Reference: 2025-066

Thank you for your email in which you requested the following information under the Freedom of Information Act 2000 (FOIA):

I am writing to request further information regarding the review of cases that used the Autonomy Introspect disclosure tool.

Please provide the following:
1. Copies of any or all documents or instructions given to those reviewing cases, for example details of the process to be followed.
2. Copies of any or all internal handbooks or similar on use of disclosure tools more broadly
3. In how many cases were term searches rerun as part of the review?
4. Details of the seniority of the staff that have reviewed the cases.

I am also seeking information about OpenText Axcelerate.
5. Which version of OpenText Axcelerate does the SFO run
6. Please detail the cost of the tool.

Response

Question 1
We have identified the documents you have requested. We consider that these documents attract the exemption under section 42 of the Freedom of Information Act 2000 (legal professional privilege). This is an absolute exemption and therefore does not require consideration of the public interest test.

However, while the documents themselves are exempt from disclosure, we recognise the public interest in providing some information about the review process. Therefore, we have provided you with some non-privileged information which explains how these reviews were conducted. This information has been shared with defendants and their legal teams already and is provided at Annex 1.

Question 2
In response to our request for clarification on 16 June, you stated: “I can confirm I am looking for both internal manuals on how Autonomy was used and current manuals providing guidance on using Axcelerate.”

We confirm that we hold this information. However, this information is exempt from release under sections 31(1)(a), (b), and (c) and section 43 of the FOIA.

Section 31 provides that: *Information which is not exempt information by virtue of section 30 is exempt information if its disclosure under this Act would, or would be likely to, prejudice—*
(a) the prevention or detection of crime,
(b) the apprehension or prosecution of offenders,
(c) the administration of justice:

How the exemption is engaged

Section 31 (a) (b) and (c) permit the exemption of information from release when the disclosure of the requested information may impact investigation and prosecution of criminal cases and the administration of justice.

As a law enforcement agency, the Serious Fraud Office (SFO) holds highly sensitive information which is of interest to others, including those we investigate. Releasing detailed information about how we approach the review of data and our strategy for investigating documents is highly sensitive and, if released to the public, could provide information which criminals could use to avoid detection and loopholes in the way in which we search material. Therefore, releasing detailed information about how we perform these searches could undermine the SFO's ability to protect our investigative techniques and information systems, thereby prejudicing the interests at (a), (b), and (c) above.

Section 43(2) provides that: *Information is exempt information if its disclosure under this Act would, or would be likely to, prejudice the commercial interests of any person (including the public authority holding it).*

How the exemption is engaged

Axcelerate is a product that is currently on the market and releasing detailed information about how it can be used and information provided to us by OpenText could undermine its commercial interests.

Public interest test

The above-referenced sections are qualified exemptions and require consideration of whether, in all the circumstances of the case, the public interest in exempting this data outweighs the public interest in disclosing it. More information about exemptions in general and the public interest test is available on the ICO's website at www.ico.org.uk.

It is recognised that there is a general public interest in publicising information about the SFO, so that the public knows that serious fraud, bribery, and corruption are being investigated and prosecuted effectively, and so that the public can be reassured about the general conduct of our organisation and how public money is spent. The SFO already takes steps to meet this interest by publishing our Annual Report and Accounts and through our ICO-aligned Publication Scheme: SFO publication scheme - GOV.UK.

However, having considered the public interest in releasing this information, we consider that the stronger interest lies in maintaining these exemptions of the FOIA. We understand the importance of publicising information about the SFO's use of public resources and funding.

However, the risk that this information could pose to the SFO's law enforcement functions against which it is essential to safeguard given the highly sensitive nature of the information held as a law enforcement agency. The SFO is a relatively small, highly specialised government department that is permitted by law to investigate only the most serious and complex cases of fraud and bribery affecting the UK. Releasing sensitive information about the way in which we conduct disclosure could risk compromising the SFO's ability to provide

and maintain data security for the cases at pre-investigation, investigation, prosecution stages and beyond.

Additionally, we do not believe that the release of those information which we hold would add significantly to these factors and therefore feel there is minimal public interest in releasing this material. Further, as outlined above, the SFO is compliant with the reporting requirements across government, which requires the publication of data surrounding procurement processes on the Contracts Finder website: Contracts Finder - GOV.UK.

Having considered the opposing arguments, we believe that the stronger public interest lies in exempting the information from release.

Question 3
At the date of your request, search terms have been re-run as part of the review on four cases. In all four cases, no material which might cast doubt upon the safety of the conviction was identified.

Question 4
All stage 1 and 2 reviews were conducted by an individual at civil service grade 6. Where a case has a decision-maker at stage 3, that individual is a lawyer of at least civil service grade 7.

Question 5
The SFO runs v23.4_430 of OpenText Axcelerate.

Question 6
The SFO paid a one-time cost for OpenText Axcelerate of £3,348,834.37.

The SFO pays £396,000 annually for renewing licence and maintenance. Additionally, the SFO pays monthly costs of around £60,000-90,000 for training and support. These payments are published in our transparency documents via our publication scheme.

Annex 1

Background to the review

The SFO, like many organisations faced with the challenges of managing vast quantities of digital material, utilises e-Discovery software. In the SFO's case its former e-Discovery platform was Autonomy, and this was used on a number of cases to store and review digital material obtained during an investigation.

As is the case with other e-Discovery platforms, operators of Autonomy could use search terms to find documents that contained specific passages, phrases, words or other identifiers. In the SFO's case, search terms were commonly used to identify evidence and relevant material (within the meaning of the Criminal Procedure and Investigations Act 1996 ("CPIA 1996")) for use in criminal proceedings.

A feature of the SFO's version of Autonomy was that when search terms were run, certain characters were treated as 'tangible' rather than 'non-tangible' ("the punctuation feature"). The characters in question were:

- % - Percentage symbol
- @ - At symbol
- / - Forward slash
- : - Colon
- . – Full stop
- £ - Pound sign

Hereafter these will be referred to as "the affected characters".

Tangible characters are ones which a search function looks for, whilst non-tangible characters are ones which are ignored. The result, as far as the SFO's version of Autonomy was concerned, was that where one of the affected characters appeared directly adjacent to a term that was being searched for, this occurrence of the term would not be found by a search for the term alone.

This is best illustrated with an example. If a search was run for the word "bribe", then Autonomy's search function would identify all documents where the five letters 'B', 'R', 'I', 'B' and 'E' appeared in the document in that order and with no other adjacent tangible characters. However, Autonomy's search function would not identify any documents where those five letters only appeared adjacent to other non-tangible characters. In other words, Autonomy would find all documents where the word "bribe" was present, but it would not, for example, find documents which only contained words such as "bribe." (i.e. with a full stop after the word bribe) or "bribes".

Instead, to account for instances where the word being searched for appeared directly adjacent to other tangible characters, including the affected characters, the user of Autonomy could either include additional characters in the search term (i.e. they could specifically search for "bribes" or "bribe.") or they could employ one of two wildcard characters:

- * - the asterisk
- ? – the question mark

The asterisk wildcard matched zero or more characters after the point at which it appeared in a search term. So, for example, if the term “bribe” was instead rendered as “bribe*” then Autonomy would return all documents where the letters ‘B’, ‘R’, ‘I’, ‘B’ and ‘E’ appeared in a document in that order and with any number of tangible characters after them (including the affected characters). That would mean in practice that the search term “bribe*” would return:

- all documents where the word “bribe” was present
- all documents where the word “bribe” was present followed by one (or more) of the affected characters (e.g. “bribe.”); and
- all documents where “bribe” is the stem of another word. For example, in the words:
    - Bribes
    - Bribed
    - Bribery

The question mark wildcard worked by returning documents where a single tangible character replaced the question mark. So, for example, if the term “bribe” was instead rendered as “bribe?” then Autonomy would return all instances where the letters ‘B’, ‘R’, ‘I’, ‘B’ and ‘E’ appeared in a document in that order and were followed by a single additional tangible character.

Significantly, the question mark wildcard would identify all the affected characters except the colon. That would mean in practice the search term “bribe?” would return:

- all documents where the word “bribe” was present followed by one of the affected characters, except the colon, (e.g. “bribe.”); and
- all documents where the word ‘bribe” was present followed by another single character. For example, in the words:
    - Bribes
    - Bribed

A question mark wildcard could also be placed in the middle of a term. So, for example if the term “bribe” was instead rendered as “bri?e”, then Autonomy would return all instances where the letters ‘B’, ‘R’ and ‘I’ were followed by any tangible character and that tangible character was then followed by the letter ‘E’. That would mean in practice this term would return documents which contained words such as: “bribe”, “brine” and “bride”.

If necessary, multiple question mark wildcards could be used, and where this was done the number of question marks would dictate how many additional characters would be matched. For example, if the search term “bribe??” was used then it would return all documents where the letters ‘B’, ‘R’, ‘I’, ‘B’ and ‘E’ appeared in that order followed by two additional characters. For example, in the word ‘bribery’. Likewise, the term “br??e” would return all documents where the letters ‘B’ and ‘R’ appeared followed by any two tangible characters and then the letter ‘E’. That would mean in practice this term would return documents which contained the words such as: “bribe”, “brine”, “bride”, “brave”, “brake”, “brace” and so on.

Two other factors are relevant when considering how Autonomy’s search function worked:

- First, for a document to be identified it only needed a term to appear once without adjacent characters. The presence of versions of the term adjacent to other tangible characters did not affect this. For example, if a search was run for the word “bribe” then Autonomy would identify all documents where that word appeared, irrespective of whether the document also contained instances where the word appeared as, for example, “bribes” or “bribe.”

- Secondly, multiple search terms could be run over the same pool of material simultaneously by combining two or more terms together with Boolean operators[1] to create a search string. It was also possible to run multiple strings at the same time. If the same document hit multiple strings, or parts of the same string, it would only appear in the search results once. This allowed operators to devise terms that intentionally overlapped (i.e. terms could be designed with minor differences with the specific intention of identifying the same type of material) without having to worry about the same documents being identified multiple times.

The SFO’s guidance and training on how to operate Autonomy took the punctuation feature into account and explained how to use wildcards. This was done so that those involved in creating search terms could specifically account for the punctuation feature if that was considered necessary in light of the volume and nature of material that needed to be reviewed and/or the way in which Autonomy was going to be used to search that material. Case teams were also encouraged to use overlapping search terms to reduce the chance that material was missed. However, despite this guidance and training it subsequently transpired that in two unrelated SFO cases the case teams were unaware of the need to account for the punctuation feature and so the search terms used on those cases were not designed with it in mind.

The SFO’s position

As the Attorney General’s Guidelines on Disclosure have long made clear, the prosecution is not required to review every individual item of electronic material in its possession in order to comply with its disclosure obligations. Instead, its obligation is to undertake reasonable lines of inquiry, including those which point away from the suspect, which can be satisfied by using appropriate search terms to locate relevant material. A consequence of this is that there is always the possibility that in cases where the volume of material is such that the use of search terms is required, the prosecution may be in possession of relevant material which remains unidentified.

This is significant as relevant material might in turn contain items which, either individually or cumulatively, have the potential to undermine the prosecution case or assist the defence case. The use of search terms therefore always has the potential to create a situation whereby material which satisfies the disclosure test is not identified. It is also worth noting that material adverse to the accused may also be missed through the use of search terms. It follows from the fact that the use of search terms is an approved methodology, that it is an accepted part of criminal litigation in cases which involve large volumes of digital material that there is always some risk involved in the use of search terms.

[1] Boolean operators are a set of commands that can be used in a search engine to set additional parameters for a search and thereby expand, limit or narrow the results of that search. For example, the operator “AND” can be used to join to two separate search terms and identify a document which contains both terms.

However, the SFO recognises that if a case team was unaware of the punctuation feature, and so inadvertently used search terms which did not account for it, then notwithstanding the overlapping nature of search terms, the risk of relevant material being missed in that case may be greater than it ordinarily would be because the approach adopted will not have been designed with full knowledge of the limitations associated with it.

As a result, and mindful of its post-conviction disclosure obligation to actively undertake further enquiries if there is a real prospect that those enquiries may reveal something affecting the safety of a conviction (as set out in *R (on the application of Nunn) v Chief Constable of Suffolk Constabulary and anor* [2014] UKSC 37), the SFO decided it was necessary to look back at all Autonomy cases where a conviction was obtained to determine, as far as possible, whether search terms were used for disclosure purposes and if they were, whether the punctuation feature was accounted for.

If the feature was not accounted for, or the review is inconclusive, then the SFO, will go on to consider, on a case-by-case basis, and with reference to the real issues in the case as they were presented during the trial, whether the duty to undertake additional enquiries necessitates further steps being taken to try and identify whether material which might affect the safety of the conviction exists. Having regard to the dicta of Lord Hughes JSC at paragraph 33 of *Nunn* that the "*contest*" for finite resources should favour current investigations over the re-investigation of concluded cases, such steps would need to be reasonable and proportionate, and so if it is decided that it is necessary to re-run search terms then this is likely to entail only re-running a selection of search terms (amended to account for the punctuation feature) unless a good reason is established to undertake a more extensive exercise.

If search terms are re-run, then the product of those searches will be reviewed to identify any material which was not identified before the trial. Then, in accordance with the second aspect of the post-conviction disclosure obligation, the SFO will consider whether any such material is capable of casting doubt upon the safety of a conviction.

As Lord Hughes observed at paragraph 35 of *Nunn*, the second aspect of the post-conviction disclosure obligation requires disclosure of "*something new which might afford arguable grounds for contending that the conviction was unsafe*", and so not all information which may come to light after a conviction therefore needs to be disclosed. Instead the SFO will consider what the consequences of any failure to account for the punctuation feature were. When making this assessment the SFO's position is:

- First, any additional material which is considered to be neutral or adverse to a Defendant, (judged in relation to the issues as they were presented during the trial) is unlikely to afford arguable grounds for contending a conviction was unsafe;
- Secondly, the fact that additional material might have satisfied the test for disclosure pre-trial, does not automatically mean that such material must be disclosed post-conviction. This is apparent from the fact that the post-conviction disclosure test is different to the statutory test set out in sections 3 and 7A of the CPIA 1996. In particular, as Lord Hughes explained at paragraph 32 of *Nunn*, the post-conviction duty is a common law duty which consciously recognises the difference in the position of a defendant on trial (who is presumed to be innocent until proven guilty) and a convicted defendant (who is presumed guilty unless and until it is demonstrated that his conviction is unsafe) and is founded on the public interest in the finality of proceedings.

What this means in practice in respect of material which would have met the test for disclosure pre-trial, is that the SFO will go on to further asses it to determine whether it provides a new evidential basis to suggest the defendant is, or may be innocent, or demonstrates that there was some serious procedural irregularity or unfairness which could have rendered the conviction unsafe. This may involve consideration of factors such as

  - Whether the material raises any new issues which were not identified before the trial;
  - the capacity of the material to support new legal arguments or submissions which a defendant may have wanted to make; and/or
  - whether the material demonstrates that the defendant was not able to present their case, or part of it, including any legal argument or submission, in its best light.

- Thirdly, above all, the SFO must be mindful of the defence perspective and where there is doubt about whether material should be disclosed, that should be resolved in favour of the defendant

<u>Overview of SFO review process</u>

The review process has three parts:

- Stage 1 – Preliminary Review - Every case is subject to a preliminary review which focuses on identifying a number of key contemporaneous documents which it is believed are likely to record and/or describe the disclosure process employed on the case and in particular the approach taken towards the creation of search terms. Those documents, where available, are then reviewed and an assessment made whether the approach adopted on the case accounted for the punctuation feature.

  If the documents make explicit reference to the punctuation feature, or to following written guidance on the use of autonomy, then the review of the case will likely conclude at this point. This is because the purpose of the review is simply to consider whether the approach adopted in relation to disclosure accounted for the punctuation feature, it is not a reconsideration of the approach itself. A review may also be concluded at this point if the documents reveal that the approach taken towards disclosure did not involve the use of search terms.

  If the documents are inconclusive, then, if they include details of the search terms used on the case, the terms themselves will be considered to determine whether they were demonstrably designed to account for the punctuation feature. If the assessment is that search terms were specifically designed with the punctuation feature in mind, or the way they were designed would have accounted for the punctuation feature (for example, because of the way they used wildcards), then again the review will likely conclude at this point.

- Stage 2 – Further Review - If there is no contemporaneous record, and/or it is unclear whether the search terms used were designed with the punctuation feature in mind and/or would have accounted for it, then a case will be subject to a more detailed review. The purpose of this review is still to ascertain whether a case was in fact affected by the punctuation issue, but instead of focusing on key documents it will take more involved steps to consider the approach adopted on the case. The precise steps taken as part of this review will vary from case to case, but typically

they will include identifying and reviewing other contemporaneous documentation and where possible discussing the case with extant SFO staff.

Where, following a further review, it is possible to conclude that search terms were designed to account for the punctuation issue then no further steps will be taken in respect of the case. If this review is inconclusive, or demonstrates that the case team were unaware of the punctuation feature, then the case will proceed to a stage 3 review.

- Stage 3 – Impact assessment – For each case that reaches this point a decision maker, who is a senior lawyer, will be appointed whose role will be:
    - First, to determine whether on the information available as a result of the stage 1 and 2 reviews there is a real prospect that any further enquiries may reveal something affecting the safety of any convictions.

      If there is a real prospect, then they will go on to decide what further enquiries should be made, arrange for those enquiries to be carried out in a reasonable and proportionate way, and subsequently review the results of those enquiries. This process will be repeated until they are satisfied that no further enquiries are necessary or all reasonable enquiries have been exhausted
    - Secondly, once a decision has been made that no further enquiries are necessary or all reasonable enquiries have been exhausted, they will consider whether they are now in possession of any additional information and/or material which might cast doubt upon the safety of the conviction.

      Where the further enquiries have led to additional material being identified by the use of amended search terms, this will entail considering the significance/status of that material and making an assessment of its impact upon the safety of the conviction.
    - Thirdly, irrespective of whether additional information and/or material which might cast doubt upon the safety of the conviction is found, they will disclose to affected defendants the nature of the review conducted, its conclusion, and the reasons for that conclusion. If the review identifies material which might cast doubt on the safety of the conviction, that material will also be provided to the affected defendant