

Consultation on the ICO's ExplAIn draft guidance

The ICO and The Alan Turing Institute are consulting on our co-badged ExplAIn guidance. This guidance aims to give organisations practical advice to help explain the processes, services and decisions delivered by AI, to the individuals affected by them.

We are looking for a wide range of views from organisations across all sectors and sizes.

The guidance is comprised of three parts. Depending on your level of expertise, and the make-up of your organisation, some parts may be more relevant to you than others. You can pick and choose the parts that are most useful.

You can answer as many or as few of the questions as you want to. You can also save your progress and return at a later date.

If you would like further information about the consultation, please email explain@ico.org.uk.

Please send us your response by 17:00 on Friday 24 January 2020. You can email it to explain@ico.org.uk or complete the [online version of this survey](#).

Privacy statement

Please note, your responses to this survey will be used to help us with our work on explainability only. The information will not be used to consider any regulatory action, and you may respond anonymously should you wish. For more information about what we do with personal data see our [privacy notice](#).

Q1. Does the guidance provide what your organisation needs when considering how to explain AI-enabled decision to individuals?

Yes

No

Please provide further detail:

We would question why this is entirely focussed on the views of organisations doing AI and appears not to wish to know the opinions of the subjects on whom AI is/may be practised

The documents could more upfront and indeed upbeat about the point that the big data sets used in AI, where there is no realistic chance of identifying any individual, have the potential to deliver huge benefits and the drive should be to "get on and do it".

That will need to keep the public on board, and our members have noted the apparent irony in having a consultation about communicating with the public where the public "don't get a look in".

On balance, we felt the guidance could be perceived by the public as taking too much of a patronising attitude towards them as "customers". The guidance should not be presented in a way which presumes that it is about how to explain issues which are way beyond the understanding of patients.

Part A specifically says that organisations need to consider "the benefits and risks of explaining AI systems to the individuals affected by their use", but makes little recognition that the process of "explanation" should equally involve the views of those whose data has been used to undertake this work, as well as those likely to be affected.

We would really emphasise the fact that this should start with understanding what the public wants to know about AI applications that affect them. We understand that would be a consultation in itself but it would have been useful to summarise the existing research on public wants and concerns so that your professional consultees could be better informed and set their responses in that context.

Overall we are surprised that the public were not included within the scope of the consultation, as equal partners. The guidance does make reference to some work which sought public views, but it was not clear

what this actually was. Our starting point should be “what do the public want to know”.

The direction of travel is towards patients having access to more and more of their data, which increases the demand and the need for health literacy and autonomy. The traditional top-down approach underpinning this consultation might not best suit this direction.

Our further comments are mostly limited to the questions at the start of this survey, but we feel that the responses apply across the consultation.

The following questions relate to ‘Part one: The basics of explaining AI’:

Q2. What other definitions, if any, should we cover to help inform the guidance?

Q3. How clear are the summaries of the relevant legislation and how they apply?

	Very clear	Clear	Not clear	Not at all clear
GDPR	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Equality Act 2010	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Judicial review	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q4. What other legislation, if any, should we cover?

Does Common Law have any relevance here?

Q5. In your experience, which of the benefits and risks we have outlined are most relevant in explaining (or not explaining) AI-assisted decisions to individuals?

Where do we cover the choice that an individual may be able to make about how their data is used in AI, or knowledge of the ways in which it has been used?

Q6. Is it clear what the different explanation types are, and why they are important for an explanation?

Yes

No

Please provide further detail:

Q7. What other explanation types, if any could we include?

Q8. We have outlined four principles:

Be accountable
Be transparent
Consider context
Reflect on impacts

Are they helpful for your organisation?

Yes

No

Please provide further detail:

Where does co-production (with patients/public) sit? How can you be transparent if you are not involving patients/public in your work; oversight, approvals, design, sign-off, review

The following questions relate to 'Part two: Explaining AI in practice'

Q9. Are there any steps missing in the summary steps?

Yes

No

What are the missing steps?

Q10. Is it clear what information should go into the explanations we have described?

Yes

No

Which explanation is unclear and why?

Q11. What other elements of the data collection and pre-processing, that contribute to explainability, if any, should we include?

Q12. In step two, is it clear how you should choose your priority explanation types?

Yes

No

Please provide further detail:

Q13. Are the examples for choosing suitable explanation types clear?

Yes

No

Please provide further detail:

Q14. After reading the guidance about selecting an appropriately explainable model in step four, how helpful do you feel this will be for your work?

Very clear

Clear

Not clear

Not at all clear

Q15. How clear is the guidance about the tools you can use for extracting rationale explanations and the limitations they have?

Very clear

Clear

Not clear

Not at all clear

Q16. What other rationale explanation extraction tools, if any, could we include?

Q17. Is it clear how you should take the statistical output of the AI system and translate it into meaningful explanation?

Yes

No

Please provide further detail:

Q18. Step five discusses how to train staff, who implement your AI system, to interpret the outputs and apply them to the circumstances of an individual. After reading this, do you feel confident about applying this training in your organisation?

Yes

No

Please provide further detail:

We have highlighted five contextual factors that influence the kind of explanations people want about an AI-assisted decision relating to them. These have come from the research carried out with the public.

Q19. Do these reflect your experiences?

Yes

No

Please provide further detail:

Q20. What other contextual factors, if any, could we include?

Q21. Do the types of explanation we have suggested for each contextual factor make sense to you?

Yes

No

Q22. How likely are you to implement the detailed proactive engagement measures in your organisation?

Very likely

Likely

Unlikely

Very unlikely

Q23. What other measures, if any, should we include?

Q24. Do you have any suggested changes to the healthcare example we have included in Annex 1?

The following questions relate to 'Part 3: What explaining AI means for your organisation':

Q25. How accurate is the characterisation of the following roles and responsibilities?

	Very accurate	Accurate	Not accurate	Very inaccurate
a. Product manager	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. AI development team	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. Implementer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d. DPO and compliance team	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e. Senior management	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q26. What other roles and responsibilities, if any, should we include?

Q27. How feasible is it for you to produce policies and procedures in the areas we have listed?

Very feasible Feasible Not feasible Not at all feasible

Please comment on any specific areas.

Q28. What other policies and procedures, if any, should we include?

Q29. Is it clear what types of information you need to document?

Yes

No

Q30. What other types of information, if any, should we include?

Q31. Do you have any other comments on this guidance?

Q32. What sector do you work in? Please tick all that apply:

- Private
- Public
- Third

Q33. What industry do you work in eg finance, health?

use MY data is a movement of patients, carers and relatives

use MY data supports and promotes the protection of individual choice, freedom and privacy in the sharing of healthcare data to improve patient treatments and outcomes.

use MY data endeavours to highlight the many benefits that appropriate usage of healthcare data can make, to save lives and improve care for all.

use MY data aims to educate and harness the patient voice to understand aspirations and concerns around the use of data in healthcare delivery, in service improvement and in research, aimed at improving patient decision making, treatment and experience.

Our vision is of every patient willingly giving their data to help others, knowing that effective safeguards to maintain the confidentiality and anonymity of their data are applied consistently, transparently and rigorously.

Q34. Where did you hear about this consultation?

- ICO Twitter account
- ICO Facebook account
- ICO LinkedIn account
- ICO newsletter

- ICO blog
- ICO staff member
- Colleague
- Twitter
- Facebook
- LinkedIn
- Other

If other please specify: