The Explanation Goodness Checklist and Scale For Explainable AI

Looking across the scholastic and scientific literatures on explanation in such disciplines as psychology, philosophy, and instructional design, we find assertions about what makes for a good explanation. There is a general consensus on factors such as clarity and precision. Thus, one can look at a given explanation and make an a priori (or decontextualized) judgment as to whether or not it is "good."

Good explanations are said to be ones that are plausible and internally consistent, have an appropriate amount of detail and a clear focus, are veridical or accurate with respect to the thing being explained, are useful for the intended user, are clear and understandable.

The Goodness Checklist is for researchers or domain experts who want to conduct an *independent, a priori* evaluation of the goodness of explanations that are generated by XAI systems. Using the Goodness Checklist, independent judges ask, *Are the researchers right in claiming that their explanations are good?*

The Checklist items are presented first in the Checklist form, and then in the form of Likert scales. For each of the items, the user can be afforded a "free response" opportunity.

The explanation helps me **understand** how the [software, algorithm, tool] works.

YES	
NO	

The explanation of the [software, algorithm, tool] sufficiently **detailed.**

YES	
NO	

The explanation of how the [software, algorithm, tool] works is sufficiently **complete.**

YES	
NO	

The explanation is **actionable**, that is, it helps me know how I can use the [software, algorithm, tool]

YES	
NO	

The explanation lets me know how **reliable** the [software, algorithm] is.

YES	
NO	

The explanation lets me know how **trustworthy** the [software, algorithm, tool] is.

YES	
NO	

1. In the explanation clear and **understandable**?

1	2	3	4	5	6	7
The explanation is not clear to me at all.						The explanation is entirely clear to me.

2. Is the explanation sufficiently **detailed**, or is it not detailed enough?

1	2	3	4	5	6	7
The explanation is not nearly detailed enough.						The explanation is at the right level of detail.

3. How **complete** is the explanation?

1	2	3	4	5	6	7
The explanation seems very incomplete.						The explanation seems complete.

4. Is the explanation **actionable**? That is, does it describe how the [software, algorithm, tool] is used?

1	2	3	4	5	6	7
The explanation does not help guide my actions						The explanation very much guides my actions.

5. Does the explanation describe the **reliability** of the [software, algorithm, tool]?

1	2	3	4	5	6	7
The explanation is of no help in telling me how reliable the [software, algorithm] is.						The explanation makes it clear how and when I can rely on the [software, algorithm].

6. Does the explanation let you know how **trustworthy** the [software, algorithm, tool] is?

1	2	3	4	5	6	7
The explanation is of no help in telling me how or when to trust the [software, algorithm]						The explanation makes it clear how and when I can trust the [software, algorithm].