



Neutral Citation Number: [2019] EWHC 3007 (Ch)

Case No: IL-2018-000002

**IN THE HIGH COURT OF JUSTICE**  
**BUSINESS AND PROPERTY COURTS OF ENGLAND AND WALES**  
**INTELLECTUAL PROPERTY LIST (ChD)**

Royal Courts of Justice  
The Rolls Building  
7 Rolls Buildings  
London, EC4A 1NL

Date: 08/11/2019

Before :

**MR JUSTICE BIRSS**

Between:

**77m LIMITED** **Claimant**  
**- and -**  
**ORDNANCE SURVEY LIMITED** **Defendant**

and between:

**(1) ORDNANCE SURVEY LIMITED** **Part 20**  
**(2) GEOPPLACE LLP** **Claimants**  
**and**  
**(1) 77m LIMITED** **Part 20**  
**(2) THE KEEPER OF PUBLIC RECORDS** **Defendants**

-----  
-----  
**Jaani Riordan** (instructed by **Laceys**) for **77m Limited**  
**Lindsay Lane QC** and **Jessie Bowhill** (instructed by **Fieldfisher**) for **Ordnance Survey**  
**Limited and Geoplace LLP**  
**The Keeper of Public Records** did not appear and was not represented

Hearing dates: 17th - 19th, 22nd - 26th July 2019

-----  
**Approved Judgment**

I direct that pursuant to CPR PD 39A para 6.1 no official shorthand note shall be taken of this Judgment and that copies of this version as handed down may be treated as authentic.

.....  
MR JUSTICE BIRSS

**Mr Justice Birss :**

<b>Topic</b>	<b>Paragraphs</b>
Introduction	1
Background	8
The witnesses	53
What happened	69
How was Matrix created?	99
The issues	129
The contracts	138
INSPIRE Download Terms	139
A1 Match Licence	176
FAP licences and scraping	208
RoS Land Values Licence	231
Authority	251
Infringement of database right	259
Defences	296
Procuring breach of contract	328
Conclusion	342
<i>Confidential annex</i> [not included]	344

*Introduction*

1. The dispute is between 77m and Ordnance Survey (OS). 77m has created a dataset called Matrix to which it wishes to sell access. Matrix consists essentially of an up-to-date, detailed and accurate list of the geospatial coordinates of all the residential and non-residential addresses in Great Britain. It contains about 28 million records. As a product Matrix would directly compete with an existing OS product called AddressBase. AddressBase is a similarly up-to-date, detailed and accurate list linking all the addresses in Great Britain to geospatial coordinates. There is more to both products than that but none of it matters at this stage.
2. 77m created Matrix by accessing, combining and processing data from a wide range of datasets which were either publicly available for free or which 77m paid to access. Some of those datasets include data which derives at least to some extent from OS. However 77m did not contract with OS for access to AddressBase. The terms on which OS would allow access to its data in a manner which would permit a third party like 77m to produce something like Matrix, would not make it economic to compete with OS in the manner 77m wishes to do. The question is whether 77m has succeeded in its aim of creating the Matrix dataset without infringing any intellectual property rights held by OS. 77m contends that it has, and OS contends that it has not.
3. In creating Matrix, 77m accessed over 50 million records from the various sources it used. At least 18 datasets from different sources were accessed. Some details of what was done at 77m are trade secrets. One of the sources used to create Matrix was data from Her Majesty's Land Registry (HMLR). 77m entered into various contracts with HMLR to receive data from or have access to HMLR databases. One of the issues relates to the terms of those contracts. Another source of data was Registers of

Scotland (RoS). Among other things RoS carries out tasks in Scotland similar to the work of HMLR in England and Wales. Another source was data made public by Lichfield District Council under a Governmental open data initiative. There are issues about the scope of the terms on which the RoS and Lichfield data was obtained.

4. Other sources which 77m used included:
  - i) the Post Office Address File (the PAF), which is a commercial dataset from the Royal Mail which 77m paid for;
  - ii) CodePoint Open, which is a freely available open dataset from OS providing a representative geospatial location for each postcode in Great Britain;
  - iii) VectorMap District and OpenMap Local, which are essentially freely available, generalised, maps of Great Britain (from OS);
  - iv) Data from the Valuation Office Agency (VOA) which 77m licensed. The VOA is concerned with the rateable value of property and produces data linking addresses to rateable value.
  - v) Data from numerous other freely available sources such as a list of sporting facilities in England produced by Sport England and a list of churches produced by the Church of England.
5. That 77m had access to all this data is not confidential whereas aspects of what 77m did with some of it is.
6. Another aspect of the dispute is a claim by 77M that OS committed the tort of procuring a breach of contract. This is said to have happened when OS found out what HMLR and 77m were doing. The allegation is that OS procured or persuaded HMLR to stop dealing with 77m in the same way as it had been before. 77M says this amounted to procuring HMLR to breach a contract it had with 77m, of which, by then, OS was on notice.
7. There are two other parties to this case apart from 77m and OS. They are the Keeper of Public Records and GeoPlace. The Keeper is responsible for managing Crown copyright and database rights, including by granting delegations of authorities to public bodies to license protected material. This role was previously held by the Controller of HMSO until July 2017. The Keeper did not participate in these proceedings. The position of GeoPlace is dealt with below.

### *Background*

#### *Some concepts*

8. Today we are used to typing an address into a computer system such as a satnav or map website, and being presented with an image of part of a map with the location marked. This case is about how that is done. Going back to basics a map is a drawing, i.e. a graphic representation of a piece of land. To find a location on a map, you could look for it on the map, but that is not very efficient. So maps often have an index or gazetteer. This is a list of locations arranged in alphabetical order, or some other manner which makes finding an individual location convenient, and against that

location in the gazetteer is some reference information which allows you to find it on the map— such as a grid reference or some other “geospatial” coordinates.

9. At the risk of labouring the point, no matter how good they are, neither a list of addresses nor a map on its own will do the job. The critical thing is the ability to link the two together.
10. In the United Kingdom the postcode system provides a scheme for uniquely identifying addressable properties based on a house number or property name and the postcode. This is administered by the Royal Mail. The dataset, called the PAF, is commercially available. It has about 30 million records. The data is regularly updated and is formatted in a consistent way to facilitate its use. There are no relevant restrictions on its use once a user has paid for it. The PAF is very useful since, for example, given a postcode the PAF will provide the full postal address for properties with that postcode. Although there are exceptions, most of the time there will be a number of individual addresses with the same postcode.
11. However the PAF is not geolocated. In other words no geospatial information is provided in the PAF. Therefore the PAF alone will allow someone to put the correct address on a parcel, but it will not, on its own, give them the information necessary to run their own delivery service.
12. Also the PAF necessarily only contains postal locations, i.e. places to which post might be addressed. There are other kinds of locations in which users might be interested. These are buildings and other places to which post is not normally addressed. They can be buildings like village halls, churches and petrol stations but also places like playing fields, ponds and electricity sub-stations. The geospatial locations of all these places can be very useful e.g. to the emergency services. While the PAF is not a source for this information, local authorities can be.
13. The UK has a standardised geospatial coordinate system (the British National Grid) based on X-Y coordinates with the origin somewhere just southwest of Scilly. These allow one to accurately specify locations on the ground and on a map. GPS coordinates are another kind of geospatial information.
14. In electronic mapping, parcels of land are defined by polygons. An individual polygon is defined by the set of coordinates of its vertices. So a square parcel of land (or building) would have four vertices and the polygon would be defined by four pairs of X-Y coordinates. In order to give a place a single geospatial coordinate a common technique is to use the centroid of the relevant polygon. For some funny shaped polygons the centroid can be outside the boundaries of the polygon, in which case one might define a point inside the polygon nearest the centroid instead. Deriving a centroid from a polygon is a one way process. A given polygon has only one centroid but given a centroid, you cannot derive the shape or extent of the polygon. Infinitely many polygons have the same centroid.
15. One can create new geospatial data in various ways. One way is to carry out a survey either by surveyors physically on the ground or by remote sensing such as by satellite or aerial photography. Today this data is or will be digitised. Another way would be to carry out data processing on existing datasets. That is what happened in this case.

16. In general terms the processing of “big data” like this involves a number of common techniques. One is matching. This way one can attribute information held in one dataset to data in another dataset. For example if one has a list of addresses, one can use the address to look up a record for that address in another dataset which may hold information about the type of building. However automatically matching data between two datasets is not always straightforward. For example one needs to write suitable code in order to identify that “One Acacia Ave, Saint Albans” is the same as “1 acacia avenue, St. Albans” and there are numerous other factors to take into account.
17. Other kinds of data processing may involve interpolation between known values to estimate missing values. The dataset may need to be processed in other ways too, such as “cleansing” to remove and correct errors, de-duplication, and reformatting. In general terms these techniques are not specific to geospatial data but apply to all datasets, although the details will vary.
18. Finally another important basic point is that keeping data up to date is crucial. Simply as an example, there are approximately 15,000-20,000 new buildings in the United Kingdom every month as well as 100,000 property transactions.
19. In these proceedings the terms dataset and database are often used interchangeably although they have different shades of meaning. A set of data can be made available to a third party by providing a copy of the entire dataset to that person. In practice (ignoring legal obligations) that gives the third party the practical ability to manipulate that dataset in any manner they wish. Alternatively access to data may be provided by allowing them to consult the database holding the data through some kind of interface. The user can ask questions and be provided with answers. This may be through a website used by the public or it could be by allowing a user to make database queries, usually by means of computer code called SQL (structured query language). This kind of access usually never provides the user with the entire dataset in one go. In describing the facts in this case it is neater to describe a set of data as a dataset and to use the term database to refer to a system which contains a dataset (or datasets) and allows queries to be made. Nevertheless one needs to recognise that the term dataset is not used in the legislation concerning database rights.

*The Ordnance Survey*

20. OS is the national mapping agency of Great Britain. It creates, maintains and disseminates geospatial and cartographic data and products. OS has been operating in various forms since the 1700s and has been under legislation dating back to the Ordnance Survey Act 1841.
21. From 1999 to 2015, OS operated as a trading fund. During this period OS was part of the government but financed its operations independently through a commercial licensing model. Any databases or copyright material created by OS vested automatically in the Crown. To ensure that OS was able to deal with this material, it was granted a number of delegations of authority from the Controller of Her Majesty’s Stationary Office. The latest delegation is dated 18 October 2007 and was in force until April 2015. The construction of that delegation is one of the issues in dispute.

22. On 1 April 2015, Ordnance Survey transferred its functions to the defendant company, Ordnance Survey Limited, a private company wholly owned by the Secretary of State for Business, Energy & Industrial Strategy (BEIS). I use the term OS to refer compendiously to the company or the trading fund. Since OS became a company, any intellectual property created by OS no longer vests automatically in the Crown. OS and the Controller of HMSO therefore entered into a Crown Rights Agreement (the “CRA”) on 31 March 2015. The construction of the CRA is another issue in dispute. For present purposes it suffices to say that certain intellectual property created by OS before 1 April 2015 is owned by the Crown but licensed to OS, and certain new material is assigned to the Crown and licensed to OS. After taking over from the Controller in July 2017, the Keeper entered into a second CRA with OS on 24 August 2017 on the same terms as before.
23. OS is required under its constitutional and shareholder documents to publish a statement of its public task which defines its core public role and function. As part of its public task OS has to make the content of its datasets widely available. The dissemination of its data and products is subject to various regulatory requirements, including the Re-use of Public Sector Information Regulations 2015 and competition legislation.

#### *OS databases and licensing*

24. OS is required to license its products on a non-discriminatory basis. It maintains a number of relevant datasets. Some datasets are made available to the public for free under the Open Government Licence (OGL) and others are available on terms. OS grants licences to public sector organisations and to private companies.
25. Licensing of the public sector by OS in England and Wales is under the Public Sector Mapping Agreement (PSMA) while in Scotland it is under the One Scotland Mapping Agreement (OSMA). These terms are negotiated with BEIS and the Scottish Government respectively. The PSMA was first introduced on 1 April 2011. OS granted a PSMA Member Licence to Lichfield District Council on 1 April 2011 and to HMLR on 12 July 2011. The OSMA became effective on 1 April 2013. OS granted an OSMA Member Licence to RoS on 29<sup>th</sup> May 2013.
26. Private sector companies who wish to supply OS data to their customers can become “Licensed Partners” under the Framework Contract (Partners). The terms of the onward supply are subject to further agreements. Direct customers are licensed under the Framework Contract (Direct Customers).

#### *The OS Topo database*

27. A major OS database product is called Topo. Topo’s full name is the OS MasterMap Topography Layer. Topo is a detailed vector spatial dataset of the whole land area of Great Britain and includes geometric representations of every feature and physical boundary. It contains over 117 million polygons, over 325 million line geometries showing boundaries of the polygons and other geographical features, and over 21.6 million cartographic text labels. Each of the geometries has been captured at some point in the history of OS from either ground or aerial survey. In other words Topo is a map of Great Britain. It is the most accurate, detailed and up to date map of Great

Britain available. OS makes Topo available to the private sector on commercial terms. It is not available outside the public sector for free.

*More generalised OS databases*

28. Another database product from OS is OpenMap Local (OML). This is released as an open dataset and is freely available for download from the OS website under the terms of the Open Government Licence (OGL). It is a “generalised” map in the sense that it contains much less detail than Topo. The geometries have been simplified and amalgamated to reduce the level of detail. This makes it useful for map making but at the expense of positional accuracy and the ability to identify distinct features. For example OML contains polygons representing buildings but they are simplified so that, for example, a terrace of properties would typically appear as one long rectangle. OML contains various other features such as roads (as sets of lines), important buildings (as polygons), railway tracks, stations and tunnels (as points and lines) and woodland (as polygons).
29. Similar to OML is VectorMap District (VMD). This is another open dataset made freely available by OS. It is designed to be viewed at a smaller (i.e. more distant) scale than OML.
30. Another open dataset is called OS Streetview (not to be confused with Google Streetview). This is a raster map of Great Britain, in other words it is simply a pictorial map rather than a set of vector data. It has been discontinued.

*Geolocated address products from OS*

31. In 1996 OS launched a database combining addresses and geospatial data. It was called AddressPoint. It was the first address database which assigned geospatial coordinates to each address. OS achieved this by matching the delivery addresses produced by Royal Mail – as recorded in the PAF – to a location within OS mapping data, including Topo. The location assigned to the address is a single point representing the address.
32. In 2001 AddressPoint was replaced by AddressLayer and in turn AddressLayer was replaced by AddressLayer 2, launched in 2006. Addresslayer 2 provided additional features. One in particular was the inclusion of non-postal addresses.
33. OS no longer markets AddressPoint, AddressLayer or AddressLayer 2.
34. Their successor is AddressBase. It has all the features of AddressLayer2 mentioned above. It was launched in September 2011 with coverage for England and Wales. It was extended to Scotland in April 2012. AddressBase is the current geospatial addressing product offered by OS. It is offered with different levels of functionality. AddressBase itself contains verified address records with geospatial coordinates. The addresses are mainly from the PAF. AddressBase Plus also includes additional non-postal addresses not in the PAF. AddressBase Premium adds further data such as historical and provisional information (e.g. from planning applications).
35. Like Topo, AddressBase and its predecessors are or were available to the private sector on commercial terms. They are not freely available.

36. However there is a freely available geolocated address dataset available from OS. This is called CodePoint Open. This dataset consists of a representative point location for each postcode in Great Britain. It is available under the OGL. Bearing in mind that a postcode normally represents a group of addresses, the point location for a postcode in CodePoint Open is derived by taking a form of “average” of all the address coordinates in that postcode. The individual address coordinates used are the ones in AddressBase. The averaging process works mathematically by taking the centroid of the group of individual address locations and then selecting the nearest actual address location to that centroid.
37. To take an example relevant to this case: a delivery van taking a parcel to a particular house using only CodePoint Open could use the postcode in the address to extract a single geolocation which represents the group of houses with that postcode but that would not get them to the door. If the location was a simple street with clearly and logically numbered houses then finding the door might be easy but the real world is often much more complicated: houses may not be numbered at all, the street plan and the numbering scheme may be more complicated.
38. There is a paid-for version of CodePoint Open called CodePoint which contains extra information but it is irrelevant.

#### *GeoPlace*

39. GeoPlace is a limited liability partnership founded in 2010 as a joint venture between OS and the Local Government Association. The aim of the joint venture was to create a single database consolidating the addressing information held by local authorities, the data held by OS, the PAF from Royal Mail and other third parties such as the VOA. The information from local authorities is recorded in a database called the National Land and Property Gazetteer. The significance of the local authorities is that they have the most up to date information about changes in their areas. Examples would be new housing developments and name changes of existing buildings. There are many others.
40. The database developed and maintained by GeoPlace is called the National Address Gazetteer (the “NAG”). Thus the NAG consolidates spatial address data from OS, local authorities, the Royal Mail and third parties.
41. OS licenses the NAG from GeoPlace. AddressBase is extracted from the NAG. The extent to which there is a relevant difference between the NAG and AddressBase and any consequences which flow from that, are questions I need to address.

#### *HMLR, INSPIRE and RoS*

42. The principal role of HMLR is to maintain the Register of Title covering ownership of land and property in England and Wales in accordance with the Land Registration Act 2002. As a trading fund HMLR must generate its own revenue. The Register of Title includes Register entries and a Title plan. For a given entry, in addition to the unique title number and date, the register contains three registers A, B and C. Register A is the Property register, B is the Proprietorship register and C is the Charges register. The property register A includes text in numbered paragraphs. The first one (A1) is the property description. This is the familiar text which generally takes the form



*“(date) The Freehold land shown edged in red on the plan of the above title filed at Land Registry and being 1 Acacia Avenue, London, AB1 3XX”.* The Title Plan held by HMLR illustrates the boundaries of the property. By law it must be based on the Ordnance Survey map (paragraph 5(a) of the Land Registration Rules 2003). The Title Plan consists of a map of the locality in which the property is situated with the boundaries of the relevant property marked.

43. In 2009 the INSPIRE regulations were brought into force in order to comply with the INSPIRE Directive (Directive 2007/2/EC). INSPIRE stands for “Infrastructure for Spatial Information in Europe”. The legislation applicable in Scotland is set out in the INSPIRE (Scotland) Regulations 2009 (SSI 2009/440 as amended by SSI 2012/284) while the legislation applicable in the rest of the United Kingdom is set out in the INSPIRE Regulations 2009 (SI 2009/3157 as amended by SI 2012/1762). The purpose of the Directive and the national legislation concerns the creation and operation of national and EU infrastructures relating to spatial information for the purposes of EU environmental policies and other policies or activities which may have an impact on the environment. In other words public authorities were required by law to make spatial datasets for which they were responsible available to the public in various ways. The regulations define various kinds of services which must be offered such as the ability to search and view the data and to download copies. The public authorities were not obliged to provide all these services for free. Some services could be charged for.
44. Part of the scheme of the INSPIRE regulations is that public authorities must create, maintain and make available certain metadata for the spatial data sets. The metadata must correspond to certain themes. One of the themes is “Cadastral parcels”, which is defined as “Areas defined by cadastral registers or equivalent”. A cadastral register is a register of title to land.
45. At all relevant times, well before the INSPIRE regulations, HMLR maintained an internal database of geospatial data in the form of polygons which indicated the boundaries of properties. These were called the Index Polygons. Some lines in the Index Polygons were taken directly from the Topo data of OS such as a line representing a wall. Other features in the Index Polygons were drawn by HMLR itself although they were drawn to be consistent with Topo. There are about 28 million Index Polygons for England and Wales showing the boundaries of each registered freehold title of land or property.
46. The Title Plan which HMLR does make available was in a graphic form and only showed the boundary in general terms. From 2008 onwards HMLR offered a paid for service whereby a user could download polygons at a rate of £1.50 per polygon. A further service called Polygon Plus provided information from the register along with the polygon for an extra £3.
47. However in order to comply with the INSPIRE regulations and make its geospatial data available, HMLR created and made available a set of polygons known as the INSPIRE polygons. These were taken from the Index Polygons and relate only to registered freehold properties. For a given freehold property, the INSPIRE polygon marks the boundary. The data was structured to meet the requirements of the cadastral data specification. Given the way the Index Polygons were and are created, the INSPIRE polygons are based to a significant extent on the Topo data. HMLR

allocated a unique reference to each INSPIRE polygon. These are called the INSPIRE IDs.

48. HMLR started an INSPIRE view service in December 2011. This allows users to view but not download map images of the 21 million INSPIRE polygons. It is subject to an end user licence. HMLR offered a paid for download service which was launched in September 2013.
49. Another service HMLR offers is called “Find a Property” or FAP. It is a statutory service which allows people to search for titles. It is made available on the internet by HMLR. A user can enter an address/postcode or title number or they can click on a point on a map. The service will provide the relevant Land Register entry. Since the INSPIRE Regulations were implemented the FAP service also allows a user to type in the INSPIRE ID. The FAP service will be addressed in more detail below.
50. Just as HMLR maintains the register of title for England and Wales, one of the tasks of RoS is to do the same thing in Scotland. It has a dataset of polygons and, following the INSPIRE initiative RoS made a dataset relating to freehold titles in Scotland available. However unlike HMLR, the RoS INSPIRE dataset did not consist of polygons, it only consisted of a set of “seed points”, one for each title, along with a unique ID. The seed point is the centroid of the relevant property. So it provides a geospatial location for that property but does not provide the boundaries. The RoS INSPIRE dataset is available for downloading under the RoS INSPIRE licence.

#### *Valuation data*

51. For England and Wales the VOA makes a dataset available which links valuation data to addresses. In Scotland a similar dataset of land values is made available by RoS. Nothing turns on the terms under which the VOA makes information available but there is an issue about the terms on which RoS made a land values dataset available to 77m.

#### *Lichfield*

52. As mentioned already, at one stage Lichfield District Council made data available under the OGL. This included addresses and geospatial coordinates as well as other data.

#### **Witnesses**

53. 77m called a single fact witness, Mr Highland. Mr Highland is the co-founder of 77m with responsibility for the management and day-to-day operations of the business. Mr Highland gave evidence on the development and commercialisation of Matrix, the licensing of the datasets in issue, and the communications with OS and HMLR in this regard. I do not trust uncorroborated evidence from Mr Highland. Two examples are sufficient to show why. His statements before trial about the reasons why there was a shift in the geocoded point in Matrix were deliberately misleading. Also Mr Highland deleted or allowed to be deleted the contents of a relevant laptop when he knew relevant documents were to be preserved.

54. OS called nine fact witnesses: Jonathan O'Meara, Santiago Jagot, Christopher Chambers, Duncan Moss, Nicholas Griffiths, Gareth Robson, Lynne Nicholson, Cara Wiles (née O'Brien), and James Cutler.
55. Mr O'Meara is the Head of Legal Services and Company Secretary at OS. The focus of his evidence was on the licensing arrangements between OS and the public sector pursuant to PSMA and OSMA. He also provided a brief overview of the history of OS and its public function and role today.
56. Mr Jagot is a Senior Pricing and Licensing Manager at OS. Mr Jagot gave evidence on the development of the commercial licensing model at OS and communications with 77m, HMLR and RoS during the relevant period. His oral evidence was defensive. I would be wary of placing weight on any uncorroborated testimony contrary to 77m's case.
57. Mr Chambers is the Strategic Programme Lead at OS. Prior to his current role, Mr Chambers was a Product Manager and Senior Product Manager for the addressing products offered by OS. He discussed the development and functionality of AddressBase and its predecessors. Mr Chambers was not cross-examined.
58. Mr Moss is the Principal Consultant for Scotland at OS. He explains how RoS is using OS data to produce and maintain its own datasets, including the centroids or "seed points" of properties in Scotland and the geographic coordinates in its land sales data reports. His evidence was not cross-examined.
59. Mr Griffiths is the Executive Director of Informatics at GeoPlace, the custodian of the NAG. Mr Griffiths explains how GeoPlace produces and maintains the NAG by consolidating and verifying address data from OS, local authorities and third parties such as Royal Mail and VOA.
60. Mr Robson is the Head of Data Capture and Management at HMLR. Mr Robson explains how various HMLR products and services rely on input data from OS, including the INSPIRE Polygons, the Register of Title and the Find a Property service.
61. Ms Nicholson is a Senior Product Manager at HMLR. Ms Nicholson was one of the primary points of contact for 77m during the relevant period, first as a Product Manager and later as a Senior Product Manager. She explained how HMLR responded to various data requests made by 77m, including in relation to the INSPIRE Polygons and the property descriptions in the A1 Match data. Her evidence also addressed the relationship between HMLR and OS and particularly the discussions with Mr Jagot concerning the data supplies made to 77m.
62. Ms Wiles (previously O'Brien) is a Product Manager at HMLR. Ms Wiles was the Account Manager for 77m from September 2013 until November 2014. She gave evidence on the negotiation of the A1 Match Licence and the supply of the A1 Match data.
63. James Cutler is the Chief Executive at emapsite.com Limited, a technology company providing geographical data to customers through a centralised platform. His company is an OS Licensed Partner and Mr Cutler is a member of the OS Partner

Advisory Council. Mr Cutler explained how he interprets the concept of “internal business use” in licensing agreements with OS. He also addressed a point that emapsite considered buying Matrix from 77m but ultimately did not do so.

64. 77m called a single expert witness, Dr Benjamin Halstead. Dr Halstead is the Chief Technical Officer of 77m. He joined the company in April 2017, replacing Dr Kieron Brown. Dr Halstead examined the source code of Matrix and explained how 77m used the INSPIRE Polygons, the A1 Match Data, the FAP service and the RoS Land Value data in the development of the product. He also considered the relationship between AddressBase and various datasets provided by HMLR, RoS and Royal Mail.
65. OS served reports from two expert witnesses: Jonathan Simmons and Dr David Greaves.
66. Mr Simmons is the Head of Data Science & Analytics at OS. Mr Simmons reviewed the Matrix processes and data sources by reference to the Product and Process Description prepared by 77m in the context of these proceedings and the disclosure material. Mr Simmons saw a brief demonstration of Matrix in October 2016 but did not have access to the Matrix source code.
67. Dr Greaves is a Senior Lecturer in Computing Science at the University of Cambridge and a chartered engineer. His report had addressed source code provided by 77m but his evidence was not relied on at trial by OS.
68. Aside from the two individual witnesses about whom I have made comments above, all the other witnesses who gave oral evidence at trial gave their evidence fairly, seeking to help the court.

*What happened*

69. 77m is a start-up company founded by Philip Highland and Graham Allison in 2010. Mr Highland has a background in data analysis, particularly in the insurance industry and had been thinking about bringing a geospatial product to the market since the mid 1990s. Two spurs to his thinking were the release of a complete dataset of addresses and rateable values by the VOA in 1995 and in 1999 the availability of high quality mapping data by a company called Getmapping. According to Mr Highland a drawback of the data produced by the VOA at that time was the need for cleansing.
70. Mr Allison had previously worked for Getmapping and a company called Streetwise Maps. Streetwise is an authorised Ordnance Survey reseller. Mr Allison remains at Streetwise. He donated all the time he could spare to 77m. A programmer called Mr Mudassir was seconded from Streetwise to 77m as a contractor in late 2010.
71. The initial idea was to create and sell an accurate geospatial database of non-residential addresses. The plan was to do this using data processing techniques and matching data from available sources. By the end of 2010 work had started to identify possible sources and initial programming and testing was being done on matching addresses.
72. At some point 77m bought access to the PAF from Royal Mail. There is a licence dated 2014. Although it is not clear to me whether that was the first time 77m had

access to the PAF, nothing turns on that. 77m has had monthly updates for the PAF ever since.

73. In December 2011 77m entered into an agreement to obtain non-domestic rating data (with weekly updates) from the VOA. I have not had my attention drawn to the terms of that licence but nothing turns on them. This data amounted to about 2 million records.
74. 77m also obtained the open OS geospatial datasets OML and VMD as well as the open address dataset CodePoint Open. For each of these datasets as I understand it 77m obtains or obtained the relevant updates. Exactly when any of this took place is not clear but does not matter.
75. The first contact between 77m (Mr Highland) and HMLR seems to have been in 2011 (with Ms Nicholson) but the first important step took place in April 2012 when Mr Highland and staff at HMLR started discussing the purchase of large numbers of INSPIRE polygons by 77m. Mr Highland met Ms Nicholson in July 2012. The staff in touch were Mr Highland and Mr Allison of 77m and were Ms Nicholson and Victoria Abbott, the account manager, for HMLR. Although the detail of much of the discussions is not relevant, it is notable that 77m, and in particular Mr Highland, were in close and regular contact with HMLR from this period until 2015.
76. Meanwhile in July 2012 77m entered into a licence with the RoS whereby RoS would provide land value data relating to property in Scotland. The data was called historic "Land Values plus House Type". 77m acquired about 1.3 million such records from RoS. The meaning and scope of this licence is disputed and will be addressed below. There is also an issue about what use 77m made of this data. 77m contended that no relevant use was made of the data obtained under this 2012 licence.
77. In September 2012 there were negotiations (including a draft contract) between HMLR and 77m whereby the entire INSPIRE polygon dataset would be released to 77m for a one off payment of £60,000. At the last minute in late November 2012 HMLR told 77M they could not go ahead. The reason was that OS considered it had rights in that data and did not accept HMLR could license it without reference to OS.
78. In 2013, negotiations between 77m and HMLR continued.
79. Also in that period there were discussions behind the scenes between OS and HMLR. Documents produced to 77m in 2015 as a result of a Freedom of Information (FOI) request shed some light on the OS/HMLR negotiations. In any event HMLR eventually promulgated a set of licence terms under which the INSPIRE polygons would be made available to the public for downloading for free. These can be called the "INSPIRE Download Terms". The meaning of these terms is one of the central issues in the case.
80. On 16<sup>th</sup> September 2013 HMLR's INSPIRE download service was launched on the INSPIRE Download Terms. On the same day 77m downloaded complete copies of the entire INSPIRE polygon dataset. This data consists only of the geospatial data defining the various polygons together with the relevant INSPIRE ID for that polygon. There is no address data and no description of the property in this data. Since then 77m has continued to download INSPIRE polygons when they are

available, at least until December 2016. The use of the INSPIRE polygons is subject to the INSPIRE Download Terms.

81. Also in September 2013 Cara Wiles (previously O'Brien) of HMLR took over responsibility for the account with 77m from Ms Abbott.
82. In an email from Cara Wiles to Mr Highland on 16<sup>th</sup> September 2013, Ms Wiles explains to Mr Highland that while the INSPIRE data which can be downloaded contains no title number or title information, customers can search for title details using the INSPIRE ID in HMLR's FAP website.
83. Also in September 2013 there were discussions between 77m and OS about a possible licence from OS relating to the INSPIRE polygons (the INSPIRE Download Terms refer to contacting OS in certain circumstances).

#### *The A1 Match data*

84. In October 2013 discussions started between Mr Highland and HMLR about a bespoke service to be provided by HMLR to 77m. This led to a contract called the A1 Match Licence which was entered into between 77m and HMLR on 14<sup>th</sup> February 2014. Various aspects of this are disputed, but what is not in dispute is that under the contract 77m would provide nearly 1 million INSPIRE IDs to HMLR (the exact number in the contract was 910,000) and in return HMLR would provide the A1 property descriptions from the Register of Title for each of those INSPIRE IDs. The price for this service was £2,500 plus VAT. The A1 property description allows one to find out what sort of property the ID refers to (e.g. an electricity sub-station) but it also gives the address for the property. The batches of A1 property descriptions were put together into response files by HMLR and provided to 77m.
85. The service under the A1 Match Licence is in effect a bulk version of what one could find out by laboriously entering all these INSPIRE IDs into the FAP website. 77m did that too. Its pleaded case is this started in January 2014 and ran until January 2017. I will come back to this below.
86. A specialist programmer, Dr Kieron Brown, joined 77m in January 2014. He left in 2017.
87. A second tranche of bulk A1 Matching was discussed in March 2014. This time 210,000 IDs were matched to A1 property descriptions for a fee of £600 plus VAT. A third tranche was undertaken in June 2014. That involved matching 288,000 IDs for £1,000. A fourth tranche of A1 property descriptions were supplied in September 2014. This time there were 2.5 million IDs and the price was £4,500. There was a fifth tranche (3.5 million IDs) and sixth tranche (about 0.9 million IDs) such that by February 2015 it seems that HMLR had supplied A1 property descriptions for a total of about 8 million INSPIRE IDs to 77m and 77m had paid about £20,000 for the service. No further written contracts were signed apart from the original one dated 14<sup>th</sup> February 2014.
88. Despite the apparent clarity in the numbers of A1 descriptions supplied this way, in fact it turned out in May 2019 that HMLR had actually provided 12.8 million A1 property descriptions to 77m this way. The discrepancy between 8 million and 12.8

million is not explained but there is no suggestion that 77m obtained the balance in any other way. All 12.8 million records were in the response files from HMLR. Nothing therefore turns on the difference between 8 million and 12.8 million. However shortly before trial (28<sup>th</sup> June 2019) it then emerged via Dr Halstead that 77m actually has in its possession 16.3 million INSPIRE IDs linked to A1 property descriptions. The source of the further 3.5 million is not at all clear and is disputed.

89. Meanwhile in January 2015 77m and OS entered into two contracts – a Framework Partners Contract and a Land Registry Polygons Contract. As I understand it, in the end nothing turns on these contracts.
90. In June 2015 RoS made its INSPIRE seed point data available for free downloading under the RoS INSPIRE licence. 77m downloaded those records (about 1.5 million). Although there was an issue about the terms of the RoS INSPIRE licence, it was dropped by OS at trial. It is accepted that whatever use 77m made of that data was lawful.
91. In August 2015 77m made a seventh request to HMLR for A1 property descriptions.
92. In September 2015 Santiago Jagot of OS emailed Ms Nicholson of HMLR to ask about the supply of addresses to 77m. He had heard from Mr Highland that 77m had an agreement with HMLR about this. HMLR never supplied any further A1 property descriptions to 77m after that. 77m contends that was a breach of contract by HMLR and that the breach was procured by OS.
93. At this stage it is convenient to say more about 77m’s use of the FAP service. 77m’s case is that it did use the FAP service to match INSPIRE IDs to addresses for about 480,000 address records. Its case is that this was done entirely manually by employees or contractors. There is clear evidence of contracts with staff working in Pakistan. Extracting 480,000 addresses manually using contractors in Pakistan is credible, not least because there are invoices to substantiate that case. However that does not explain the recently discovered 3.5 million extra A1 property descriptions. Those 3.5 million did not come from the automated A1 match service provided by HMLR therefore, contends OS, they must have come from the FAP service and that can only be explained by the use of an automated tool to access records via the FAP website, i.e. a process called “scraping”. Scraping would be outside what OS contends are the applicable terms for the FAP service.
94. Turning back to the chronology, in October 2015 77m launched Matrix, offering licences to customers. At about the same time 77m’s then solicitors Gordon Dadds wrote to OS. The thrust of the letter was to assert 77m’s case that it had acted lawfully, that OS was seeking to prevent the completion of the Matrix product, and that OS was abusing its dominant position contrary to competition law.
95. 77m downloaded the open datasets posted by Lichfield District Council in April 2016. There were about 52,000 records.
96. Following a letter of claim in April 2016, proceedings were issued by 77m against OS on 13<sup>th</sup> September 2016. The action was in the Intellectual Property Enterprise Court. The main relief sought was declarations that 77m had acted lawfully and Matrix did not infringe any relevant OS rights. OS served a Defence and Counterclaim, claiming

infringement of copyright and database right. By later amendments GeoPlace and the Keeper were added as parties and the claim for procuring breach of contract was included. The order permitting the amendment to bring a claim for procuring breach only allowed that claim to be brought based on procuring after September (or, in practice, August) 2016 (which was when OS obtained the annexes to the Particulars of Claim) on the basis that the claim before that date was unarguable. A belated attempt in the middle of trial to amend the procuring plea to cover a much earlier period was dropped.

97. There was a demonstration of Matrix to OS in October 2016. Also in October 77m supplied a version of Matrix to Hastings Insurance. This version included street addresses and building types but according to 77m the data provided did not include INSPIRE polygons. There is no reason to doubt that latter point. In December 2016 a similar version of Matrix was supplied to a different customer Constable Homes. The customer then downloaded the INSPIRE polygons themselves and these were integrated into the version of Matrix supplied to Constable.
98. As I understand it development of Matrix effectively stopped sometime in 2017/2018. Staff such as Dr Brown left, 77m says due at least in part to the pressures of this litigation and uncertainty caused by OS.

*How was Matrix created?*

99. Matrix was created by processing and combining data from a wide range of datasets including but not limited to some of the datasets referred to above. However explaining what was done in more detail is not easy. This dispute has been made harder to resolve because 77m has never been entirely frank about how Matrix was created. There are some grounds for sympathy in that the details are complex and highly technical, individuals have left, steps have not been well documented, and I do believe Mr Highland, who is not a computer programmer, genuinely does not know every minute detail of the entire story because much of the detailed work was done by others. Another problem is that 77m believes key parts of its method are secret and clever ways of effectively getting around restrictions OS seeks to place on access to the more accurate information it holds. OS is the last organisation 77m wishes to have access to some of its ideas. Also although some of the secrets are technically complex, the detail of them does not matter and so they can be protected without too much trouble. The practical problems are caused by some relatively simple ideas, which 77m says are secret, but which could be easily given away. Insufficient thought was given to this at the outset and it significantly complicated the materials in this case and the conduct of the trial. Overall however, despite making the various allowances above, the cause of most of the difficulty has been a lack of candour by 77m and, in at least some cases, deliberate falsehood (e.g. moving the geocoded points – see below).
100. The description which follows is my own, based on the expert evidence and in part on explanations provided by OS and 77m. Presented in this manner the steps described are intended to focus on matters relevant to this dispute rather than represent a comprehensive description of the whole process. The tasks were not necessarily undertaken in this order, and many of the steps were repeated as new or updated data became available. I think 77m does not agree with the way this description characterises certain sets of data. 77m says the data was grouped in different ways.



However despite counsel's best effort in closing, I was not left with a clear explanation of what is materially wrong with this description. Setting it out this way is intended to aid comprehension and I believe it is a fair approach to understanding Matrix.

101. It is also worth noting that the evidence does not make it possible to identify the dates when various of these steps were taken but that does not matter. Overall Matrix was created over the period from 2010/11 until 2017/18.
102. The steps are:
  - i) Process address data
  - ii) Process geospatial data
  - iii) Link address data and geospatial data together
  - iv) Generate more data associated with geolocated addresses

(i) Process address data

103. 77m has a file called the Master Address List. This was made by starting from the PAF. Then those addresses were merged with addresses from other freely available lists of addresses. Examples of other sources of addresses include the VOA, the list of sporting facilities from Sport England and the Church of England's list of churches.
104. The Master Address List data was cleansed (which for this purpose is also a form of verification) and standardised according to 77m's own standards. This exercise included making edits to data from the PAF. To achieve this 77m created its own proprietary address matching software. 77m also used address matching to obtain further information about the addresses in its Master Address List by matching them against datasets e.g. which recorded building type or listed building status. The cleansing/verification process is an ongoing one so that when new data is made available, it is verified against existing data.
105. Although the PAF address data was cleansed, the A1 property descriptions, which is also a source of addresses, were not used in that exercise.

(ii) Process geospatial data

106. 77m had a number of sources of geospatial data. The most important geospatial data was the INSPIRE polygons. Taken as a whole these polygons are effectively a map of England and Wales by freehold boundary. Although freehold boundaries do not necessarily always correspond to physical features, in practice they almost always do. However on its own, of course, the set of INSPIRE polygons does not tell you what the addresses are nor does it say anything about what is actually situated on the land (land use etc).
107. 77m took a number of steps to process its geospatial data.
108. 77m categorised all the INSPIRE polygons by what they represent. The polygons were put into three classes. One class was so called "garbage" polygons. These had

an area of less than 3m<sup>2</sup>. In effect they were no use to 77m. The remaining polygons were divided into “land only” polygons which had no buildings on them and polygons which intersect with a building, which were classed as “operational”. It seems clear, although nothing turns on it, that further categorisation work was done as well, e.g. to identify and eliminate polygons representing features which are under the ground, at least in London, and categorising different sorts of land use.

109. One way of doing this categorisation was to use the A1 property descriptions, which are indexed against INSPIRE IDs, in an attempt to identify non-addressable polygons in order to eliminate them. The text in the A1 property description might indicate that the site is non-addressable (e.g. “playing field” or “woodland”).
110. Another step 77m took was to create a good quality set of building polygons. The way this was done is confidential and is explained in the confidential annex.
111. Another step taken by 77m was to derive a set of five items of data associated with each INSPIRE polygon. The first two items of data are the centroid and the area of the polygon. In order to do this one needs to perform a calculation which uses all the X-Y coordinates which define the polygon geometry. These results were stored against the INSPIRE polygon.
112. Further calculations were done based on the INSPIRE polygon and the building polygons which are in the same location. They are confidential and set out in the confidential annex.

(iii) Link address data and geospatial data together

113. As a result of the previous steps, at this stage 77m now had three sets of processed data: a good quality Master Address List with some further information but no spatial or coordinate data linked to the addresses; a set of spatially derived information indexed to the INSPIRE polygons; and a set of building polygons with further information attached including some related to relevant INSPIRE polygons. These are “processed” data because 77m has done extensive processing to generate these datasets in this form.
114. At this stage the Master Address List and the information index to INSPIRE polygons were not linked together. The third set of processed data can be treated as linked to the INSPIRE polygons. The next task was to link the address data in the Master Address List with the geospatial data indexed by the INSPIRE polygons so as to end up with a comprehensive set of geolocations for all the addresses in the Master Address List. 77m did this using a technique involving “anchor points” and “streetwalking”. Essentially this involves two steps. The first step is to identify a set of what 77m calls “anchor points”. 77m uses a scoring system to record the quality of the geolocation for a given address. An anchor point is an address in the Master Address List for which 77m has an accurate geolocation, i.e. a sufficiently high score. Given those anchor points 77m can then carry out the second step, referred to as streetwalking. This second step used interpolation and inference to work out the remaining locations which are not anchor points.

*Anchor points*

115. 77m emphasised that many sources were used to determine accurate geolocations of an address in the Master Address List which did not involve the A1 property descriptions. I accept that. One example is Codepoint Open. This was used to assign a default location to the address by postcode. For reasons explained above, that location is not necessarily very accurate.
116. However as a result of the A1 matching exercise (and by use of the FAP website) one of the sets of data 77m had acquired was a list of A1 property descriptions indexed by INSPIRE ID. So if an address in the Master Address List can be found to match an address in an A1 property description, that gives you the INSPIRE ID corresponding to the address in the Master Address List and that in turn allows you to link the address in the Master Address List with the five items of good quality geospatial data referred to above.
117. Thus, as well as providing a way to identify non-addressable polygons, the A1 property descriptions from HMLR gave 77m a way of creating anchor points by address matching. It is plain that today the reason many millions of addresses in the Master Address List have accurate geolocations is as a result of this process of matching the address in the Master Address List with the address in the A1 property description.
118. An aspect of the matching exercise which 77m emphasised was that when a match was identified between the two addresses - in the Master Address and in the A1 property description, the address in the Master Address List did not change. In other words nothing from the A1 property description itself was taken into the Master Address List address. The A1 property description address was only used to find a match and then link the INSPIRE ID with the Master Address List. The reason 77m emphasised this was that the address in the Master Address List had been derived from the PAF – which was fully licensed from Royal Mail – and cleansed by processing by 77m in a manner free of any rights of OS. OS had a claim, which I will address below, to rights in the addresses in the A1 property descriptions because they were said to have derived from AddressBase or its predecessors and ultimately from the NAG. 77m’s case is that even if OS did have such rights, what was done was merely reading the information and did not infringe any rights of OS even if it was outside the terms on which HMLR had licensed the A1 property descriptions.

### *Streetwalking*

119. Given the anchor points, the remainder of the addresses in the Master Address List were assigned geospatial coordinates by streetwalking. This involved the computer identifying an anchor point – say a house in a street (1 Acacia Ave.) and another nearby anchor point (say 30 Acacia Ave.) and then notionally “walking” along the street between these two properties and filling in the blanks. In simple cases this is straightforward but in the real world it is often not simple at all. The details of 77m’s streetwalking algorithms are secret but irrelevant. It is common ground that there are innovative techniques here which are not publicly known in the industry.

### *Anchor points and streetwalking together*

120. Either as an anchor point or by streetwalking, the end result is that the individual addresses in the Master Address List are linked to the appropriate INSPIRE ID and

therefore to the geospatial data associated with that INSPIRE ID. That means that the good quality geospatial coordinate data in the five items of data stored against each INSPIRE ID is now linked to the address. One of those data is taken as the geospatial location of the address. I refer to it as the geocoded point. Exactly how it is produced can only be understood in the light of the information in the confidential annex.

121. The number of anchor points actually generated and used by 77m was never clear and inconsistent evidence was given about it. Also unclear was the minimum number of anchor points necessary to allow the dataset to be completed accurately by streetwalking.
122. As far as I know the numbers do not matter to any level of precision. The total number of addresses is about 30 million. The total number of A1 descriptions which 77m hold, indexed to INSPIRE IDs is about 16 million. I am satisfied that many millions, well over 50%, of the A1 property descriptions correspond to an anchor point. Mr Simmons' view was that 11.6 million anchor points were identified based on A1 property descriptions obtained under the A1 match licence (i.e. not via FAP). I accept that evidence.
123. The witnesses did not agree about how many anchor points were needed to reach the tipping point beyond which good results for the rest of the dataset can be provided by streetwalking. Mr Highland's view was that the number was about 6 million but Mr Simmons' view was that it was at least 9 million. If I had to decide on a figure I would find it was 9 million but irrespective of that issue, I am satisfied that even though there will be quite a number of anchor points which were not created using the A1 matching, nevertheless Matrix in its current form would not be a viable dataset of addresses and accurate geospatial coordinates without the A1 property descriptions obtained under the A1 match licence.

#### *Anchor points in Scotland*

124. Anchor points in Scotland were created using the land values data and the INSPIRE cadastral seed points. Essentially, 77m assigned locations to addresses in the Master Address List by address matching with the Land Values data and then using coordinates provided in the Land Values data to assign a location. It then used those coordinates to search for the nearest RoS seed point and then replaced the coordinates from the land values data with the coordinates of the seed point. About 600,000 anchor points were created this way.

#### (iv) Generate more data associated with geolocated addresses

125. The previous step (iii) was the most important one. After the full Master Address List had been geocoded further processing was done. This included deriving building heights in a manner which is confidential and irrelevant, and obtaining a greater set of land use classifications by applying queries to the INSPIRE polygons.

#### *Conclusion on the Matrix dataset*

126. This overall processing was not just undertaken once. Updates were provided and further runs of some or all of these steps were carried out. The result is that Matrix consists of a good quality Master Address List for which each address in the list has a

single geocoded point. Further information is also linked to the address, e.g. about land use and other geospatial information, such as the area of the relevant plot freehold. All of it is indexed by the Master Address List and associated geolocation.

*The shift in the geocoded point*

127. A further step 77m took was to introduce a very small, random, shift in the coordinates of the geocoded point for each address. The true purpose of this shift was to ensure that the geolocation coordinates for a given address in Matrix were not the same as the geolocation coordinates for the same address in OS's AddressBase dataset. 77m feared that if its coordinates were the same as the ones in AddressBase, OS would assume, wrongly, that AddressBase or its predecessor(s) had been stolen by 77m. In fact the coordinates were not copied (77m did not have access to AddressBase), rather they had been derived in the manner described above. However 77m thought OS would regard the close similarity of millions of coordinates as evidence of copying. As a matter of fact, as I understand it, without the random shift the coordinates would indeed be very similar or identical to the ones in AddressBase, for a very large number of addresses. OS did not suggest this indicated direct copying of coordinates. In my judgment the similarity is due to the way 77m derived the coordinates.
128. Mr Highland's evidence about this shift was not reliable. He suggested that the random shifts were introduced as a form of "Easter Egg", by which he meant a piece of information added to a database which would act as a tell-tale sign of copying if it appeared in a third party's data. It is true that 77m's random shift would have that effect but it was not the true reason why it was done. The true reason was the one described above. It was not mentioned until he was cross-examined at trial. The reason was because Mr Highland did not trust OS but that mistrust led him to give evidence deliberately which gave a false impression.

*The issues*

129. The questions to be decided can now be understood in context.
130. OS does not maintain a claim based on copyright but contends that 77m has infringed OS database rights in three respects:
- i) Infringement of the database right in the Topo database (also called MAIA in the pleading) by the use of INSPIRE polygons;
  - ii) Infringement of the database right in the Addressing databases (that is AddressBase and its predecessors AddressPoint, AddressLayer and AddressLayer2) by:
    - a) the use of centroids from the RoS Land Values dataset; and
    - b) the use of addresses from HMLR and Lichfield D.C.
131. Although there were some references in the pleadings to addresses in the RoS Land Values dataset, no claim to infringement of database right has been made relating to

that data. The database right infringement claim relating to RoS Land Values dataset is about centroids (OS pleading para 107f).

132. The first set of issues to deal with relates to the meaning and scope of various contractual licences entered into between 77m and HMLR (the INSPIRE Download Terms, the A1 Match Licence and the FAP terms) or RoS (the RoS Land Values Licence). The argument about the meaning of the INSPIRE Download Terms is important because those terms represent the licence given to 77m relating to the INSPIRE polygons, which are at the heart of what 77m did. The argument about the terms of the A1 Match Licence is important because those terms represent the licence given to 77m relating to most (but not all) of the A1 property descriptions held by 77m. The argument about the FAP terms relates to the issue of scraping and the outcome determines the relevant licence terms (if any) governing several million further A1 property descriptions. The argument about the RoS Land Values licence also involves looking a bit more closely at what 77m actually did with the relevant data.
133. Once the scope of the licences is determined one can see whether 77m's actions fell within their terms or not. Even if the actions are within the licences there are also issues about alleged limits on the power of the bodies granting these licences (HMLR, Lichfield DC and RoS). OS contended that if the licences were so broad as to render lawful what 77m did then HMLR and RoS had no power to grant them and therefore they were not effective in law to grant such a licence.
134. Once all that has been decided one can address whether 77m has infringed OS's database rights. At trial 77m took a further point, asking whether the acts which were carried out needed a licence at all. To deal with that one needs to identify exactly what rights OS has and then examine the scope of the definition of infringement in database rights law and the scope of various defences. OS did not accept that these points (or some of them) were open to 77m and I need to consider that.
135. 77m also maintained separate defences of estoppel by representation.
136. In relation to Lichfield DC, a host of issues were raised. The number of records obtained from Lichfield DC was very small compared to the numbers from the other sources with which this case is concerned and I doubt the numerous issues unique to Lichfield are worth the candle. Nevertheless the points were maintained. In closing 77m submitted that OS had not established that any geocodes from Lichfield DC had been transferred into Matrix. I agree but the issue is about addresses, not geocodes.
137. Finally there is the procuring breach of contract issue.

*Construction of the various licences*

138. The relevant principles of contractual interpretation are well established – see ***Rainy Sky SA v Kookmin Bank*** [2011] 1 WLR 2900; ***Arnold v Britton*** [2015] AC 1619; ***Wood v Capita Insurance Services Ltd*** [2017] UKSC 24. There is no need to restate them here. I would only add that where an agreement reflects a public standard form contract, factual evidence regarding the circumstances surrounding an individual instance of that contract will be of limited, if any, importance: *Lewison* 5<sup>th</sup> Ed at [3.18] and *Chitty* 33<sup>rd</sup> Ed. at [13-051].

*The INSPIRE Download Terms*

139. The INSPIRE Download Terms are set out in a guidance document accessible from the relevant part of the gov.uk website concerned with downloading the INSPIRE polygons. The document has various sections, dealing with how to download the polygons and the metadata. The next section is headed “About INSPIRE polygons” and starts with the following sentence:

“INSPIRE Index Polygons is an open source dataset, developed to comply with the INSPIRE Directive (*[hyperlink ref]*). It contains the locations of freehold registered property in England and Wales and a sub-set of our Index Polygons for all freehold land and property.”

140. The same section contains some other descriptive information, including a reference to each polygon having a unique INSPIRE-ID. There is then a link to the Land Registry’s FAP service, explaining what that will give the user, as follows:

“Use a Land Registry-INSPIRE ID with our Find-a-Property service (*[hyperlink ref]*) to get the title registration and plan for each polygon.”

141. After some technical guidance about the “open data format” which the data will be supplied in, the guidance document has a section headed “Conditions of use”. This section is worth setting out in full. It is as follows:

“Your use of the INSPIRE Index Polygons service is governed by conditions.

The INSPIRE Index Polygons and attributes provided in this service are available for use and reuse under the Open Government Licence (OGL) (*[hyperlink ref]*). This licence enables public bodies to make their data available free of charge for reuse.

Use under the OGL is free. If you fail to comply with any of the conditions of the OGL then the rights granted to you under the licence will automatically end.

Under the OGL when reusing data you must acknowledge the source of the data and include an attribution statement. You must:

1. Display the following statement ‘This information is subject to Crown copyright and is reproduced with the permission of the land registry.’
2. In addition when reusing the polygons (including the associated geometry, namely x,y coordinates) display the following Ordnance Survey copyright/ database right notice

‘© Crown copyright and database rights [year of supply or date of publication] Ordnance Survey 100026316’

3. Where possible you must provide a link to these conditions.

Under the OGL, Land Registry permits you to use the data for commercial or non-commercial purposes. However, as the licence says, OGL does not cover the use of third party rights which we are not authorised to license. Land Registry uses Ordnance Survey data in the preparation of the polygons. Therefore you should contact Ordnance Survey for the relevant licence conditions if you need to:

1. use the polygons (including the associated geometry, namely x,y co-ordinates) for a purpose other than personal, non-commercial use or commercial or non-commercial use within your organisation; or
2. sub-license, distribute, sell or make available the polygons (including the associated geometry, namely x,y co-ordinates) to third parties.

For Ordnance Survey’s licence conditions, contact the Ordnance Survey (*[hyperlink ref]*)”

142. The OS hyperlink at the end is to the Ordnance Survey in general rather than any specific licence conditions. The INSPIRE Download Terms document has a short section after these “Conditions of use” which provides some further information, none of which has a bearing on the issues to be decided.
143. In terms of the overall structure of these conditions, the starting point is that the data is available on wide terms but there are exceptions. The wide opening language indicates that the data is free to be used in any way and is free of charge. That is based on the terms of the OGL itself. Those terms are very wide and make clear that users are permitted to publish and transmit the information freely and to exploit the information covered by it “commercially and non-commercially”. That expressly includes combining it with other information or “including it in your own product or application”.
144. Nevertheless despite the width of these terms, the clauses also provide for exceptions. One exception in the OGL itself is that it does not license third party rights. There are two major sets of exceptions expressed in the INSPIRE Download Terms themselves. One set is the acknowledgement terms which form numbered paragraphs 1 to 3. They are not significant. The second set of exceptions form numbered paragraphs 1 and 2 with the preamble. These are the critical ones. It is clear that they are related to the rights of Ordnance Survey. It is not controversial that they define a set of acts for which the INSPIRE Download Terms provide that no licence is being given and so, if the user wishes to undertake any of them, they need to contact Ordnance Survey. The debate is about the scope of those expressly unlicensed acts.

*Relevant factual matrix*



145. Before going further I turn to deal with the factual matrix.
146. OS contended that there were two key elements of the factual matrix which have a bearing on the true construction of the INSPIRE Download Terms. One element is the licensing structure OS already uses. The other element is based on the submission that “internal business use” was a term of art in the industry.
147. As to the first element, OS suggested 77m had admitted this on the pleadings but in my judgment the admission (paragraph 28.1 of the Reply and Defence to Counterclaim) is qualified in a manner which does not assist OS.
148. The argument from OS amounts to a submission that buried in the terms of the licences OS grants to public sector organisations – the PSMA and OSMA – there are express clauses restricting competing and commercial activities which are in wide terms. In the PSMA this is clause 13 of Appendix 1 (and see definition clauses 2.1.1 and 2.1.2). Commercial activity means any activity for financial gain and competing activity includes anything which competes with a product or service of Ordnance Survey. The argument is therefore that armed with this knowledge one can construe the exceptions in the INSPIRE Download Terms to mean that those sorts of activities are excluded.
149. I do not accept that submission for three reasons. These INSPIRE Download Terms are in a standard form addressed to the world at large, although I am prepared to accept that the only people genuinely interested in thinking about these exceptions in INSPIRE Download Terms are those working in geospatial data in general. I am also prepared to accept that such people will have a broad understanding of OS licensing, including the existence of the PSMA. Nevertheless I do not accept that the level of knowledge of the specifics of OS public sector licensing arrangements necessary to make this point could properly be deployed as a relevant part of the factual matrix for the interpretation of this contract. Secondly, even if a putative downloader of the INSPIRE polygons who was aware of the details of the PSMA paused to wonder if it shed any light, the fact is that the Download Terms make no reference to the PSMA (or OSMA). Even if that putative downloader went on to consider whether the terms were consistent or inconsistent, they would conclude that whatever was in the PSMA was not determinative. What matters are the terms of the exceptions themselves.
150. Thirdly, the reader would assume OS agreed with HMLR about the text of the INSPIRE Download Terms (as indeed it did). That would mean that it would not matter even if the terms were not aligned completely with some other OS licence.
151. The second item of factual matrix is about “internal business use”. OS said this expression referred to the use of data for the internal administration and operation of the licensee’s business as opposed to use for the creation of products / services for licensing or supply to third parties (whether on a commercial basis or otherwise). So the argument goes, someone versed with that understanding of “internal business use” as a concept would identify that even though the term does not appear expressly in the INSPIRE Download Terms, that concept is what the exceptions are referring to and they should be construed accordingly.
152. OS relied on the evidence of Mr Cutler to make that good. His evidence was that the understanding of those in the industry was “*party A cannot use party B’s data to*

*create a new product which is then sold without royalties being paid to party B*". His view was that that is what the concept of internal business use is about. He maintained his view in cross-examination and did not agree with counsel for 77m that the knowledge was confined to a narrow group of OS partners. I accept Mr Cutler's evidence that there is a widely held belief in those working in the geospatial data industry along the lines he described. However I am far from convinced that sheds useful light on the interpretation of the INSPIRE Download Terms. First, those terms do not use the expression "internal business use". Second, the generally held understanding in the industry means that a person would not be surprised to read a contract which contained a provision with a meaning which fitted with their understanding. However, equally well, if they saw terms which appeared to make a different provision, there is no evidence their understanding was so strong that it would displace that different meaning. Moreover 77m is right to make the point that such a person might well conclude that the INSPIRE Download Terms are drafted the way they are in order to comply with the Government's INSPIRE obligations, in which case the general understanding would not be relevant.

153. OS also contended that 77m in particular was well aware of the meaning of "internal business use". That is true (and Mr Highland accepted Mr Cutler's evidence). However since this is a contract on standard terms I do not accept that is relevant to its construction.
154. 77m also relied on a number of factors said to be part of the factual matrix -such as the existence of other possible licences of government data (not the OGL) which could have been used. I was not persuaded any of that helped either.

*The rival constructions*

155. Both sides start with the terms of the OGL. OS submits, and I accept, that in the OGL and therefore in the Download Terms:
  - (a) 'Information' means information protected by copyright or database right (for example, literary and artistic works, content, data and source code) offered for use under the terms of the OGL.
  - (b) 'Information Provider' means the person or organisation providing the Information under this licence. Accordingly, in the Download Terms, the Information Provider was HMLR.
  - (c) 'Use' means doing any act which is restricted by copyright or database right, whether in the original medium or in any other medium, and includes without limitation, distributing, copying, adapting, modifying as may be technically necessary to use it in a different mode or format.
156. Turning to the disputed clause, 77m's case is that it would be seen as a specific and therefore limited exception or carve out from the general permission to use the data. The terms do not expressly say anything about prohibiting the use of data derived from the polygons in a product or service. Instead the words focus on preventing use of the polygons outside the organisation and making available the polygons themselves (included associated geometry). That is consistent with a prohibition on resale of the polygons but says nothing about sale of a product or service based on the

polygons but which does not include them. The reference to “commercial or non-commercial use within your organisation” is effectively tautologous but would be understood as emphasising that internal use is not restricted even if it is commercial. So 77m contends that while it obviously did use the INSPIRE polygons for commercial purposes, its use of them was purely internal and it did not and does not supply the polygons to third parties. It says that its position is that the licence “means what it says”.

157. A separate construction issue relates to the ambit of the term “associated geometry, namely x,y coordinates”. The question is whether the INSPIRE polygon centroid, area and the three items of geospatial data calculated by 77m fall within that term. It matters because of course 77m does (or would) supply that data.
158. OS contends that it is convenient for the purpose of analysis to split the disputed part into three conditions. If any one of those conditions apply then the user has to contact OS – and necessarily has no licence under the INSPIRE Download Terms. They are:
  - i) if the user needs to use the polygons (including the associated geometry, namely x,y co-ordinates) “for a purpose other than personal, non-commercial use” (“Condition 1”).
  - ii) if the user needs to use the polygons (including the associated geometry, namely x,y co-ordinates) “for a purpose other than... commercial or non-commercial use within your organisation” (“Condition 2”).
  - iii) if the user needs to sub-license, distribute, sell or make available the polygons (including the associated geometry, namely x,y co-ordinates) to third parties (“Condition 3”).
159. OS says that Condition 1 is straightforward in that, turning it on its head, it means that no further licence from OS is required if the user merely needs to use the polygons for personal, or non-commercial uses. OS’s point is that this does not apply here.
160. The dispute is about condition 2. OS contends that these words prohibit the use of the relevant data (polygons etc.) for the creation of a product that is then sold or distributed to third parties. In effect this is the condition which corresponds to the concept of “internal business use” said to be a term of art.
161. OS argues that the clause requires one to look at the purpose of the use and then, if the purpose of the use is something outside the organisation, then it is not permitted. This is one of OS’s major points and, it contends, 77m’s construction ignores this aspect of the way the clause is written.
162. OS also contends that its interpretation makes sense of the INSPIRE Download Terms as a whole because if the clause permits exploitation of OS rights both commercially and non-commercially, both internally and externally then there would never be a need to seek a separate licence from OS, and its construction accords with business common sense because it ensures OS’s investment is protected and there is no conflict with the exploitation that OS or any OS Licensed Partner makes of OS licensed data.

163. Condition 3 prohibits the polygons (etc.) from being sub-licensed, distributed, sold or made available to third parties. Both sides agree that this means polygons and relevant associated geometry cannot be distributed to third parties. The issue is the scope of the term defining associated geometry.
164. I will start by putting to one side the question about the meaning of associated geometry. The clause means the same thing whether it is concerned with polygons themselves or associated geometry. I agree with 77m that it is relevant to see the context as one in which the INSPIRE polygons are being made available for use and that this clause is an exception. I do not believe OS's points about consistency with the terms as a whole or business common sense are significant. On both constructions the clause fits into the INSPIRE Download Terms as a whole and makes business sense.
165. Next, a significant aspect of this clause is that some sort of commercial activity does not need a further licence from OS. This does not prohibit commercial organisations from downloading the polygons and using them internally. Quite the opposite.
166. I do not agree with OS that 77m's approach gives no meaning to the idea of purpose. The clause provides that no further OS licence is needed if you use the polygons for "a purpose" which is personal or non-commercial. It also provides that no further OS licence is needed if you use the polygons for "a purpose" which is "commercial or non-commercial use within your organisation". If that is the purpose for which you use the polygons then no further OS licence is needed. It may not be the purpose for which the user uses other things, but if that is all the polygons are used for then that use is licensed. Otherwise the clause has no meaning.
167. The question is what to do about a case in which polygons are used internally but some information derived from that use is itself used in a product or service which is sold externally. Mr Jagot (of OS) gave an example of a financial services company using the data to assess flood risk and determine a premium to be paid. The point of the example was to illustrate a case of "commercial use within an organisation" of the INSPIRE data and to put it forward as something permitted by the INSPIRE Download Terms. In cross-examination it emerged that Mr Jagot distinguished between a situation in which what the company tells the customer is just the price of the insurance based on its assessment of the flood risk (and a value for that risk) and a case in which the company also says words to the effect of "we have calculated this risk because the property is X yards from the river". Mr Jagot's view was that the former was acceptable and within the scope of the licence given by the Download Terms but the latter was not, because it involved the company giving out geographic data to a third party – i.e. the customer.
168. Now of course the exercise of construction of the clause is an objective one and Mr Jagot's views are not determinative. All the same it seems to me that this example does allow one to see how the terms work. I find that the terms do not prevent a licensee company from using the polygons or the associated geometry for any purpose at all and that includes developing products or services for resale to customers, always provided that the licensee has no licence to supply the polygons or associated geometry to a third party, such as a customer. The licensee company is entitled to supply any information (or a product) created using the polygons or the associated geometry as long as that supply does not involve giving the polygons or associated

geometry to a third party. So, looking at Mr Jagot's example, I would accept it as a correct example of the ambit of the terms if the piece of information "*the property is X yards from the river*" either is a polygon (which it plainly is not), or is the associated geometry (which I address below), but not otherwise.

169. There is no other qualification on what a licensee company may or may not do. Mr Jagot did not see the terms in that way. His evidence was that the terms prevented a licensee from using the data to develop products and services for resale. Whether or not OS's "internal business use" concept achieves that does not matter. Mr Jagot is wrong insofar as he is referring to the effect of the INSPIRE Download Terms.
170. Other examples were addressed in argument but it is not necessary to deal with them.
171. 77m has used the INSPIRE polygons internally to create Matrix. I find that in doing this 77m was operating within the INSPIRE Download Terms.
172. 77m's case is that it does not and does not intend to provide INSPIRE polygons to customers of Matrix. However it does wish to provide the five items of data associated with an INSPIRE polygon. The question is whether these are within the scope of the associated geometry. If they are then 77m is not licensed to provide that data to customers.
173. The operative words are "the polygons (including the associated geometry, namely x,y co-ordinates)". 77m pointed out that when polygons are provided, the manner in which is done is by providing a series of x,y co-ordinates which specify the vertices of the polygon and therefore define the polygon itself. I believe that is what a reasonable objective person knowing the relevant factual matrix would understand the clause to be referring to. Mr Jagot's view seems to have been that this covers any geographic data derived from the polygons. I do not accept that. It is far too wide. An intermediate position was the submission that geometric information like a polygon centroid (which is after all an x,y co-ordinate) and can be said in a general sense to be geometric information associated with the polygon, fell within the clause. I do not agree with that either, for two reasons. First the words themselves are apt to refer to the existing x,y co-ordinates which are already associated with the polygons when those polygons are provided. There is no hint of a reference to derived data. Second, at its heart what is being protected is the polygons. Derived data, whether it happens to be a centroid with an x,y co-ordinate or other geometric information such as the area of the polygon, does not allow the recipient to recreate the polygon. It does not give it away.
174. The same conclusion applies to the other three, confidential, items of geospatial data 77m store with the INSPIRE polygon.
175. Therefore for the INSPIRE Download Terms I reject OS's case and find in favour of 77m. 77m's use of the INSPIRE polygons has been authorised by HMLR.

#### *The A1 Match licence*

176. The general circumstances of the A1 Match licence were summarised above. 77m contends that the relevant factual matrix includes the following particular points:

- i) The parties' mutual understanding of 77m's commercial need for the A1 Match Licence. This included a need to cleanse non-addressable sites from the INSPIRE polygon datasets.
  - ii) HMLR was well aware that 77m wanted to link the INSPIRE IDs to addresses.
  - iii) HMLR knew that a large number of records needed to be matched, a figure of 3.5 million records was mentioned. It was the parties' common intention that it would be preferable to request them in stages rather than all at once. There was also mention of records changing on a monthly cycle and a question whether daily transactions would be available in the future.
177. The significance of these submissions is twofold. One aspect relates to the scope of the licence HMLR granted to 77m. The question is what did 77m tell HMLR about why 77m wanted the data. OS submits that HMLR only understood that 77m wanted to "cleanse garbage polygons". That is much narrower than what 77m actually used the A1 Match data for. 77m's case is that HMLR's understanding was much wider (or at least ought to have been, given what 77m said). The second aspect relates to the procuring breach case. 77m argues that the supply was to be on a more or less continuous basis and that HMLR knew and accepted that. OS does not agree. Both aspects are matters of construction of the licence but 77m relies on the context as part of its case on construction.
178. On the first aspect I prefer the evidence of Ms Wiles and Ms Nicholson to Mr Highland that what 77m led HMLR to believe was that the purpose of the matching was for suppressing and removing irrelevant and potentially incorrect information. This can be summarised as cleansing garbage polygons and as removing non-addressables. Although strictly speaking non-addressable sites and garbage polygons are distinct, the difference does not matter. These expressions convey the same sense of what HMLR was told. Moreover they convey a materially different concept from the idea of using the address matching to link INSPIRE polygons to the addressable locations in 77m's data and then go further and associate high quality geospatial information derived from the INSPIRE polygons with those addresses. 77m never told HMLR that that was something they intended to do (or had done).
179. It was not false to say that 77m wanted to use the results of the A1 matching process in order to cleanse garbage polygons and/or remove non-addressables. I accept that that is something 77m did intend to do and did in fact do. However it was not all that was done. I do not believe it was an accident that 77m as an organisation and Mr Highland in particular did not tell HMLR the whole story.
180. It is true, as was put in cross-examination on 77m's behalf, that creating a link is a consequence of matching. But I accept Ms Wiles' evidence that that was not discussed with her. I also accept Mr Simmons' evidence that garbage records can be identified without any need to link the addressable records (my emphasis) within the A1 Match data to 77m's full Master Address List.
181. I find that the only mutual understanding of 77m's commercial need for the A1 Match Licence which the parties had was to cleanse garbage polygons/non-addressable sites from the INSPIRE polygon datasets. HMLR did not know that 77m had any other purpose in linking INSPIRE IDs to addresses.

182. As for the second aspect, OS contended that Ms Wiles gave clear evidence that she and Mr Highland discussed a one off bespoke service before entering into the contract. I do not accept the position is that simple. It is true that in internal emails Ms Wiles refers to the arrangement as a one off bespoke arrangement but those emails were not sent to Mr Highland. I find that HMLR did not tell 77m that HMLR had decided to treat the service as a one off for their own internal purposes (which is what happened, at least at the start). HMLR knew, because of what Mr Highland told Ms Wiles, that more sites were likely to be identified as needing to be matched over time. 77m certainly never told HMLR that there was only ever going to be one request. One conversation was about keywords and Ms Wiles knew and understood its relevance. The point is that a keyword (such as “village green”) was being used by 77m to identify a non-addressable site and so if more keywords were to be used in future (which is what Ms Wiles was told) more matching would be required (which is what Ms Wiles understood).
183. Turning to the A1 Match licence itself, the contract was dated 13<sup>th</sup> February 2014. The contract is a bespoke document agreed between Mr Highland and Ms Nicholson. It is headed Contract Schedule. Clause 1 refers to and incorporates HMLR’s standard terms on its website. 77m says that the standard terms are significant, because they define “Services” and then provide the framework within which the Services of the agreement will be performed. The Services are “*The service or services that we supply to you as set out in the Contract Schedule or Schedules or any additional services required by you from time to time.*” 77m contends that this supports its case about the ongoing nature of the agreement and means that the scope of the Services includes further requests made by 77m under the A1 Match Licence in addition to those set out in the contract schedule.
184. The “Term” is defined in clause 1 of HMLR’s standard terms as “*The period during which we agree to provide each Service. Unless otherwise specified to the contrary in the Contract Schedule[s] the term of each Service shall be ongoing unless terminated sooner by either party in accordance with these terms and conditions ...*”. 77m argues that the A1 Match Licence itself does not specify “to the contrary” and that this clause also indicates that the “Services” were to be ongoing and of indefinite duration.
185. 77m also referred to clause 3 of the standard terms which deals with variations and additional services, providing that if the customer wants HMLR to provide additional services to those already agreed then supplemental terms will be agreed and put into a further contract schedule. Since no other relevant contract schedule was entered into despite the further supplies of A1 Match data after the February 2014 contract, 77m contends that this shows that no new or different services were regarded as being supplied by HMLR.
186. No other parts of the standard terms are relied on.
187. Clause 4 of the A1 Match Licence itself provides:
- “In consideration of You paying to us the Price, we will provide You with the Services on a continuous basis unless terminated sooner by either party in accordance with the Terms and conditions.”

188. Clause 6 is entitled “Service Description” and states:

“We will return to you a Response File (in Microsoft Excel format), identifying for all the INSPIRE IDs within the Customer File the property descriptions, as extracted from the register.”

189. The terms “Customer File” and “Response File” are defined in clause 5. The effect of clauses 4 and 6 is that the Service consists of the provision of a Response File processed against the Customer File, in response to a request from 77M.

190. Clause 6.4 (there are no numbered clauses 6.1 to 6.3) is entitled “Pricing” and states:

“The Bespoke price is volume based; £2,500 plus VAT.”

191. Clause 7 deals with encryption and states that “*it may be necessary to vary the form or method of encryption during the Term. We will endeavour to provide reasonable notice of any proposed changes*”.

192. Clause 8.1 grants a “non-exclusive, non-transferable, revocable, perpetual” licence to use the information supplied by HMLR. Clause 8.2 provides for termination and revocation of the licence for breach. There are also termination clauses in the standard terms but nothing turns on them.

193. Clause 9 sets out the scope of the “Permitted use” as follows:

“You have a business requirement to verify the INSPIRE data held on your internal systems and to confirm whether 910,000 INSPIRE IDs relate to non addressable sites. The Information will be limited to this use and you should confirm destruction of the data following completion of your cleansing process.”

194. Clause 11 and 11.4 (there are no clauses 11.2 to 11.3) are:

“11. We will Not:

Warrant that the Information provided on the Response File will be fit for your particular purpose nor do we warrant the completeness or accuracy or error free nature of our delivery of any Information on the Response File

You undertake:

11.4 Not to copy, sell, distribute, send or make use of the Information provided on the Response File (or any other information we provide You with as part of our delivery of the Full Service) other than for the Permitted use.”

195. There are no other relevant clauses.

196. 77m contends that on its true construction the contract between HMLR and 77m provided that the services provided by HMLR were continuous. That explains why,



after the matching service for the 910,000 INSPIRE IDs referred to on the contract were provided, 77m made a further five requests for the service, sending a total of about 7 million more INSPIRE IDs to HMLR, and HMLR responded by providing Response Files and charged fees for each request broadly in proportion to the “volume based” fee in the A1 Match licence document. No other contract documents were signed. Thus, contends 77m, it was a breach of that term when HMLR failed to provide a matching service in response to 77m’s seventh request made in August 2015.

197. OS suggested that an example put to Ms Wiles in cross-examination showed an alternative purpose of clause 4, which did not involve future supplies. This was said to be related to sorting out problems if the Response File from HMLR was jumbled so that 77M had to repeat the request. I cannot read that clause in that way. Clause 4 of the A1 Match Licence, when read in the context of the documents as a whole and the circumstances leading up to its signing, only makes sense as a reference to a continuing relationship into the future whereby 77m will provide more INSPIRE IDs and HMLR will respond with matching A1 property descriptions. No specific price was agreed but pricing clause 6.4 shows that the parties were in agreement that the volume would determine the price. The terms of the encryption clause 7 also only make sense if further supplies from 77m (i.e. lists of INSPIRE IDs called Customer Files) were to be made in future. The standard terms also lend some support to 77m’s case although they are not determinative on their own. The circumstances leading up to the signing of the agreement did not include anything by way of context which would rule out such a construction. If anything, they support it.
198. The fact the licence granted to use the information is perpetual (clause 8) is nothing to do with whether HMLR are accepting an obligation to provide further information in future.
199. I find that the terms of the A1 Match Licence did oblige HMLR to provide its match service in response to future requests from 77m. It follows that HMLR’s refusal to supply a Response File in response to the seventh request was a breach of that contract.
200. The scope of the licence granted is defined by the combined effect of clauses 9 and 11.4. 77m contends that clause 9 refers to two distinct uses, first verifying the INSPIRE data held on 77m’s internal systems and second confirming whether 910,000 INSPIRE IDs relate to non-addressable sites. OS contends that the clause is referring to a single use for the Information, not two uses, and that that single use is the same as the one made known to HMLR in advance – i.e. cleanse garbage polygons/non-addressable sites from INSPIRE polygon data held by 77m.
201. I prefer OS’s construction for two reasons. First, read literally, verifying INSPIRE data does not mean anything on its own. 77m has already acquired INSPIRE data from HMLR (and may acquire more in future if more becomes available) but it does not need to be verified in the abstract. On the other hand for 77m to provide a list of INSPIRE IDs which it thinks are or may be non-addressable sites and for the A1 property descriptions linked to those INSPIRE IDs then to be given, that would allow 77m to verify that it was right that they were indeed non-addressable sites. In other words the terms “to verify” and “to confirm” are duplicative. Second, bearing in

mind the context, reading the clause as referring to a single use makes sense since HMLR had only been told about a single use for the data.

202. I reject 77m's case on the scope of Permitted Use.
203. 77m's use of the A1 property descriptions is explained in the section describing Matrix. 77m did use the data to identify non-addressable sites/garbage polygons but that is not all that was done. It used the link between the INSPIRE ID and address given in the A1 property description for addressable sites (my emphasis) as a way of specifying the geospatial coordinates of its own list of addresses in the Master Address List. It did this to create its anchor points. This activity was not permitted by the Permitted Use.
204. It is worth noting that the INSPIRE IDs provided to HMLR by 77m over the period in which the A1 matching took place were not picked at random by 77m. The evidence about the reasons for picking INSPIRE IDs was not clear. I am not sure it matters but in case it does, I find that 77m did select some IDs on the basis that they may have been non-addressable or garbage polygons but also selected others for other reasons. One reason for selecting some IDs was to help with the problem caused by addresses in a given postcode which were only building names and not numbers. These cause difficulties for streetwalking because they have no inherent order. Mr Highland accepted this was a reason for selecting IDs. Its significance is that these are being selected to obtain a good geolocation for addressable sites, not for cleansing non-addressable sites.
205. I also reject what I understand to be a further aspect of 77m's case namely that all the licence does is give permission to carry out certain acts such that if the licensee does acts wider than the scope of the permission granted, there are no consequences as far as the contract is concerned. That is not right in terms of clauses 9 and 11.4. They make clear that as a matter of contract 77m would commit a breach of the contract if it did something wider than the permitted use. Therefore I find that 77m has breached the A1 Match licence.
206. There was also a point on deleting the A1 Match data. Clause 9 refers to it. 77m gave two reasons why it has not yet deleted the A1 property descriptions. One was because of the existence of these proceedings. The other was described in 77m's skeleton argument (paragraph 7.36) as because "*the cleansing process is not yet complete, primarily because OS intervened to instruct HMLR to stop supplying Response Files, so that 77M has at least one outstanding request for data*". There is some truth in this insofar as it is linked to the failure by HMLR to answer the seventh request but it is not accurate to call the process "cleansing" for the reasons already given.
207. It is worth recalling that this action is between OS and 77m, not HMLR and 77m. Breach of the A1 Match licence by HMLR is relevant to the procuring claim. Enforcement of the A1 Match Licence against 77m is a matter for HMLR not OS.

#### *FAP licences and scraping of A1 property descriptions*

208. It is not in dispute that 77m used HMLR's FAP service to obtain addresses for some INSPIRE IDs. There are apparently four issues to be decided relating to HMLR's

FAP service but each side characterised them in different ways. The first issue is what terms apply to FAP. The second is what is the scope of whichever terms apply. The third is a factual question, whether about 3.5 million records were obtained by automated scraping. There seemed to be an argument about the effect of the way in which HMLR invited 77m to use the FAP service but as I understand 77m's case, it does not say that the invitations purported to give 77m a licence different in scope from the terms to be decided on under the first issue.

209. As a preface, 77m is right that HMLR did draw attention to the FAP service as being something to use given an INSPIRE ID. The same INSPIRE guidance document which contains the INSPIRE Download Terms expressly does so.

*What terms apply?*

210. The FAP service works through a public website. A user goes to the appropriate webpage and is presented with boxes to fill in in order to do searches. One term which can be entered is an INSPIRE ID. When that is entered the system returns an address, which is the A1 property description. At that point a user can purchase a copy of the title register, title plan or certain other things. In order to get those further things a user has to register. Once they do, they can purchase what they want. Thus a user can in fact obtain an address in return for an INSPIRE ID without registering. That is what 77m did. 77m calls the part in which one can get an address without registering Stage 1 and the second part after registration Stage 2.
211. 77m's case is that the terms applicable to stage 1 are different from those applicable to stage 2. The point arises because if you look carefully, it can be seen that the relevant webpage contains two different cross-references to what are in fact two different sets of terms. One might have thought the situation was tolerably clear but it is not.
212. Neither side drew my attention to any authority on this but in principle it seems to me that the way a question like this should be approached must be to consider objectively the point of view of a reasonable user coming to the website to use it. That is because these licence terms are not open for negotiation. They are standard terms which the operator of the website is putting forward to all users.
213. The position is that when a user accesses the webpages, in particular at Stage 1, there is a prominent menu which includes an option for Terms & Conditions. Those are the terms which, at first sight, an objective user of the service would think were applicable to the FAP service. In fact this option links to terms specific to the FAP service, which varied over time. The link to other terms relied on by 77m is at the bottom of the page in smaller type and with much less prominence. It is in fact a link to the OGL.
214. There is a critical difference between the terms specific to the FAP and the terms of the OGL. The FAP expressly prohibits using automated software agents to access the system (i.e. prohibits scraping) whereas the OGL does not. There was a second issue about whether the terms applicable at the time permitted commercial use of the information or not. 77m maintained that while the version in place in May 2016 contained a reference to private and domestic use, the versions at the relevant times (2011 onwards) were not so limited in any event. I do not understand OS now to

contend otherwise. That has the consequence that the fact that the 480,000 addresses manually downloaded from FAP were used for commercial purposes was within whichever terms were applied by HMLR to those addresses.

215. Turning to the issue of which sets of terms are applicable, a user would be more likely to see the prominent link rather than the small reference at the bottom. Furthermore I doubt anyone would find the link at the bottom without seeing the other much more prominent link, and if a user found both, they would see that the OGL terms were generic while the prominent terms were specific to the FAP service. Absent anything else, these points would favour OS's case.
216. However 77m is right that the FAP specific terms are drafted in such a way that really only makes sense in the context of Stage 2. For example clause 3 of the 2016 terms provides:
- "In order to obtain services through Land Registry's Find a Property service you will need to register for the services. You must register via our 'Find a Property' website through the 'Find a Property' login."
217. There are other similar examples. Despite clause 3, the FAP website does not in fact require a user to register before they can use the service at Stage 1.
218. A further point arises from the requirements for how the user signifies that they agree. The terms state immediately under the heading: "*To agree these terms click "Agree". If you do not agree these terms do not click "Agree" and do not use the Find a Property Service*". The only context in which an "accept" button is displayed is when a user proceeded to register for an account in order to use the FAP Stage 2 service.
219. The FAP specific terms obviously apply to the service provided to a registered user. If there was no alternative set of terms and conditions, I believe that despite real uncertainty caused by the drafting and despite the fact that the system in fact provides information to a user without them registering (i.e. Stage 1), an objective user would conclude that the Stage 1 information was also provided subject to the restrictions in the FAP terms. They would assume that some terms and conditions applied and that would have to be the FAP specific terms.
220. The OGL terms, which are not inconsistent with what happens at Stage 1, are available as an alternative. An objective user presented with this website might conclude that the OGL terms provided the answer to the puzzling inconsistency between the FAP specific terms and the way the website actually operates. For one thing, why else have them available? Moreover, from the user's point of view, both the website itself and the FAP specific terms are entirely under HMLR's control. If HMLR wanted to prevent Stage 1 being used without clicking on an "accept" button and without registration it could easily have done so.
221. Nevertheless, standing back and looking at the circumstances as a whole, I believe that objectively, a reasonable user would conclude that what HMLR intended was for the FAP specific terms to be the terms of the licence applicable to the FAP however it was accessed, at Stage 1 or Stage 2. The terms are clearly addressed to the FAP service in particular, unlike the OGL. A reasonable user using Stage 1 of the FAP

service would not think they were obtaining information from the FAP service provided by HMLR free of the licence in HMLR's FAP terms and conditions.

*What is the scope of the applicable terms?*

222. As I have said, the only point seems to be that while the OGL did not prohibit scraping, the FAP specific terms do.

*Was FAP data scraped?*

223. As explained above, in addition to the millions of A1 property descriptions which 77m has which came from the service under the A1 Match licence, shortly before trial it emerged that there are about 3.5 million more such records whose origin cannot be accounted for. OS contends they must have been scraped. Mr Highland could not account for where they came from but maintained he was convinced 77m did not scrape.

224. I find that these records did come from HMLR. There is no other credible source. It is more likely than not that they came from the FAP service and are not from a missing batch or batches under the A1 Match service. Ms Wiles gave evidence that further instances of A1 matching beyond the ones already found were highly unlikely, and Mr Highland agreed with that. Dr Halstead also agreed they did not come from A1 Match.

225. There was a point about semicolons, OS contending that the presence of semicolons in the 3.5 million indicated that they came from FAP rather than A1 Match. I was not convinced that this proved as much as OS contended but it does not matter.

226. There was a suggestion in cross-examination on 77m's behalf that the data might have come from three other sources published by HMLR – data about property owned by corporate entities, the Business Gateway and data obtained from Mr Petty under a FOI. However there is no link to INSPIRE IDs in any of these three and so they cannot help 77m.

227. Given the finding that the data came from FAP, there are only two credible possibilities. Either the data was obtained by automatic scraping or it was obtained by individuals working on a large scale. The former is technically easy to do, particularly given the IT expertise available to 77m. However the latter is not so improbable since 77m did indeed use a service whereby individuals in Pakistan were paid to undertake the task of manually downloading this kind of data. The invoices were exhibited to Mr Highland's evidence. The documented work in Pakistan was not on a scale as large as 3.5 million records but it was done.

228. However the work was undertaken, it is likely that Dr Brown was involved. The contents of Dr Brown's laptop were deleted when he left the company, on Mr Highland's instruction and despite Mr Highland knowing that he ought to preserve electronic documents. On the other hand if the deletion was done as part of an effort to hide scraping one might have thought more would have been done to conceal or at least disguise the 3.5 million records themselves. There is no evidence for that.

229. Of the two possibilities, procuring the data by scraping is more likely than the idea that it was procured by a hidden army of individuals for which no other trace survives, such as email communications, invoices etc. If there was such evidence on Dr Brown's laptop and it has been lost to 77m then Mr Highland only has himself to blame. It is true that the matter was raised very late but it relates to such a large number of records that I would be very surprised if the sole source of evidence relevant to this was one laptop. I find that the relevant records, which seem to be 3.5 million in number were acquired by 77m using an automated tool, i.e. by scraping.

*Consequences if use was unlicensed*

230. The 3.5 million records were obtained in breach of the relevant licence terms. 77m made an open offer to delete them but that does not deal with any consequences of their use in creating Matrix, which I have found therefore was unlicensed.

*The RoS Land Values licence and what did 77m actually do with the data?*

231. 77m obtained two different datasets from RoS, the RoS Land Values data and the RoS INSPIRE cadastral seed points. The terms applicable to each are different. The terms applicable to the Land Values data are the RoS Land Value Licence from 2012. The terms applicable to the RoS INSPIRE data are the RoS INSPIRE terms from 2015. There is not now any issue about whether or not 77m's actions were permitted by the terms of the RoS INSPIRE terms. The issue is whether they were permitted by the RoS Land Values licence.
232. The RoS Land Values data combines addresses and a geolocation for each address. The geolocation is the x,y coordinate of a centroid of an OS polygon. The problem for 77m was that this centroid/geolocation consisted of OS data and 77m knew that. So, as explained in the section dealing with Matrix, what 77m did was this. 77m matched the addresses in the Master Address List with the address data in the Land Values data, then found the nearest INSPIRE cadastral seed point (which is itself an x,y coordinate) to the relevant centroid/geolocation in the Land Values data for that address and then, instead of storing the Land Value centroid/geolocation data for that address, stored instead the INSPIRE cadastral seed point against the address in the Master Address List. So 77m ended up with a link between an address in the Master Address List and an INSPIRE cadastral seed point (which is itself a good quality geolocation and, critically, was available for use by 77m). The link was provided by using the Land Values data. OS contends this use of the Land Values data was outside the 2012 licence.
233. The relevant terms of the 2012 RoS Land Value Licence are as follows:

- (a) Data is referred to as "Land Values plus House Type" data from 1 January 2005 to 31 May 2011 for every registration county in Scotland. This is in the first part of clause 1.1.
- (b) The second part of clause 1.1 provides:

"The Data will be provided to you to allow you to develop a web service containing house sale information on the following website:

<http://www.77m.co.uk/>. This website will be offering a one stop information service that aggregates many datasets together. The aim is to provide a comprehensive overview for any location. House price and house type information from the RoS Data will be made available via this website. Other parts of the Data will only be used for internal modelling purposes.”

(c) The third and fourth parts of clause 1.1 refer to the Data being provided on the understanding that 77m will be purchasing ongoing “Land Values plus House Type” Data and provide for a review after two years. Nothing turns on this.

(d) The fifth part of clause 1.1 provides:

“This licence grants you the following non-exclusive rights:

- To reproduce the Data in computer-readable form on your Website (Subject to the conditions of Clause 3 below)
- To link your Website to the Internet so that it may be accessed by your customers...
- To charge your customers for access to the Data
- To publish records and hardcopy publications using the information contained in the Data...”

(e) The other relevant clause is at 1.5, as follows:

“The Land Values data is provided for your exclusive use and will not be published, assigned or sold on in any way except as indicated within the agreed use shown above.”

(f) Clause 3 deals with data protection and requires 77m not to publish certain personal data.

234. The first issue is what does the Data mean? From the first part of clause 1.1 it appears to be confined to land value (i.e. house price) and house type. That is OS’s case. However 77m argued it meant the whole of the dataset provided by RoS. It is clear that the dataset did contain more than house prices and house types both in fact (it contained an address and the centroid/geolocation) and as contemplated in the licence. The licence refers to other “parts” of the Data apart from house price and house type at the end of the second part of clause 1.1 and it seems from data protection clause 3 that the house owner’s name would also be included.

235. OS suggested that 77m had agreed with OS’s narrower definition in the Re-Amended Particulars of Claim at paragraph 24U but that is not right. In that paragraph 77m contends that the term refers to the data set, which is 77m’s case.

236. The first part of clause 1.1 reads as OS contends but it is fair to note that the last sentence of the second part of clause 1.1 uses the capitalised term in a manner which is not so limited. If that was all there was to go on then I might accept 77m's case. However the bullet points in the fifth part of clause 1.1 do not make sense if Data there refers to everything in the dataset but do make sense if all they refer to is house price plus house type. That is because the first bullet permits reproduction of Data on the website but the second part of clause 1.1 makes clear that it is only house price and house type which is to be made available via the website. Other parts of the Data are to be used for internal modelling. The same problem arises for the third and fourth bullets.
237. I find as follows. In the bullet points "Data" is limited to house price and house type. In the first sentence of the second part of clause 1.1 the term Data refers to the whole dataset but the scope of what is permitted by the wide terms of that first sentence is qualified by the rest of the paragraph. Read as a whole the paragraph allows 77m to use house price and house type to develop a webservice which makes those two items of information available but it only allows 77m to use other parts of the dataset for internal modelling purposes in the development of the website. There is no other permitted use of information which is not house price or house type apart from internal modelling purposes as part of developing the website.
238. 77m also suggested that its approach to the definition of Data had to be correct because it would make no sense if, for example, the licence permitted 77m to do the various acts defined by the bullet points (reproduce the Data on the website, charge its customers for access to any of the Data, etc.) but could not use the Data itself save for very limited purposes. I do not agree. The terms as a whole make sense for the reasons I have explained.
239. Rather like the A1 Match licence, the oddity about these bespoke terms is that they contain a clear indication of an ostensible reason why 77m wanted the data which was on offer, but they do not mention the crucial second purpose which 77m had for the data, i.e. to ignore the land values and house type and use the simple combination of centroid/geolocation and address obtained from RoS to ascribe good geolocations to 77m's own Master Address List. The ostensible use emerging from the RoS Land Values licence is to operate a web service containing house sale information. Matrix is not itself a web service, although no doubt it could be used to provide one; and it is by no means limited to being a house sale information service, although again, as I understand it, could be used to provide one. What is not mentioned in the licence at all is the second use concerned with the centroid/geolocations.
240. Using Matrix to provide house price information on a web service would be licensed but that is not the limit of what 77m has done and wishes to do.
241. Irrespective of the point on a web service, the issue turns on what "internal modelling purposes" allows for. The question is whether that permits 77m to use the data to do what it did in creating Matrix.
242. OS contends it does not. It argues that a model is a smaller scale representation of something or a test/simulation. In the context of a website, as discussed in the RoS Land Value Licence, OS contends that a sensible construction, consistent with business common sense is that the other data could be used internally for a test or



small scale data processing, but not as a wholesale input to create a commercial product. Had the parties intended to cover such use then OS contends very different wording would have been used.

243. 77m contends that it does permit 77m to do what it did. 77m argues that the words have to be understood in the context and that context is an overall agreed use for the Data which is the creation of a commercial product which will provide “a comprehensive overview for any location” and “aggregates many datasets together”. These are both phrases from the second part of clause 1.1. 77m argues that this would not be possible unless the Data could be linked to a particular location, and matched to other data sets.
244. 77m also argued that the terms permit it to use all of the data to create its “one stop information service”, but that if 77m publishes the data on this service it will only make available the house price and house type but that can be after 77m has successfully linked the particular address to a location. The only concern is to ensure that individual names linked to addresses are not published— hence clause 3.
245. In my judgment OS is wrong in its submission that the terms do not permit use of the data to create a commercial product. Whatever was permitted was clearly permitted in the context of the licensee developing a commercial product, not least its web service.
246. I agree with OS that modelling is the process of making and using a representation of something larger, or it is a testing or simulation exercise. In context this means that the other data could be used internally to make a model of the service it intended to provide or to test it. The model or test could be of a commercial product but the clause did not permit 77m to use the data to produce a new geolocation for all or substantially all of the properties in the dataset. That was not an act of internal modelling.
247. However it is fair to say that one could not provide house price information on a web service without allowing a user to find the relevant house. That must be by the user entering an address or a location. For that to work the system has to be able to accept a user typing in an address, use that to find the right house and then display the right house price. This necessarily at least involves using the address and centroid/geolocation in the RoS Land Values data. So the term about internal modelling cannot, on an objective construction of the contract, have been meant to prevent that. 77m seek to extend this logic to include the sort of linking which this case is concerned with and to argue that the reference to aggregating datasets together supports that approach.
248. I do not agree. The reference to aggregating datasets does not help. The clause explicitly forbids use of other parts of the Data save for internal modelling. I agree that there must also be an implied term to allow the web service expressly contemplated to work, but the limit of such an implied term must be to what is necessary to achieve that, nothing more. Such a term does not have to extend so far as to include what 77m actually did. The test is one of necessity, and that is not necessary.

249. 77m is correct that if that is the limit of the licence it could not have done what it wanted to do, but that does not alter the analysis. There was no evidence to which I was directed about the circumstances leading up to the signing of this licence which might shed light on the issue. If 77m did not tell RoS what it intended to do in sufficient detail, it only has itself to blame if the contract does not permit it.
250. Accordingly 77m's use of the RoS Land Values data was not licensed by RoS.

*Authority*

251. There is no question that HMLR had authority to offer a licence of the scope of the INSPIRE Download Terms since, irrespective of the scope of the INSPIRE legislation, OS in fact agreed to it. Therefore the claim for infringement by the use of INSPIRE polygons (and data derived from them) fails because of the finding that 77m's activity was within the INSPIRE Download Terms.
252. In relation to the RoS Land Values licence, OS did have a pleaded case that the RoS Land Value licence fell outside the terms of the OSMA Member Licence and therefore RoS was not able to provide the centroids to 77m for the purpose of commercial re-use by 77m. Since I have found that what 77m did was not permitted by the RoS Land Values licence anyway the point is academic. There are no findings of primary fact based on live evidence which relate to this point and so I will say no more about it. The Re-Re-Re-Amended Defence and Counterclaim did also refer to addresses from RoS but as far as I can tell the infringement alleged was based on centroids.
253. In relation to the use of addresses from HMLR, this arises under the A1 Match Licence (OS pleading para 100J) and the FAP terms (OS pleading para 100BA). OS did have a pleaded case that the A1 Match licence and the FAP terms fell outside the terms of the PSMA Member Licence and therefore HMLR was not able to provide the addresses to 77m for the purpose of commercial re-use by 77m. Since I have found that 77m's scraping was not permitted by either of these licences the point is academic as far as those addresses are concerned. For reasons explained below it is also academic in relation to the 480,000 manually downloaded FAP addresses.
254. There is one point of primary fact based on the live evidence which relates to this. That is the allegation that various representations had been made to 77m in particular. I am not concerned with whatever might be made of representations to the public at large made in public documents but to a series of specific matters which Mr Highland gave evidence about. These did not withstand counsel's cross-examination. The only representations relied on in evidence by Mr Highland were in his 6<sup>th</sup> witness statement. However when pressed on them, it emerged that either he had not seen them at the time, or they were not about addresses at all or, to the extent they were said to relate to the FAP, they in fact did not. I find that no relevant representations addressed specifically to 77m were made at all. There was a further point, namely whether Mr Highland knew that OS's rights went beyond polygons (he accepted knowledge of rights to polygons) and on to addresses. I do not accept OS's case on that. In the cross-examination (at T1/140-142) Mr Highland did not accept he thought OS had rights in addresses. There is no reason why he should have and the point put to him did not prove it. I accept Mr Highland's evidence on that.

255. Among the points taken by 77m relating to Lichfield DC, there is one which resonated with me more strongly than the others. I agree with 77m that the Lichfield DC circumstances can be analysed in terms of apparent/ostensible authority from the point of view of a member of the public (77m). OS cited *East Asia Company Ltd v PT Satria Tirtatama Energindo (Bermuda)* [2019] UKPC 30 as authority for the principle that:

“ostensible authority is a relationship between a principal and a third party created by a representation made by the principal, which the third party can and does reasonably rely upon, that the agent of the principal has the necessary authority to enter into a contract on its behalf”.

256. This means that to be effective a representation must be by the principal and not the agent. But the Lichfield DC data was made available to the public on a gov.uk website under the OGL terms. On the face of it this was part of a government wide open data initiative. I hold that this does amount to a representation on behalf of the government as a whole, including the Crown and a wholly owned company like OS, that the data was being lawfully made available on those terms. This representation was not just on behalf of the agent (Lichfield DC), it was by the principal. Moreover it would be reasonable for a third party (77m) to rely on it. There is no ostensible reason not to. As regards OS itself, the fact that GeoPlace is not wholly owned by the government is irrelevant since OS is the entity entitled to grant licences in respect of AddressBase.

257. A further OS specific reason why reliance would be reasonable was because before 77m downloaded the Lichfield data in 2016, in February of the previous year OS had a publicly promulgated scheme called presumption to publish which was itself part of the wider government open data initiatives. Under that scheme a public body such as a local authority could publish certain kinds of data to be made available under the OGL. The public authority was simply required to give prior notice to OS. In fact in this case Lichfield DC did not do so but publication of the existence of such notice was not part of the scheme so there would no reason for a member of the public to think it had not been given. The details of the scheme were subject to various details and exceptions. They do not matter. In my judgment 77m is right that from the point of view of the public, it would have been likely that the Lichfield DC data which was made available would fall within the scheme. They would have every reason to assume it was within the scheme and no reason not to. Although he did not accept the conclusion, the answers given in cross-examination of O’Meara of OS on this issue supported 77m’s case. His major qualification was that it depended on the data. I agree but the Lichfield data was the kind of data which the presumption to published was concerned with.

258. Therefore I find that from 77m’s point of view, the terms of the OGL apply to the data downloaded from Lichfield DC. The OGL permitted 77m to do what it did with that data.

*Infringement of database right by 77m*

259. Having got this far the database right claims which are still live are infringement of the database right in the Topo database or in the Addressing databases (such as AddressBase) by
- i) the use of centroids from the RoS Land Values dataset; and
  - ii) the use of addresses from HMLR.
260. Turning to the law, database right is a *sui generis* property right originating in the Database Directive 96/9/EC. The Database Directive was transposed into domestic law by the Copyright and Rights in Databases Regulations 1997 (SI 1997/3032).
261. Article 7(1) of the Database Directive provides that database right subsists where “there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents”.
262. Database right is a negative right which prevents the “extraction and/or re-utilization” of the whole or a substantial part of a database.
263. The concepts of “extraction” and “re-utilization” are defined in Article 7(2).
- i) “Extraction” means “the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form”: Article 7(2)(a).
  - ii) “Re-utilization” means “any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission”: Article 7(2)(b).
264. Relevantly for this claim, the scope of database right is subject to a number of important limitations.
265. First, database right does not constitute an extension of copyright protection to mere facts or data: recital 45 to the Database Directive. This was reaffirmed by the Court of Appeal in *Football Dataco Ltd v Sportradar GmbH* [2013] FSR 30, in which Jacob LJ explained that database right protects the collection of data, not its constituent elements.
266. Second, the protection afforded by Article 7(1) only concerns acts of extraction and re-utilisation. Where the creator of a database makes the content of the database accessible to the public, the consultation of that database does not, by itself, constitute an infringement: *British Horseracing Board Ltd v William Hill Organisation Ltd* (C-203/02) at [54]-[55].
267. Extraction and consultation were considered in *Directmedia Publishing GmbH v Albert Ludwigs-Universität Freiburg* (C-304/07). This was about the manual and selective transfer of titles of German language poems from a database into another medium after the defendant had consulted the database onscreen. The CJEU held that this transfer constituted an act of extraction within the meaning of Article 7(2)(a). The decisive criterion was the existence of an act of transfer of all or part of the contents of the database to another medium, regardless of the particular mode of transfer or the fact that the contents of the original database might be arranged

differently in the new medium: [36]-[40]. However the CJEU also reaffirmed that the protection conferred by database right did not cover consulting a database which had been made publicly available for information purposes:[51]-[53]. In this the CJEU described consultation as being something which a database owner who has made the database accessible to third parties (free or paid for) cannot prevent them from doing for information purposes.

268. The CJEU in *Apis-Hristovich EOOD v Lakorda AD* C-545/07 followed the court in *DirectMedia*. At [41] the court described the transfer which amounted to an act of extraction as implying that the contents of the original database were to be found in a medium other than that database. At [44] to [45] the court noted that the difference between permanent and temporary transfer was the duration of storage and gave the operating memory of a computer as an example of a place in which temporary storage might take place.
269. The distinction between consultation and re-utilisation was examined in *Innoweb BV v Wegener ICT Media* C-202/12 where the defendant operated a meta search engine which enabled its users to simultaneously search various third party websites in real time. The CJEU's approach to consultation was the same as in *DirectMedia*. The court held that the defendant was not merely consulting the third party websites for information purposes – instead, it provided a form of access to those websites which was different from the access route intended by their owners: [47]. The defendant had therefore made available the contents of the websites to the public within the meaning of Article 7(2)(b).
270. The third limitation to database right arises from a series of defences set out in paragraph 3 of schedule 1 to the Database Regulations. I will address them in context below.
271. 77m originally pleaded exhaustion but this was dropped.

#### *Assessment*

#### *Centroids from RoS, extraction and consultation*

272. It was not disputed that the centroids in the RoS Land Values dataset were derived ultimately from Topo. The centroids therefore are part of the contents of an OS database. The processing which 77m carried out to use the centroids to find the nearest INSPIRE Cadastral seed point involved the temporary transfer of those centroids to the operating memory of a computer. That is another medium. The “other medium” to which they were transferred was the part of 77m's system which carried out the processing to work out the nearest INSPIRE cadastral seed point. Subject only to the issue of consultation which I address below, I would hold that this amounted to an infringing act of extraction of the contents of a database. The fact the centroids were not stored permanently in a new database does not matter because extraction can involve temporary transfer. The fact the individual items of data were transferred one address at a time and then, presumably, discarded before transferring the next one, makes no difference. The fact that the database from which 77m extracted them was the RoS Land Values rather than from Topo itself does not matter. Extracting contents from one database (A) which are themselves extracted from another database (B) is an act of extraction of the contents of database B as well as the

contents of A. The legal test requires extraction of all or a substantial part of the contents of the database. To be an infringement of database right in database B, it would not be enough for those extracted to be a substantial part of database A, they must be a substantial part of database B. In other words, in this case, a substantial part of Topo. The set of centroids for the land in Scotland is a substantial part of Topo (and 77m did not suggest otherwise).

273. I should make clear that by discarded I mean discarded as part of the process. 77m kept the data in the original form provided to it by RoS but that is not relevant to this issue.
274. 77m contended that what was done was only an act of consultation and not extraction. 77m emphasised that the relevant data (centroids) did not end up in a new database and the public who used Matrix would never be presented with the centroid. The geolocation in Matrix is not the (unlawful) centroid it is the (lawful) seed point. 77m argued that it was using each given centroid for information purposes – i.e. to draw an inference about other data (the closest seed point).
275. OS seemed to think it was obvious that what 77m was doing was not consultation but extraction. It was not so obvious to me. It is not too hard to see why 77m's actions fall within the definition of extraction itself (see above) but what is less clear is why they do not also fall within the scope of consultation. If they do then I take it from the cases that there is no infringement.
276. In the end I have decided that what 77m have done is not consultation. The reasons are these. What is apt to confuse the issue is that in economic terms what OS is trying to stop is Matrix, but Matrix does not contain the relevant contents (the centroids). Matrix is not the "other medium". The fact that the process which involved the putative act of extraction by 77m of the centroids was ultimately something which led to another database being produced, is irrelevant to the analysis. If what 77m did is an act of extraction that must be true whatever it is 77m went on to do having used the centroids for its purposes and discarded them. This is I think what the court is referring to at [47] of *BHB v William Hill*. That case also used the term appropriation to refer to extraction, which has been picked up later (see [51]). Moreover while re-utilisation involves making available to the public, extraction does not have to.
277. There may be a simple answer, as follows. In *DirectMedia* at paragraph 60 the court refers to on screen consultation. It may be that what the CJEU is talking about there is that a situation in which a person reads the data on a computer screen and does nothing else is consultation. It is not extraction because the only possible "other medium" into which the contents have been transferred is the individual's brain and that is not a relevant sort of medium. If things are written down and then it is done on a large scale then there may be extraction but the act of on screen consultation is not infringing. If that is what the CJEU means then clearly 77m does not do this sort of consultation. However I have misgivings about this way of reading the cases. It is not clear that that is what is meant and, for example, I do not see why consultation by a user themselves sitting by a terminal should be exempt while consultation by a user accessing data through their own device like a mobile phone might not be.

278. 77m is using the centroids from the original database for information purposes in a sense (to draw an inference) and that chimes with the references to consultation in *DirectMedia* and *Innoweb*. Moreover 77m was given access to the database albeit what 77m then did was not licensed. However what I think deprives 77m's activity of the character of mere consultation is its scale. When a member of the public, or a commercial user, wishes to consult the database to learn something about a particular entry or to learn something about particular entries, they consult the database. By contrast someone who takes all or a substantial part of all the contents, and transfers them into another medium so that they can use them, is appropriating to themselves a substantial part of the investment which went into creating the database. Protecting that investment is what database right is for. That is what 77m did and that is why it is extraction not consultation. There may be a grey area between the sort of commercial consultation I refer to and wholesale activity of the kind carried out by 77m but the scale of 77m's actions puts them firmly on the extraction side of the line.
279. I will deal with the paragraph 3 exceptions at the end of this section.

*Addresses from A1 Match and FAP*

280. The other relevant act alleged to infringe database right is the use of the addresses from the A1 Match process and from the FAP. This raises two issues. The first is whether 77m has committed an act restricted by database right. 77m used the addresses provided via A1 Match and FAP in order to match them with an existing address 77m already had in the Master Address List. To do this the A1 Match/FAP address had to be copied into temporary computer memory.
281. However we know that another reason for acquiring these addresses was to cleanse garbage polygons/non-addressables. That act was permitted. Since the A1 Match data consists of addresses corresponding to INSPIRE IDs which 77m selected, I take it that an attempt was made to match at least every A1 Match address which had not been identified as a non-addressable or a garbage polygon. That will have represented a substantial part of the A1 Match data. It is more likely than not that the same applies to the FAP addresses, in other words that 77m must have defined a set of INSPIRE IDs for which it sought addresses by accessing FAP (manually or by scraping) and so the same conclusion follows, that at least a substantial part of the FAP addresses were themselves tested against the Master Address File to try and match them.
282. In other words the number of addresses used, aside from the task of cleansing garbage polygons/non addressables, represents a substantial part of the address data from HMLR.
283. For each attempt at matching a given address, the address in temporary computer memory was discarded afterwards. What changed was that for a successful match the Master Address File entry was now linked to an INSPIRE ID. 77m has also kept the A1 Match and FAP addresses in the dataset(s) it acquired from HMLR. At least as far as A1 Match is concerned, that act of keeping is licensed under the A1 Match licence albeit 77m may owe a contractual obligation to HMLR to delete the A1 Match data at some point.

284. I find that this matching activity amounts to extraction. It is not mere consultation. The numbers of addresses extracted is very high – well into the millions. I will come back to substantial part below. There is no act of re-utilisation because 77m has not made the addresses acquired from A1 Match or FAP available to the public.
285. The second issue raised by the addresses is the relationship between the A1 Match and FAP addresses acquired by 77m and OS's database rights relating to the Addressing databases. The issues can be dealt with by focussing on the current AddressBase product alone. The simple starting point is that 77m got the addresses in issue via the A1 Match process and FAP from HMLR, not from OS. What, one might ask, has OS got to do with it? The answer is as follows. The addresses acquired via the A1 Match process and FAP were the A1 property descriptions. There are two sources for the A1 property descriptions. Most of them have come from something called the Land Registry Property Gazetteer (LRPG), which is HMLR's own in house address database. The remainder have come from information given by HMLR's "customer", i.e. a person seeking to register a title. OS can have no claim to database right relating to the addresses from HMLR customers, the issue is the addresses from the LRPG.
286. The evidence of Mr Robson of HMLR was that the primary source of the addresses in the LRPG since April 2014 has been AddressBase, and since 2001 has been its predecessors. Before 2001 the primary source was the PAF. Mr Robson's evidence, which I accept, was that for the significant majority of titles which are a single addressable property, the A1 property description will be the AddressBase address. For multiple occupancy buildings the situation is more complicated (but does not matter) and for non-addressable sites the A1 property description is likely to be HMLR created.
287. However the position is not as simple as it might seem. Mr Robson also exhibited the current edition of HMLR's Practice Guide 1 which states in terms in para 4.3.2 that when a person applies for a first registration "We [HMLR] will generally enter the address in the register from the Post Office address file". He explained in his witness statement that this reference to the Post Office address file was to "the verified PAF addresses which exist within AddressBase". In cross-examination Mr Robson accepted that the ultimate data which HMLR is putting into the LRPG is the Post Office address, albeit as Mr Robson added, "from AddressBase". Therefore as 77m puts it, the text in these A1 property descriptions comes from the PAF, albeit via AddressBase.
288. To someone versed in the idea that Ordnance Survey is a mapping organisation, one might think that the real value of AddressBase is that the addresses are accurately geolocated, and one might think that its major distinguishing feature over the PAF is that the addresses in the PAF are not geolocated. Thus once the PAF addresses find their way into AddressBase the value added is an accurate geolocation. Viewed this way, one might wonder how it is that OS can effectively appear to have acquired rights over the PAF itself. 77m contends that is what is going on in this case. 77m's point is that the geolocation information it uses comes from INSPIRE and is lawful. The only thing 77m is using which derives from AddressBase is the address text itself and that came from the PAF.



289. To examine this point one needs to look more closely at AddressBase. In this respect its predecessor databases were the same. For all relevant purposes in this case AddressBase is simply a copy of the NAG. Database right in relation to AddressBase is therefore owned by GeoPlace and licensed exclusively to OS. OS has authority to license it on. The bulk of the addresses in the NAG are from the PAF but there are other sources too. The main other source of addresses is the local authorities' database called the National Land and Property Gazetteer (NLPG). Of course GeoPlace is itself a joint venture between OS and local government. OS emphasises that the address data from the NLPG is important because it is more up to date than the PAF. No doubt some NLPG origin addresses in the NAG have found their way into HMLR's database LRPNG and on into the data given to 77m but there is no clear evidence from which to draw even an approximate conclusion about scale. Mr Robson's evidence, who was called by OS, is also against it.
290. The critical point which OS makes is that the PAF addresses in the NAG are not merely from the PAF, they have been verified. This verification includes a number of steps. The relevant one is to verify the PAF addresses against the local authorities' NLPG data. Thus the set of PAF addresses in the NAG is a subset of all PAF addresses. A significant part of the value of a PAF address appearing in the NAG (and therefore in AddressBase) is that it has been through this verification process. Thus, contends OS, a database right exists in the NAG/AddressBase in relation to the addresses themselves, even though the bulk of them came from the PAF.
291. 77m submitted that all this cannot change the fact that what was being consulted by HMLR was in reality just PAF data. I do not agree.
292. The verification process was described in detail by Mr Griffiths of GeoPlace and Mr Chambers of OS. Mr Chambers was not cross-examined. Mr Griffiths was cross-examined by 77m. The point 77m submitted this cross-examination established was that the addresses in the NAG had to match the PAF (or other third party data) and that, as 77m put it, "the NAG was in essence a cleansed version of PAF". These points are true as far as they go but they seriously understate the effort which goes into the verification process and its significance. GeoPlace receives PAF files from Royal Mail every day. On average 90% of the new records can be matched to addresses already in the NAG. The 10% which cannot be matched are sent to local authorities at the end of each month to check. Mr Griffiths said that after years of continual investment and improvement by GeoPlace 99.7% of PAF records and 99.9% of VOA council tax records have found a home in the NAG. If a record cannot be matched it will not be included and so records are not inserted directly from the PAF, but only enter the NAG once they have been matched and verified. Mr Griffiths explained that the investment in maintaining the NAG is the majority of GeoPlace's operating expenditure, which in recent years has been about £6 million pa. I accept his evidence.
293. This kind of investment in verifying the contents of a database is exactly what the database right was created to protect – see recitals (7) and (39) on investment and Article 7 of the Directive along with recital (40) which explains that the object of the *sui generis* database right is to ensure protection of any investment in obtaining, verifying or presenting the contents of a database.

294. Therefore the argument advanced by 77m, which emphasises that the text of the addresses comes from the PAF is true but ignores the considerable investment in verification of these addresses which is undertaken by GeoPlace to maintain the NAG. The fact the text comes from the PAF does not mean that GeoPlace has no database right in the collection of addresses in the NAG. It does.
295. Given the number of addresses extracted by 77m, I hold that this represents a substantial part of the contents of the NAG and, subject to the defences below, this would represent an infringement of database right.

*Defences*

296. 77m relied on the defences in Paragraph 3 of Schedule 1 to the Database Regulations. They are applied by paragraph 20(2) of the Database Regulations themselves. Paragraph 3 provides as follows:

**Material open to public inspection or on an official register**

(1) Where the contents of a database are open to public inspection pursuant to a statutory requirement, or are on a statutory register, database right in the database is not infringed by the extraction of all or a substantial part of the contents containing factual information of any description, by or with the authority of the appropriate person, for a purpose which does not involve re-utilisation of all or a substantial part of the contents.

(2) Where the contents of a database are open to public inspection pursuant to a statutory requirement, database right in the database is not infringed by the extraction or re-utilisation of all or a substantial part of the contents, by or with the authority of the appropriate person, for the purpose of enabling the contents to be inspected at a more convenient time or place or otherwise facilitating the exercise of any right for the purpose of which the requirement is imposed.

(3) Where the contents of a database which is open to public inspection pursuant to a statutory requirement, or which is on a statutory register, contain information about matters of general scientific, technical, commercial or economic interest, database right in the database is not infringed by the extraction or re-utilisation of all or a substantial part of the contents, by or with the authority of the appropriate person, for the purpose of disseminating that information.

(4) In this paragraph –

‘appropriate person’ means the person required to make the contents of the database open to public inspection or, as the case may be, the person maintaining the register;

‘statutory register’ means a register maintained in pursuance of a statutory requirement; and

‘statutory requirement’ means a requirement imposed by provision made by or under an enactment.

297. Neither party was able to find any cases which consider the scope or applicability of these exceptions.
298. Each sub-paragraph defines acts which may be done in relation to a database despite the existence of database right. I will refer to these three provisions as the authorised extraction defence (paragraph 3(1)), the time and place shifting defence (paragraph 3(2)) and the general information dissemination defence (paragraph 3(3)). None of these labels is a perfect summary of the provisions but it makes the exercise of dealing with them a bit less indigestible.
299. Starting with the authorised extraction defence (sub-paragraph 3(1)), the way it works is that database right in a given database is not infringed by the extraction of all or a substantial part of the contents provided the following elements are satisfied:
- i) The contents of the relevant database are open to public inspection pursuant to a statutory requirement, or are on a statutory register,
  - ii) The contents extracted are contents containing factual information of any description,
  - iii) The extraction was by or with the authority of the appropriate person,
  - iv) The appropriate person being the person required to make the contents of the database open to public inspection, or the person maintaining the statutory register,
  - v) The extraction was for a purpose which does not involve re-utilisation of all or a substantial part of the contents.
300. The way the definitions of “statutory requirement” and “statutory register” work means that the first element (i) applies to two cases: (a) the contents are open to public inspection pursuant to a requirement imposed by provision made by or under an enactment, or (b) contents are on a register maintained in pursuance of a requirement imposed by provision made by or under an enactment. Notably case (a) requires the contents to be publicly open while limb (b) does not. The term register is not defined.
301. In putting the elements this way I have read the qualification about “containing factual information” as applicable to the contents extracted rather than applicable to the database as a whole because that makes more sense. Also I have read the reference to the authority of the appropriate person as a reference to the act of extraction (rather than to the contents being open to public inspection etc.) because that makes most sense anyway and because of the way appropriate person is defined.
302. Finally, for the authorised extraction defence, the extraction can be for any purpose whatsoever as long as it does not involve re-utilisation of all or a substantial part of

the contents. Thus the authorised extraction defence will not apply if the extraction was in order to make the contents available to the public.

303. Turning to the time and place shifting defence (sub-paragraph 3(2)), this works in a similar but not identical way. Under this limb, database right in a given database is not infringed by acts of extraction or by acts of re-utilisation. That is broader than the authorised extraction defence. The acts of extraction or re-utilisation of all or a substantial part of the contents of the database do not infringe provided the following elements are satisfied:
- i) The contents of the relevant database are open to public inspection pursuant to a statutory requirement,
  - ii) The extraction or re-utilisation was by or with the authority of the appropriate person (as defined above),
  - iii) The extraction or re-utilisation was for the purpose of
    - a) enabling the contents to be inspected at a more convenient time or place; or
    - b) otherwise facilitating the exercise of any right for the purpose of which the statutory requirement requiring the database to be open to inspection was imposed.
304. Unlike the authorised extraction defence, the time and place shifting defence only applies to a publicly open database (pursuant to an enactment). The definition of publicly open is the same as before and the relevant part of the definition of appropriate person still applies. Also unlike the authorised extraction defence, the time and place shifting defence is not limited to contents which contain factual information, this exception is wider and applies to any contents.
305. Finally whereas with the authorised extraction defence the extraction could be for any purpose as long as it did not involve re-utilisation, to be covered by the time and place shifting defence the act of extraction or re-utilisation has to be for either of two particular purposes. One purpose is to allow a more convenient inspection of the contents (more convenient in time or space). The other is to facilitate an exercise of the right which was the reason why the database was open to public inspection in the first place.
306. The general information dissemination defence (sub-paragraph 3(3)) again works in a similar but not identical way to the other two. Under this limb, like the time and place shifting defence, the general information dissemination defence can apply to acts of extraction or re-utilisation. The acts of extraction or re-utilisation of all or a substantial part of the contents of the relevant database do not infringe provided the following elements are satisfied:
- i) The contents of a database which is open to public inspection pursuant to a statutory requirement, or which is on a statutory register,

- ii) The contents contain information about matters of general scientific, technical, commercial or economic interest,
- iii) the extraction or re-utilisation is by or with the authority of the appropriate person (as defined above),
- iv) the extraction or re-utilisation is for the purpose of disseminating that information.

307. Thus like the authorised extraction defence and unlike the time and place shifting defence, the general information dissemination defence applies to publicly open databases and statutory registers which need not be publicly open. The definitions of these concepts and the appropriate person are as before.
308. Also similar to the authorised extraction defence and unlike the time and place shifting defence, the exception is limited to contents of a particular kind. In this case it is information about matters of general scientific, technical, commercial or economic interest.
309. Whereas in a manner similar to the time and place shifting defence and unlike the authorised extraction defence, to be covered by the general information dissemination defence the act of extraction or re-utilisation has to be for a particular purpose, that is the purpose of disseminating the relevant information.
310. An issue which applies to all three defences is about the correct way to identify the database to which this all applies and the appropriate person. The circumstances in the present case are that 77m acquired the data it did from a public authority (RoS or HMLR) and as a result of access to a database held by that public authority (RoS Land Values for the centroids and the Register of Title for the addresses). However the database right which would be infringed arises from another database – Topo in the case of the centroids and the NAG or AddressBase in the case of the addresses. How do these provisions work in such a case?
311. I am not aware of any separate statement, apart from the terms of the regulations themselves, which indicates what the purpose(s) of these defences is or are, but I think it is not too hard to see what they might be. The defences are there so that database rights do not frustrate the purpose of making various sorts of information open to the public or on a statutory register in the first place. Once information is put on a statutory register or, pursuant to a statute, made publicly open, the public as a whole, including commercial organisations as much as private persons, ought to be free to use that information at least to some extent (defined in the defences) without fear of trespassing on database rights, always assuming the member of the public has acted with the authority of the appropriate person. In that regard it is notable that the appropriate person is not defined as the person who holds the relevant database right. On the contrary it is the person required to make the contents publicly open or to maintain the register. That means that these provisions must have been intended by Parliament to provide a defence irrespective of whether the acts were consented to by the database right holder.
312. Accordingly, the right way to approach this is to identify the database in which database right is claimed and also to identify the contents of that database which have

been used by the person seeking to take advantage of the defence. I refer to this person as the user. The issue is not whether the database itself has been made publicly available or is on a statutory register, rather the question is whether the relevant contents used by the user have been made publicly available or are on a statutory register. Next one has to identify the “appropriate person”. That is not the owner of database right, it is the person required to make the contents public or maintain the register. Once all these things are identified one can apply the provisions. As long as the user carries out acts within the defences and has the authorisation of the appropriate person, the user’s use of the contents does not infringe any database right.

313. Each party sought to deploy Paragraphs 4 and 6 of Schedule 1 of the Regulations to support their submissions but I do not believe these paragraphs advance the issues. Paragraph 4 is concerned with what the Crown can do in certain cases when the contents of a database have been communicated to the Crown. Paragraph 6 provides that doing an act authorised by an Act of Parliament does not infringe database rights (unless the Act of Parliament provides otherwise).
314. Now I need to mention the Berne three step test. OS submitted that these defences had to satisfy the three step test in Art 9(2) of the Berne Convention for the Protection of Literary and Artistic Works. The argument is that these defences are derived from Art 6 of the Database Directive which itself stipulates at Art 6(3) that the exceptions need to be interpreted in accordance with the Berne three step test. Art 6 does refer to the Berne three step test but the Article is Chapter II of the Directive dealing with copyright in databases (and refers expressly to the Restricted Acts in Art 5). *Sui generis* database right is in Chapter III. Chapter III defines different acts protected by database right (extraction and re-utilisation) from the acts defined in Art 5 of Chapter II which relate to copyright. Nor does the Berne Convention itself purport to apply to *sui generis* database rights. Accordingly, although the terms of parts of Art 6 are similar to aspects of these defences, I believe OS is mistaken to say that the defences in Paragraph 3 of Schedule 1 of the Database Regulations have to satisfy the Berne three step test. 77m submitted the three step test did not apply but also argued that even if it did, the defences comply with it anyway. That was on the basis that they do not conflict with normal exploitation of the work if Parliament has required by statute that such contents be made available. The need for authorisation means there is no unreasonable prejudice to the rights holder’s legitimate interests and since the defences are limited to data made available in public datasets by statute, they relate to a certain special case. I can see the force in 77m’s point but I do not have to decide this question.

*Application of the authorised extraction defence (paragraph 3(1))*

315. The first question is whether the contents of the relevant database are open to public inspection pursuant to a statutory requirement, or are on a statutory register. To address this I remind myself that the relevant contents are the OS centroids which 77m used from the RoS Land Values dataset and the addresses 77m acquired from HMLR by the A1 Match and through FAP. I am not concerned with INSPIRE data.
316. 77m contended that the RoS Land Values data was open to public inspection or are on a statutory register pursuant to the Land Registration (Scotland) Act 2012 and that the addresses from HMLR were open to public inspection or are on a statutory register

pursuant to the Land Registration Act 2002 and the Land Registration Rules 2003 (see e.g. paragraph 5(5)). OS did not challenge any of that but submitted it was the wrong question based on the point of law I have decided above. OS contended that the right question was whether Topo (which is where the centroids came from) or the NAG/AddressBase (which is where the addresses came from) were publicly open or were a statutory register.

317. For the reasons already explained, the right approach is to identify the database in which database right is claimed and also to identify the contents of that database which have been used by the person seeking to take advantage of the defence. The answers are that for the Topo database the contents are the centroids in the RoS Land Values dataset and for the NAG/AddressBase the contents are the addresses on the Register of Title.
318. In each case these contents are open to public inspection or are on a statutory register under the applicable land registration legislation explained above. The appropriate person in each case is RoS and HMLR respectively. It is not OS.
319. The next question is whether the contents extracted are contents containing factual information of any description. They clearly are and I do not think this was disputed.
320. Was the extraction carried out with the authority of the appropriate person? Starting with RoS Land Values and the A1 Match, the answer is no. It is true that those contents were obtained with the authority of the appropriate person and that therefore the initial downloading, itself an act of extraction, was authorised. However the relevant act of extraction in each case is the one dealt with above – using the centroids to find the nearest seed point or matching addresses to ascribe a geolocation to an address in the Master Address File. For the reasons I dealt with in relation to the various licences, those acts were not authorised. None of these defences provide that once the contents have been acquired by an authorised initial act of extraction, the user is free to do anything at all with them irrespective of the authorisation of the appropriate person.
321. The same applies for the scraped FAP data, even more so since the initial extraction of that data was not authorised either. However for the 480,000 addresses manually downloaded from FAP, the relevant act of extraction was the putting of that data to commercial use. That activity was authorised by the FAP terms of service promulgated by HMLR. The appropriate person (HMLR) authorised commercial use of the data. Moreover all 77m did was an act of extraction. No re-utilisation was involved because the FAP addresses were not made available to the public. OS disputed whether HMLR had authority to grant such a licence but I hold that it does not matter for these defences to apply. This example shows the authorised extraction defence working in the manner it should. The defence allows the public to take at face value the terms of an authorisation granted by a public body over the use of data that public body has made available on a public database it is responsible for.
322. The only act relevant to the authorised extraction defence was an act of extraction. OS suggested 77m had admitted to some acts of re-utilisation but that admission was in general terms and did not relate to the centroids or addresses in issue.

*Application of the time and place shifting defence (paragraph 3(2))*

323. The conclusions I have reached about the relevant contents, databases, appropriate persons and authorisation apply to this defence too and mean that it cannot help 77m any further than the previous defence. In any case I was not convinced the relevant use by 77m in either activity was for the purposes relevant to the time and place shifting defence.

*Application of the general information dissemination defence (paragraph 3(3))*

324. The conclusions I have reached about the relevant contents, databases, appropriate persons and authorisation also apply to this defence too and mean that it cannot help 77m any further than the authorised extraction defence. OS accepted that the contents amounted to information about matters of general scientific, technical, commercial or economic interest but denied that what 77m did was for the purpose of disseminating that information because it was instead for profit. That was relevant because the *sui generis* database right is to protect investment. I have to say I am not convinced that the fact that a user wanted to operate for a profit must necessarily rule out the idea that they were doing acts for the purpose of disseminating information about matters of general scientific, technical, commercial or economic interest. Publishing an encyclopedia springs to mind. However it is not necessary to explore that any further.

*Estoppel*

325. 77m sought to rely on various estoppels by representation allegedly made. There were two categories. One related to authority and does not arise. The other relates to the use which 77m may lawfully make of data pursuant to the licences already considered. Most of the representations relied on related to the INSPIRE polygons or INSPIRE IDs and therefore do not arise. Although paragraph 12.66 of 77m's written closing includes a reference to relying on representations made by RoS, as best I can tell those relate to authority only and so do not arise.
326. The only representation relied on which relates to addresses is what are said to be encouragements by Ms Nicholson that 77m was able to use the FAP portal to verify its data. The idea that this could apply to addresses did not survive cross-examination (T1/177). I find that no relevant representations were made by Ms Nicholson or anyone else at HMLR that 77m was entitled or permitted to use addresses obtained from the FAP in the way it has done. I leave aside the point that what 77m did by matching addresses was not merely verifying data on any view – it was creating new links to geolocations from addresses. Also it is not and cannot be suggested that any encouragement of this sort could have led 77m to believe it could scrape the FAP data. Thus at best this could apply to the 480,000 manual FAP addresses. I have not got into that because of my conclusion under the authorised extraction defence.

*Pleading points*

327. In dealing with the issues this way I have not upheld certain pleading objections taken by OS. The main one was about 77m's case that even if 77m was unlicensed it was still necessary for OS to prove that 77m had committed acts of infringement of database right. Purely in terms of the pleadings I have a little bit of sympathy with OS, particularly given the bewildering way in which this case was pleaded, but not that much. OS has the advantage of a team of experienced intellectual property lawyers. Its case is that 77m infringed an intellectual property right. One might



expect the legal team for a rights holder to be able to articulate what that case was, which indeed they did when they had to. The position of 77m was never so clear cut as to absolve the team for the rights holder of needing to be in a position to do that. There was a list of issues too but that list itself was part of the problem.

*Procuring breach of contract*

328. The leading authority is *OBG Ltd & Anr v Allan & ors* [2007] UKHL 21 in which the House of Lords restated the principles relating to the tort of inducing a breach of contract. There are four elements that must be established:

- i) a contract between a claimant and a third party must have been breached by the third party. An actual breach of the contract must take place, mere interference is not enough.
- ii) the defendant must have known of the existence of the contract and its terms. It is not enough for the defendant to know they are procuring an act which in fact is a breach. To be liable the defendant must actually realise it will have that effect. Knowledge of circumstances which would indicate that fact to a reasonable person is not enough whereas turning a blind eye and consciously deciding not to enquire could well be.
- iii) the defendant(s) must have intended to induce or procure a breach of that contract.
- iv) such breach of contract must have been induced by the defendant. The question is whether the defendant's acts of encouragement, threat, persuasion and so forth have sufficient causal connection with the breach to attract accessory liability.

329. A further point is that the tort is not actionable *per se* but only when loss or damage is proved.

330. OS contended that even if these elements were established, it had a defence of justification. That defence was recognised by Lord Nicholls in *OBG v Allen* at paragraph 193. He explained that the defence of justification may be available to a defendant in inducement tort cases and gave the example of a defendant who interferes with another's contract in order to protect an equal or superior right of his own, citing *Edwin Hill & Partners v First National Finance Corpn plc* [1989] 1 WLR 225.

331. In the latter case Stuart-Smith LJ explained the principle as follows at p.233:

“Justification for interference with the plaintiff's contractual right based upon an equal or superior right in the defendant must clearly be a legal right. Such right may derive from property, real or personal, or from contractual rights. Property rights may simply involve the use and enjoyment of land or personal property. To give an example put in argument by Sir Nicolas Browne-Wilkinson V.-C., if X carries on building operations on his land, they may to the knowledge of X

interfere with a contract between A and B. to carry out recording work on adjoining land occupied by A. But unless X's activity amounts to a nuisance, he is justified in doing what he did. Alternatively, the law may grant legal remedies to the owner of property to act in defence or protection of his property; if in the exercise of these remedies he interferes with a contract between A and B. of which he knows, he will be justified. If, instead of exercising those remedies, he reaches an accommodation with A, which has a similar effect of interfering with A's contract with B, he is still justified notwithstanding that the accommodation may be to the commercial advantage of himself or A or both. The position is the same if the defendant's right is to a contractual as opposed to a property right, provided it is equal or superior to the plaintiff's rights. In my judgment that is the position in this case; I therefore agree with the judge's conclusion and would dismiss the appeal."

332. In my judgment this principle is capable of applying to intellectual property rights as much as to rights in land. Counsel for the claimant did not suggest otherwise.

*Assessment*

333. To recap, I have found that the terms of the A1 Match contract between HMLR and 77m were continuous such that HMLR was obliged to supply A1 property descriptions in response to a request from 77m, for a volume based fee. In August 2015 77m made a seventh request to HMLR but in September 2015 Santiago Jagot of OS emailed Ms Nicholson of HMLR to ask about the supply of addresses to 77m. He had heard from Mr Highland that 77m had an agreement with HMLR about this. HMLR never supplied any further A1 property descriptions to 77m after that.
334. I find that that refusal by HMLR to supply a response a reasonable time after August 2015 (a matter of weeks or a few months) was a breach of contract by HMLR.
335. Following a hearing before Arnold J at which the procuring case was almost struck out, the judge instead allowed the claim to proceed but only on the basis that it was limited to a case after receipt by OS of Annex 16 of the Particulars of Claim which happened on 12<sup>th</sup> September 2016. For reasons which do not matter the date can be treated as August 2016 but not earlier. During the trial 77m raised the possibility of amending its pleading to plead an earlier date but that would in all probability have required an adjournment and when I put that to 77m, the application to amend was dropped.
336. So 77m's case on procuring is based on the submission that HMLR's continuing breach after August/September 2016 was procured by OS, once it had notice of the terms of the licence.
337. Dealing with the facts first, I am quite sure that in about September 2015 Mr Jagot of OS did indeed procure or induce HMLR not to supply a response file to 77m in response to the seventh A1 Match request. Given the pleaded case it is not necessary for me to find what Mr Jagot knew at that time about the contract between HMLR and

77m but doing my best I must say I think HMLR were downplaying to OS what they were doing with 77m and in all likelihood emphasising the bespoke and, as they put it, “one off” nature of the relationship. I doubt Mr Jagot at that time knew that there was the kind of continuous supply obligation which I have found to exist.

338. In reaching this conclusion I reject part of OS’s case that HMLR had already decided on its own initiative not to supply 77m with any more A1 Match data. It is true that internally HMLR had started having concerns and may even have suspended supply but I am not satisfied a final decision was made until after Mr Jagot intervened.
339. What did happen was that in August/September 2016 OS were given annex 16. This is not a copy of the A1 Match licence nor does it explain enough detail to put OS on notice. OS finally received a copy of the A1 Match licence in April 2017. I find that OS were on notice of the contract terms only after April 2017.
340. If the procurement/inducement Mr Jagot undertook in 2015 had happened after April 2017, then things might be different, but there is no evidence that took place and I find it did not. 77m suggested that at no time after August/September 2016 did OS withdraw its “instruction” to HMLR not to supply a response file to 77m. By April 2017 these complex legal proceedings had been on foot since September 2016. I do not accept that in those circumstances not doing something after April 2017 satisfied the legal test, even if in other cases it would be possible for inaction to amount to procuring.
341. It is not necessary to examine the defence of justification. I will say only that I can see the force in such a defence in circumstances in which the holder of intellectual property rights (OS) which have been sub-licensed by a licensee (HMLR) to a third party (77m) and as part of which sub-licence the licensee is supplying something to the third party in which the IP rights subsist (addresses) for the third party to use under the sub-licence, if that rights holder then finds out that the third party is acting in breach of the terms of the sub-licence and using what has been supplied in a manner which infringes the underlying rights, the rights holder might be justified in seeking to procure the licensee to stop any further supplies to the third party even knowing that the licensee was bound by contract to do so.

### *Conclusion*

342. I conclude:
- i) 77m has not breached the INSPIRE Download Terms. Therefore the claim for infringement of OS database right relating to 77m’s use of INSPIRE polygons or data derived from INSPIRE polygons fails.
  - ii) 77m has breached the A1 Match licence.
  - iii) 77m scraped 3.5 million addresses from HMLR’s Find a Property Service and breached the applicable terms.
  - iv) 77m’s use of the centroids breached the RoS Land Values licence.

- v) 77m's use of the centroids from the RoS Land Values dataset and the addresses acquired from HMLR via the A1 Match licence and scraping the FAP service amounted to acts of infringement of database right held by Ordnance Survey or GeoPlace in the Topo database and the NAG/AddressBase database respectively.
  - vi) 77m's use of the addresses manually downloaded from FAP, which were about 480,000 in number, was within paragraph 3(1) of Schedule 1 of the Database Regulations (the authorised extraction defence) and therefore did not infringe any OS database rights. 77m's other activity did not fall within any of the defences in paragraph 3 of Schedule 1 of the Database Regulations.
  - vii) No estoppel by representation relating to 77m's actions arises.
  - viii) No issue about authority to grant licences arises.
  - ix) The claim for procuring breach of contract fails.
343. Although 77m achieved a measure of success, the winning parties in this case are the claimants on the counterclaim, Ordnance Survey and GeoPlace.