

Freedom of Information Act 2000 (FOIA)

Decision notice

Date: 23 January 2017

Public Authority: Department for Environment, Food and Rural Affairs

Address: Noble House
17 Smith Square
London
SW1P 3JR

Decision (including any steps ordered)

1. The complainant requested from the Animal and Plant Health Agency ("APHA") information about tests on cattle. APHA applied section 12 to the request.
2. The Commissioner's decision is that APHA has correctly applied section 12 to the request and so does not require it to take any further steps to ensure compliance with the legislation.
3. As an executive agency of the Department for Environment, Food and Rural Affairs, APHA does not constitute a public authority for the purposes of FOIA and so this notice is issued to its parent Department.

Request and response

4. On 9 May 2016 the complainant requested the following information from APHA under FOIA:

"Please send to me the annual number of cattle reactors per 1000 skin tests on a bovine in England when performing routine whole herd tests on the following herd sizes in each of the last 2 years.

1-10

11-50

51-100

101-200

201-300

>300

As such, please send to me 6 numbers for 2014 and 6 numbers for 2015."

5. APHA responded on 2 June 2016. It denied holding the requested information.
6. The complainant requested an internal review on 2 June 2016. APHA sent him the outcome of its internal review on 28 June 2016. It revised its position and applied section 12 to his request.

Scope of the case

7. The complainant contacted the Commissioner on 11 July 2016 to complain about the way his request for information had been handled. He specifically complained about APHA's application of section 12 to his request.
8. The Commissioner considered whether APHA had correctly applied section 12 to the complainant's request.

Reasons for decision

Section 12 – Cost of compliance exceeds the appropriate limit

9. APHA argued that section 12 applied to the complainant's request.
10. Section 12(1) of FOIA states that:

"Section 1(1) does not oblige a public authority to comply with a request for information if the authority estimates that the cost of complying with the request would exceed the appropriate limit."

11. The appropriate limit is currently set out in the Freedom of Information and Data Protection (Appropriate Limit and Fees) Regulations 2004 ("the Fees Regulations"). A public authority may take into account the cost of locating, retrieving and extracting the requested information in performing its calculation. The cost limit is currently set at £600 for central government. Under the Fees Regulations, a public authority is

required to cost their spending on the relevant activities at £25 per person per hour. Consequently, the appropriate limit would only be exceeded if it is estimated that it would take longer than 24 hours to carry out the relevant activities in order to comply with a request.

12. Under regulation 4(3) of the Fees Regulations, a public authority may, for the purposes of estimating the cost of complying with a request, only take account of the costs it reasonably expects to incur in:
 - a. determining whether it holds the information;
 - b. locating a document containing the information;
 - c. retrieving a document containing the information; and
 - d. extracting the information from a document containing it.
13. APHA informed the Commissioner that it had estimated that it would take a total of 27 hours to provide the complainant with the information that he requested. It explained that at the initial stage of the request, in determining whether the information was held, two members of staff spent time analysing a way it could provide this to the complainant. Its initial conclusion was that this could not be provided without making certain assumptions when joining datasets together, which therefore amounted to the information not being held. Consequently, APHA gave its response, dated 2 June 2016, that the requested information was not held.
14. APHA went on to explain that, on review, it spent more time analysing what had been requested and concluded the information was 'held' in accordance with FOIA but to provide it to the requester would exceed the appropriate limit, therefore engaging Section 12 of the Act. It confirmed that the total amount of time spent by staff analysing whether it held the requested information was 3 hours.
15. The Commissioner was informed by APHA that the relevant data is captured and stored in its internal computer system called SAM, but it is not structured in a way that facilitates a single extraction in order to answer the request. APHA explained that it did not routinely report on this data in the manner of the request, so there is no existing process for creating the requested dataset. The request additionally asked for a significant volume of detailed data. APHA stated that it would need to run an extract of TB test data for England for the 2 years in question; run a separate extract of TB reactor data for England for the 2 years in question; and join the data together using the primary key of the TB tests to arrive at an aggregated data test dataset. This would then need to be manually manipulated to extract the data in the required format.
16. APHA provided the Commissioner with the details below of the process that it would need to follow, broken down into the relevant activities.

Activity 1 – Extract TB Skin Test data from the SAM data warehouse

17. Run APHA's pre-canned 'SAM TB statistics - Cattle/Buffalo/Bison TB Test Stats CPH version' Business Objects report for 2014 and 2015. This report contains details of all TB skin tests completed in Great Britain in the calendar year, and contains geographical information, the CPHH for the herd, the test type, the overall result of the test (but not the number of TB reactors on the test), and the number of animals available to test. Each yearly instance of the report will need to be scheduled in the Business Objects reporting timetable, and would take approximately 3 hours per year to run. After that, manual extraction would need to take place which would take approximately 30 minutes. This makes a total of 6.5 hours.

Activity 2 - Extract TB Skin Test Reactor data from the SAM data warehouse

18. Run pre-canned 'Sam R6 - TB Reactor Details - Skin/Gamma Status' Business Objects report for 2014 and 2015. This report contains details of TB Skin Tests where reactors were found, containing 1 row for each reactor, and includes the test type and the CPHH for the herd. Each yearly instance of the report will need to be scheduled to run, which would take approximately 2 hours per year. Again, after that manual extraction would need to then take place which would take approximately 30 minutes. This makes a total of 4.5 hours

Activity 3 - Merge data sets from Business Objects reports in MS Excel to produce 2 distinct data sets

19. Manually manipulate TB skin test data to remove superfluous data and to merge part test data in order to identify herd size. Number of animals presented at test is recorded against each part-test so in order to determine herd size, this data would need to be merged and reduplicated. Manually manipulate reactor data to remove superfluous data and to aggregate to herd level. This would take approximately 10 hours.

Activity 4 - Load MS Excel data sets into MS Access and merge by CPHH to create a new relational database

20. This would take approximately 1 hour.

Activity 5 - Export merged data sets for each year into MS Excel and filter or pivot data to output in the requested format

21. This would take approximately 2 hours.

22. APHA confirmed that its total estimate for time spent in supplying the information to the complainant in the requested format would be 27 hours, which includes the 3 hours already spent on the request to establish whether the information was 'held'.
23. APHA explained that a sampling exercise was not carried out because the processes for producing a small quantity of data were similar to those needed for a large dataset. It stated that running the necessary reports from the system would be quicker, as the time period could be less. However, to produce the final figures sought in the original request, there was a significant amount of planning and data manipulation that would be required; for example, extracting the data on herd sizes, when they were last routinely tested, how many had reactors, then putting that number of reactors into brackets per thousand cattle tested and so on.
24. The Commissioner was informed by APHA that this part of the process would be the same no matter what period of data was tested. It was therefore believed that doing a sample set of data was not feasible. APHA explained that the estimates of timings provided above were based on a senior member of staff's experience working with large datasets and their knowledge of the various parts of the task that would be required to gather this data together. It confirmed that the activities described above were the quickest methods of gathering the information to supply it in the requested format.
25. The complainant was provided with details of APHA's explanation for the application of section 12. He explained that his profession was designing and managing databases and went on to state that:

"The following comments refer to Route 2 only in APHA's response because I am unfamiliar with the pre-canned system which APHA has set up and referred to in Route 1 of their response. If I understand APHA's arguments correctly, APHA considered Route 2 to be not feasible and that Route 1 requires 30 hours effort if the 3 hours analysing whether the information is held is included.

In their response regarding Route 2 they state that their estimates of timings are based on a senior member of staff's experience working with large datasets. I contend that the work will require the time of someone who is competent at querying a database and familiar with the databases involved and how they are structured. For someone with relevant training and qualifications that expertise should be built up over a period of no more than 2 years.

For well structured and maintained database(s) designing SQL queries with required joins should be very straightforward and simple, and for appropriately indexed large tables execution should not take more than a few minutes. The bulk of the filtering could be achieved efficiently by adding suitable clauses to the queries. Any remaining, more intricate filtering and formatting requirements could be achieved by executing a specially written small script written in a high level language such as ASP.NET. The results would then need to be checked for errors and the script updated and re-run accordingly. For a competent analyst this should not require more than a day's work."

26. The complainant explained that APHA had provided him with bespoke data in the past and provided an example of this.

27. In response to the complainant's comments, APHA stated that:

It seems [the complainant] may have misunderstood the processes required to produce the information requested. He refers to 'Route 1' and 'Route 2', which seems unclear in what is meant. We believe he refers to the 5 'Activities' described in our letter dated 12 September. This describes the challenge of extracting the underlying data and transforming it into the format requested. These 'Activities' are inclusive of one another, therefore all need to be completed. The total approximate time spent in supplying the information to [the complainant] in the requested format would be 27 hours, which includes the 3 hours already spent on the request to establish whether the information is 'held'. It would not be the 30 hours quoted.

The senior staff member within APHA was referred to in the same response for the purposes of experience of the system only, not for completing the actions. This is to give an approximate well judged time in dealing with the request. Regulation 4(4) Freedom of Information and Data Protection (Appropriate Limit and Fees) Regulations 2004, sets the limit of £25 per hour irrespective of the seniority.

The detailed data needed to fulfil [the complainant's] request will need to be extracted from the APHA IT system called Sam. The data structures of the Sam operational data stores and Sam Data Warehouse were designed by IBM incrementally over many years in response to APHA's detailed functional and reporting requirements. The Sam Data Warehouse is maintained and indexed by a Database Administrator employed by IBM under a commercial hosting agreement. It is acknowledged that the

structure of the Sam Data Warehouse delivers sub-optimal performance for reporting but the cost of developing subject area data marts has proved to be prohibitive. APHA specialists developed and maintain the semantic reporting layer (Business Objects Universe) which has been designed to ensure that the reporting requirements of policy customers (Defra, Welsh Government & Scottish Government) and APHA's own operational teams, can be met.

APHA uses SAP Business Objects (BO) as its business intelligence software suite. The BO components are used by APHA specialists to deliver the reporting requirements of APHA and policy customers. Business Objects reports run underlying SQL queries to extract data and these are customised by APHA specialists to ensure business requirements can be delivered. The pre-canned reports referred to in our response to you, dated 12 September 2016, take several hours to run due to the complexity of the SQL queries, the data structure and the volume of data involved.

APHA considered developing bespoke SQL queries to extract the data. Our knowledge and experience of the data and the network infrastructure, led us to believe that this would cause disruption to our business customers. They would be unable to run or schedule other reports whilst the queries were running. The solution we proposed, whilst time consuming, would minimise the impact on APHA BO users. The BO software used by APHA (Enterprise Edition 3.1) is now out of support and this adds to the challenges we would experience. Defra, APHA's ministerial department, is not planning to upgrade to the latest supported version of BO and is exploring the use of other business intelligence tools such as Power BI, Qlik and Tableau.

With regards to [the complainant's] comments on previous FOI/EIR requests to APHA, they have all been responded to in accordance with the relevant legislation. [The complainant] has made frequent requests to APHA over the years, and the majority of them have involved complex data that has taken several hours to produce. This has caused a level of disproportionate or unjustified disruption. However, we have always maintained our duty to be open and transparent in accordance with the legislation. When release of information to [the complainant] has been given, in some cases the data requested has taken close to the cost limit to supply."

28. Commenting on paragraphs 4 and 5 of APHA's response (above), the complainant stated that:

"I interpret this to mean that the pre-canned reports would take several hours to run and running bespoke SQL queries would cause unacceptable disruption to other customers.

Let me first comment on the pre-canned reports. It should be possible to measure run times in seconds and not hours. This is provided the following applies.

- *Table columns which are joined are properly indexed.*
- *Queries are kept simple through appropriate use of temporary tables to store data between queries.*

I contend that APHA's ability to respond to customer requirements is being crippled by grossly inefficient processing. Simple queries consisting of carefully crafted joins of properly indexed SQL tables inserting into appropriate, temporary tables typically filter through millions of records in the first step in seconds. If the SQL queries are complex and the tables are not appropriately indexed for columns being joined, this will have a huge impact on performance and may well lead to run times measured in hours. If pre-canned reports are taking this long to run it sounds like it is necessary to develop bespoke SQL queries to avoid the bottlenecks. This should then give APHA the performance which they need and due to much quicker run times would avoid disrupting other users. This assumes that tables are structured properly and a competent administrator (wherever he/she is employed) has indexed the columns which are used in joins."

29. The APHA's view of the complainant's comments were that:

"[The complainant's] arguments are his views on the "grossly inefficient processing" of our system, and this may not be his choice for such processing. This does not, however, rescind the factual aspects on how it actually does perform, and we cannot simply speed up the processing.

In accordance with Section 12 of the FOIA, and in line with Regulation 4(3) of the Fees Regulations, APHA have explained the costs it reasonably expects to incur in carrying out the activities to respond to the request. These are set out in our letters dated 12 September and 7 October 2016.

APHA can only work with the resources that are available, and in this case the processes have been broken down into detailed time allocations. Any estimate must be reasonable in the circumstances of the case, and APHA feel that cogent oral

evidence has been given to address this by a member of staff experienced in working with the system on a day to day basis.

So in conclusion APHA feel that [the complainant] has not addressed the application of Section 12, but instead given his views on the performance of our system.

We therefore still maintain the use of Section 12 of the FOIA, in that the cost of complying with this particular request exceeds the appropriate limit and would cause a strain and disruption on APHA's everyday business objectives."

30. The Commissioner notes the complainant's concerns over the potential inefficiency of the computer system operated by the APHA which holds the information that he is seeking to obtain. However, under the Act, she is not in a position to require public authorities to put in place different, potentially more efficient, computer systems or to assess whether a public authority might be able to provide information without exceeding the appropriate limit under section 12 if it had potentially better computer systems in place. She is limited to determining whether, a public authority's estimate for the purposes of section 12 is reasonable, given the computer systems that it actually operates.
31. The Commissioner has reviewed the explanation provided by APHA of the processes that it would need to follow to provide the information requested and the amount of time that this would take. She is satisfied that, based on this explanation, it is likely to take in excess of 24 hours to provide the information requested by the complainant. She is therefore of the view that it was reasonable for APHA to estimate that responding to the request would have exceeded the appropriate limit and has consequently decided that it has correctly applied section 12 to the request.
32. Where a public authority has applied section 12 to a request, the Commissioner would expect it to consider whether it would be reasonable to provide advice and assistance to a requester under section 16 with regard to whether some information could be provided within the cost ceiling.
33. APHA informed the Commissioner that it was aware of section 16 and the duty on a public authority to provide advice and assistance to try to help the applicant refine their request so that it fell within the appropriate limit in Section 12. However, it believed the explanation that it had provided as to why it had not carried out a sampling exercise, that the processes for producing a small quantity of data were similar to those needed for a large dataset, was also applicable to the reason why it had not provided advice and assistance to the

complainant. It was of the view that any suggestion made in order to narrow down the request, for example reducing the period of time that it covered, would still have resulted in an estimate that responding to the request would exceed the appropriate limit.

34. In light of APHA's explanation, the Commissioner is satisfied that it has not breached its obligations under section 16 of the Act.

Right of appeal

35. Either party has the right to appeal against this decision notice to the First-tier Tribunal (Information Rights). Information about the appeals process may be obtained from:

First-tier Tribunal (Information Rights)
GRC & GRP Tribunals,
PO Box 9300,
LEICESTER,
LE1 8DJ

Tel: 0300 1234504

Fax: 0870 739 5836

Email: GRC@hmcts.gsi.gov.uk

Website: www.justice.gov.uk/tribunals/general-regulatory-chamber

36. If you wish to appeal against a decision notice, you can obtain information on how to appeal along with the relevant forms from the Information Tribunal website.
37. Any Notice of Appeal should be served on the Tribunal within 28 (calendar) days of the date on which this decision notice is sent.

Signed

Rachael Cragg
Group Manager
Information Commissioner's Office
Wycliffe House
Water Lane
Wilmslow
Cheshire
SK9 5AF