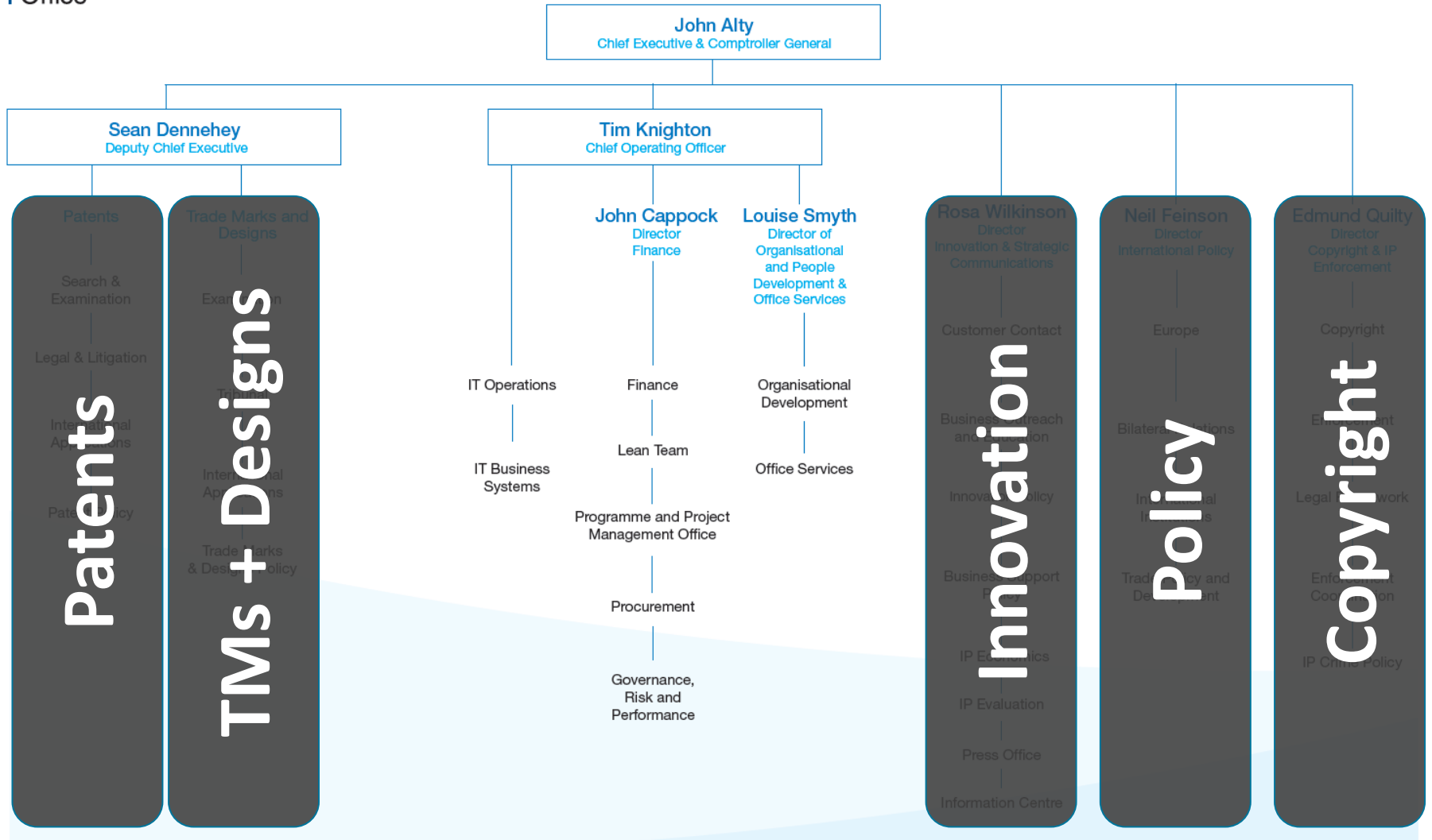


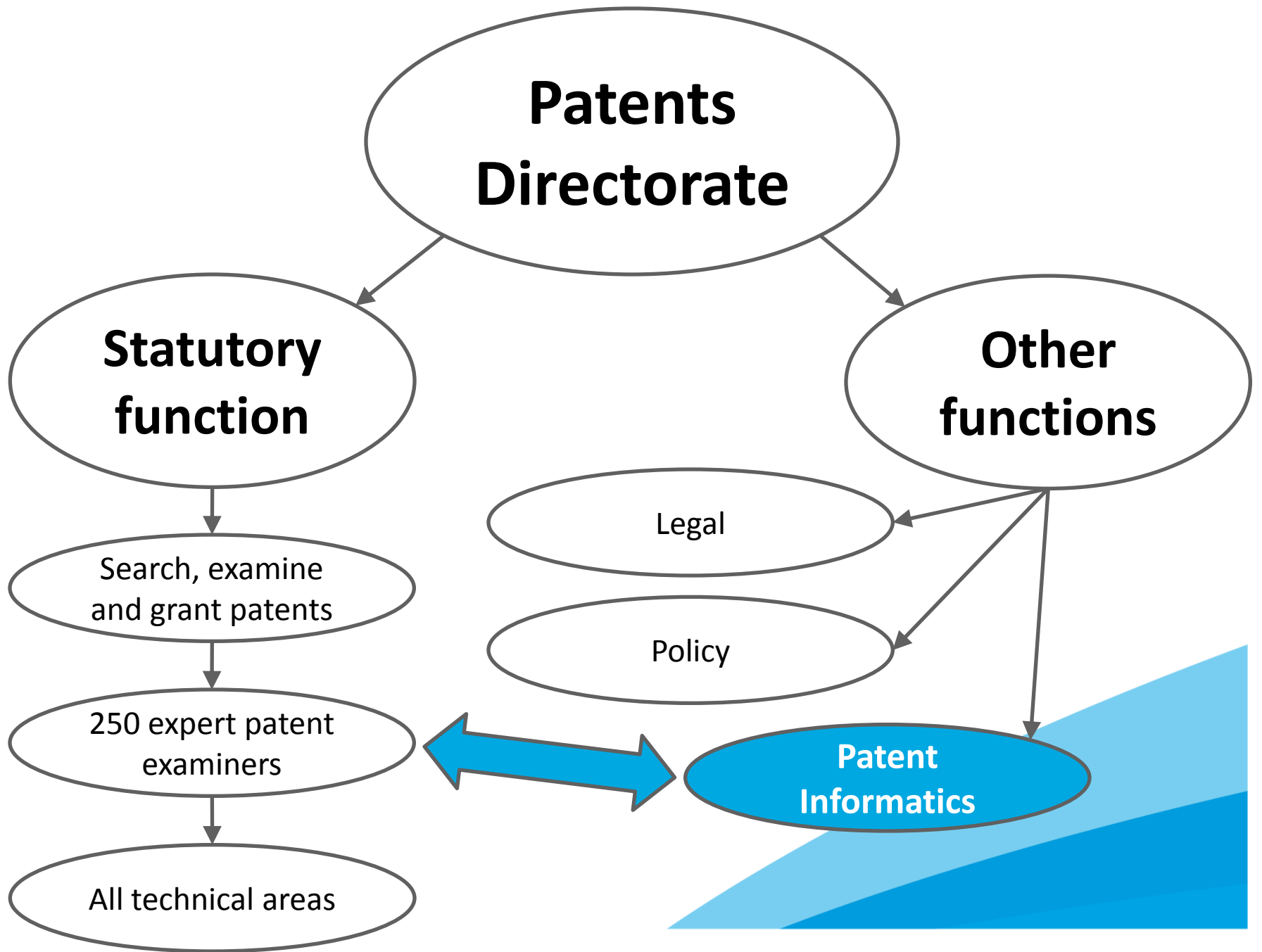
Intellectual
Property
Office

Formerly known as The Patent Office

- Established in 1852
- Located in Newport, South Wales since 1991







What we do

(12) UK Patent Application		(19) GB	(11) 2496587	(13) A
		(43) Date of A Publication		22.05.2013
(21) Application No:	1119305.9	(51) INT CL:	F16C 33/20 (2006.01) F16C 33/14 (2006.01)	
(22) Date of Filing:	09.11.2011	(56) Documents Cited:	GB 2465852 A GB 2436601 A EP 1717469 A2	
(71) Applicant(s):	Mahle International GmbH Pragstraße 26-46, 70376 Stuttgart, Germany Mahle Engine Systems UK Limited 2 Central Park Drive, Central Park, Rugby, Warwickshire, CV23 0WE, United Kingdom Mahle Composants Moteur France Sas PO Box BP104, 27 Route de la Foins, Chavanod 74856, France (including Overseas Departments and Territories)	(58) Field of Search:	INT CL F16C Other: WPI; EPODOC	
(72) Inventor(s):	Ian Matthew Laling Jonathan Forder Cedric Fortune			
(74) Agent and/or Address for Service:	Harrison Goddard Foote Belgrave Hall, Belgrave Street, LEEDS, SL2 8DD, United Kingdom			

(54) Title of the Invention: Sliding bearing
Abstract Title: A sliding bearing comprising three layers of material

(57) A sliding half-bearing 1 is provided. The sliding half bearing 1 comprises: a strong steel backing 2; a Cu-based or Al-based first layer 3, an Al-based second layer 4, and a polymer-based bearing material third layer 5. The second layer 4 contacts and is directly located between the first layer 3 and the third layer 5. The first layer 3 is other than a sputter coated layer, and the second layer 4 is a sputter coated layer.

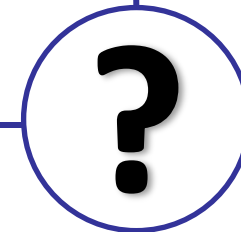


Figure 1

GB 2496587 A

When

What



Who

Where

Patent policy

Government

Academia

Industry

Patent professionals



Data sources

Patents

- European Patent Office
- OECD
- Thomson Reuters
- STN

Other IP data

- UK TM Register
- UK Design Register
- OHIM

Non-patent literature

- Web of Science
- INSPEC, IEEE, BIOSIS,
- IOP Journals, ...

Businesses

- Bureau van Dijk



Analysis tools

Search Technology: VantagePoint

- Custom scripting, NLP, fuzzy logic, clustering

Thomson Reuters: Thomson Innovation

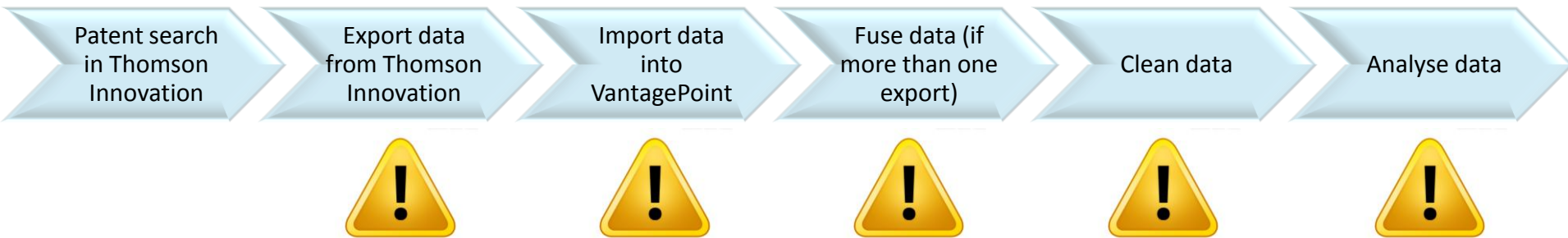
- Themescape mapping, citation analysis, clustering

Custom in-house software

- Data cleaning, linking/matching, text mining



Patent analysis



Big Data

Thomson Innovation

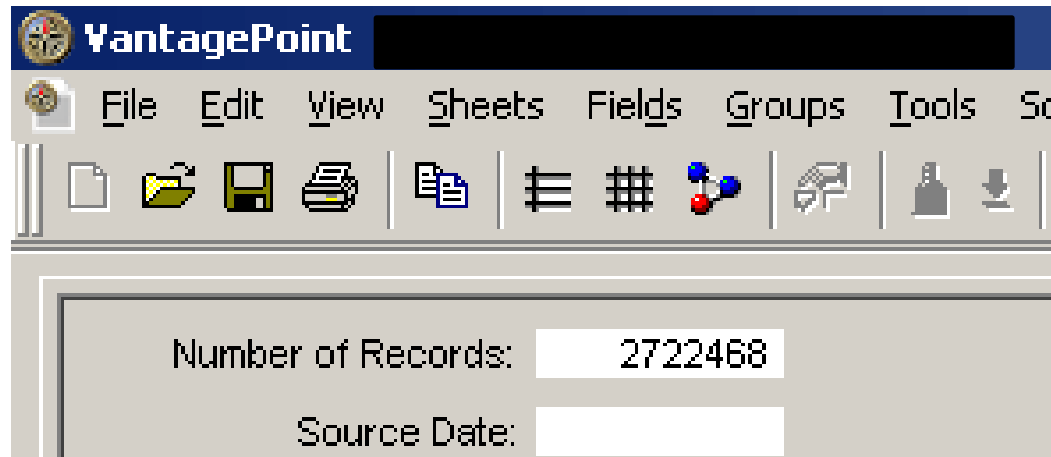
- Limited to 60k records per export
 - Multiple exports and data fusion required
- But can export one record per family, so 60k families (may result in >150k records)
- Themescape maps limited to 60k actual records



Big Data

VantagePoint

- Most projects 10k-50k records
- However...



Limitations

Computer hardware

We were running

- 32-bit VP on Windows XP single core CPU, 2GB RAM VM

Then

- 64-bit VP on Windows 7 dual-core CPU, 6GB RAM VM

Now we use

- 64-bit VP on Windows 7 dual-core CPU, SSD HDD, 16GB RAM PC

Is this enough?



Limitations

CPU/RAM issues

- Importing data
- Running scripts
- Data/Record fusion

VantagePoint file sizes

- VP v7 vs. VP v8



VantagePoint Memory Manager

VantagePoint

File Edit View Sheets Fields Group

0 Titles, 0 Selected

Title of the invention

Hotkey Manager

Categories

- Edit
- Fields
- File
- Groups
- Help
- Keyword List
- List
- Matrix
- Misc
- Navigation
- Scripts

Command De
Opens the me

Press New Hotke

None

Reset All Defa

Memory Manager

Dataset is using 7,353,092 bytes of RAM
Fields are using 6,843,768 bytes of RAM

Name	Active	Loaded	MemUsed	NumViews	Modified
~Raw Record	No	No	132	-	No
Abstract of the invention in English	No	Yes	2,181,701	-	Yes
Abstract of the invention in English (NLP)	No	Yes	2,252,908	-	Yes
Applicant Country	No	No	132	-	No
Applicant Country Expanded	No	Yes	86,927	-	Yes
Applicant Name	No	No	132	-	No
Applicant Name (long)	No	No	132	-	No
Applicant Name (long-with country)	No	No	132	-	No
Applicant Name Count	No	No	132	-	No
Application Country	No	No	132	-	No
Application Date	No	No	132	-	No
Application Number	No	No	132	-	No
Application Year	No	No	132	-	No
EPO classification	No	Yes	329,943	-	Yes
EPO family number	No	No	132	-	No
Inventor Country	No	No	132	-	No
Inventor Country Expanded	No	Yes	73,743	-	Yes
Inventor Name	No	Yes	787,655	-	Yes
Inventor Name (count)	No	No	132	-	No
Inventor Name (long)	No	No	132	-	No
Inventor Name (long-with country)	No	No	132	-	No
IPC -Advanced Level- Invention	No	Yes	679,099	-	Yes
Priority Countries	No	No	132	-	No
Priority Date (oldest)	No	No	132	-	No

Minimize Memory Use

Unload Field From Memory

OK

Publication number

3104

100%

Publication Number - no kind

3104

100%

Summary





Intellectual
Property
Office



pit@ipo.gov.uk

www.ipo.gov.uk/informatics

Intellectual Property Office is an operating name of the Patent Office