

Annex E. The analysis on the coding fixed before collection, and the correction record

The final analysis applies the correction record to the original record; the analysis on the coding as frozen, which the protocol fixed, is kept unchanged beside it (Section 3.7). This annex prints its headline results whole. The protocol's canonical sentence, filled by the analysis program on the coding fixed before collection, reads:

Of the 21 providers in this audit's ranked frame whose Article 53(1)(d) duty had already fallen due at our read, 7 published a training-content summary that served live and that we could reach without an account. Of the 2 providers whose duty does not fall due until 2 August 2027 under Article 111(3), 0 published one. For 3 providers the documents did not settle when the designated model became publicly available, and those are reported separately. A further 0 providers are not in any of those three groups, because the dependency rules removed the Article 53(1)(d) variables from their records.

In that sentence, "the designated model" is the model studied. Table E1 is Table 5 on that coding. Table E2 gives the six sensitivity analyses on it. Table E3 lists every figure that differs between the two analyses, found by program, and Table E4 the correction record whole.

Table E1. Training-content summaries, by deadline group, on the coding fixed before collection.

What is published and reachable, as read on the dates in each row. This study does not determine compliance.

What we found	Deadline had passed (n = 21)	Not yet due, deadline 2 August 2027 (n = 2)	Release date not settled (n = 3)
Published, identified as the summary	7	0	0
Related document only	6	2	1
Summary mentioned, with no location a reader can reach	1	0	0
The provider states it does not publish this	0	0	1
Nothing found in the places we looked	6	0	1
Could not be opened	1	0	0
Removed by a rule	0	0	0

Note: each cell counts providers; each column is one deadline group. Source: the analysis program's frequency output on the frozen coding.

Table E2. The sensitivity analyses fixed in the protocol, on the coding fixed before collection.

What is published and reachable, as read on the dates in each row. This study does not determine compliance.

Analysis fixed in advance	Figure it recomputes	Result
Index items weighted equally	The index	Differs by more than 5 points for 3 of 26 providers
Every statement of non-publication scored zero	The index	Differs by more than 5 points for 1 of 26 providers
Signatories of the code of practice only	Identified summaries, deadline passed	Signatories: 3 of 5 (60.0%); non-signatories: 4 of 16 (25.0%); signatory status not known: 0 providers; removed by a rule: 0 and 0
Providers whose deadline had passed only	Index bands	Full public record 7 of 21; Substantial public record 3 of 21; Partial public record 1 of 21; Thin public record 2 of 21; Minimal public record 8 of 21
Most favorable reading: answers we could not open count as found	Identified summaries, deadline passed	7 of 21 as published; 8 of 21 on this reading, counting 1 answer that could not be opened
Release dates assigned from a period, such as a season	Deadline groups	0 records carry such a date, so the check is untested

Note: none replaces a published figure or reorders a table. The index rows rest on a flagged item (less certain: our two readings agreed below a threshold set in advance on the item this rests on; Section 4.11). Source: the analysis program's sensitivity outputs on the frozen coding.

Table E3. Every figure that differs between the analysis on the coding fixed before collection and the final analysis.

Figure	Original analysis (coding fixed before collection)	Final analysis
Training-content summary: related document only (deadline had passed)	6 of 21	4 of 21
Training-content summary: published, identified as the summary (deadline had passed)	7 of 21	11 of 21
Where the summary or related document sits on the provider's site: no document to read (deadline had passed)	7 of 21	6 of 21
Where the summary or related document sits on the provider's site: could not be determined (the crawl was not completed, or not run to a document reached after the window) (deadline had passed)	7 of 21	9 of 21
The document names the model it covers: names specific models (deadline had passed)	12 of 21	13 of 21
The document names the model it covers: no document to read (deadline had passed)	7 of 21	6 of 21
The document states the kinds of training content: no document to read (deadline had passed)	7 of 21	6 of 21
The document states the kinds of training content: lists them (deadline had passed)	10 of 21	12 of 21
The document states the categories of sources: no document to read (deadline had passed)	7 of 21	6 of 21
The document states the categories of sources: lists them (deadline had passed)	11 of 21	13 of 21
The document names particular sources: gives examples of kinds of source only (deadline had passed)	2 of 21	3 of 21
The document names particular sources: names datasets (deadline had passed)	9 of 21	10 of 21
The document names particular sources: no document to read (deadline had passed)	7 of 21	6 of 21
The document refers to the AI Office template: no reference to the template (deadline had passed)	11 of 21	13 of 21
The document refers to the AI Office template: no document to read (deadline had passed)	7 of 21	6 of 21
The document carries a date or a version: dated and versioned (deadline had passed)	10 of 21	11 of 21
The document carries a date or a version: no document to read (deadline had passed)	7 of 21	6 of 21
The document carries a date or a version: neither dated nor versioned (deadline had passed)	1 of 21	2 of 21
Statement on the provider's own web crawling for training: nothing found in the places we looked (all providers)	11 of 26	10 of 26
Statement on the provider's own web crawling for training: web collection stated, crawler not named (all providers)	8 of 26	9 of 26
Training-content summary: summary mentioned, with no location a reader can reach (deadline had passed)	1 of 21	0 of 21
Training-content summary: could not be opened (deadline had passed)	1 of 21	0 of 21
Where the summary or related document sits on the provider's site: could not be opened (deadline had passed)	1 of 21	0 of 21
The document names the model it covers: could not be opened (deadline had passed)	1 of 21	0 of 21
The document states the kinds of training content: could not be opened (deadline had passed)	1 of 21	0 of 21
The document states the categories of sources: could not be opened (deadline had passed)	1 of 21	0 of 21

Figure	Original analysis (coding fixed before collection)	Final analysis
The document names particular sources: could not be opened (deadline had passed)	1 of 21	0 of 21
The document refers to the AI Office template: could not be opened (deadline had passed)	1 of 21	0 of 21
The document carries a date or a version: could not be opened (deadline had passed)	1 of 21	0 of 21
The document names the model it covers: covers the provider in general (deadline had passed)	0 of 21	1 of 21
Served live when read (the documents of Qwen and Tencent on 15 October 2026; Tencent's file opened by the AI-assisted re-check, the identical file saved by the study on 16 October 2026), of those with a document address (deadline had passed)	13 of 14	15 of 15
Not retrieved by our reading, live or from an archive copy, of those with a document address (deadline had passed)	1 of 14	0 of 15
Located documents, deadline had passed: could not be determined (the crawl was not completed, or not run to a document reached after the window)	7 of 13	9 of 15
Located documents, deadline had passed: not reached within three clicks of the home page	2 of 13	2 of 15
Index band: full public record (deadline had passed)	7 of 21	10 of 21
Index band: substantial public record (deadline had passed)	3 of 21	1 of 21
Index band: partial public record (deadline had passed)	1 of 21	2 of 21
Index band: minimal public record (deadline had passed)	8 of 21	6 of 21
Index of Microsoft	77.0 (substantial public record; 0 not determinable)	90.0 (full public record; 0 not determinable)
Index of Qwen	5.0 (minimal public record; 0 not determinable)	57.0 (partial public record; 0 not determinable)
Index of Tencent	0.0 (minimal public record; 7 not determinable)	90.0 (full public record; 0 not determinable)
Index of Thinking Machines Lab	77.0 (substantial public record; 0 not determinable)	90.0 (full public record; 0 not determinable)
Answers not settled, all providers (including answers that could not be opened)	76	70

Note: rows are read by program from the two analyses' frequency, archive, findability, band and index outputs; a value with no row in one analysis counts zero there. The index rows rest on a flagged item (less certain: our two readings agreed below a threshold set in advance on the item this rests on; Section 4.11). Source: the analysis program's outputs on both sets.

Table E4. The correction record: every value the final analysis applies in place of an original one.

Company	Item	Original record	Value the final analysis applies	Under
Cohere	A page about the EU AI Act	Mentioned, with no location a reader can reach	Could not be determined	deviation 33
Google	Statement on the provider's own web crawling for training	Nothing found in the places we looked	Web collection stated, crawler not named	deviation 34
Google	Own reading: marking as far as technically feasible (Article 50(2))	Nothing found in the places we looked	"we are concerned that adding more regulatory complexity — as technical solutions are still evolving — could contradict Europe's goals for competitiveness and simplification"	deviation 34
Microsoft	Training-content summary	Related document only	Published, identified as the summary	deviation 36
Microsoft	The document names the model it covers	Names specific models	Names specific models	deviation 36
Thinking Machines Lab	Training-content summary	Related document only	Published, identified as the summary	deviation 36
Thinking Machines Lab	The document names the model it covers	Names specific models	Names specific models	deviation 36
Tencent	Training-content summary	Could not be opened	Published, identified as the summary	deviation 36
Tencent	The document names the model it covers	Could not be opened	Names specific models	deviation 36
Tencent	Summary or related document served live when read	Not retrieved by our reading, live or from an archive copy	Served live on 15 October 2026 (opened by the AI-assisted re-check; the study saved the identical file on 16 October 2026)	deviation 36
Tencent	The document states the kinds of training content	Could not be opened	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38

Company	Item	Original record	Value the final analysis applies	Under
Tencent	The document states the categories of sources	Could not be opened	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Tencent	The document names particular sources	Could not be opened	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Tencent	The document refers to the AI Office template	Could not be opened	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Tencent	The document carries a date or a version	Could not be opened	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Qwen	Training-content summary	Summary mentioned, with no location a reader can reach	Published, identified as the summary	deviation 36
Qwen	The document names the model it covers	No document to read	Names a model family	deviation 36; replaced under deviation 40
Qwen	The document states the kinds of training content	No document to read	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Qwen	The document states the categories of sources	No document to read	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Qwen	The document names particular sources	No document to read	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Qwen	The document refers to the AI Office template	No document to read	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Qwen	The document carries a date or a version	No document to read	Could not be determined (the document was not coded)	deviation 36; replaced under deviation 38
Qwen	Summary or related document served live when read	Not asked, because a rule removed the question	Served live on 15 October 2026 (saved by the study's reader)	the note on deviation 36
Tencent	Where the summary or related document sits on the provider's site	Could not be opened	Could not be determined (the link crawl was not run to the document)	the note on deviation 36
Qwen	Where the summary or related document sits on the provider's site	No document to read	Could not be determined (the link crawl was not run to the document)	the note on deviation 36
Tencent	The document states the kinds of training content	Could not be opened	Lists them	deviation 38
Tencent	The document states the categories of sources	Could not be opened	Lists them	deviation 38
Tencent	The document names particular sources	Could not be opened	Names datasets	deviation 38
Tencent	The document refers to the AI Office template	Could not be opened	No reference to the template	deviation 38
Tencent	The document carries a date or a version	Could not be opened	Dated and versioned	deviation 38
Qwen	The document states the kinds of training content	No document to read	Lists them	deviation 38
Qwen	The document states the categories of sources	No document to read	Lists them	deviation 38
Qwen	The document names particular sources	No document to read	Gives examples of kinds of source only	deviation 38
Qwen	The document refers to the AI Office template	No document to read	No reference to the template	deviation 38
Qwen	The document carries a date or a version	No document to read	Neither dated nor versioned	deviation 38

Company	Item	Original record	Value the final analysis applies	Under
Sakana AI	Earliest documented date of outside access to the model	11 September 2026	24 April 2026	deviation 39
Qwen	The document names the model it covers	No document to read	Covers the provider in general	deviation 40

Note: rows are in the order of the correction record; the original record is never edited; the correction record carries both values, and the release's correction-evidence file gives the source of each. Source: the final analysis's corrections file.