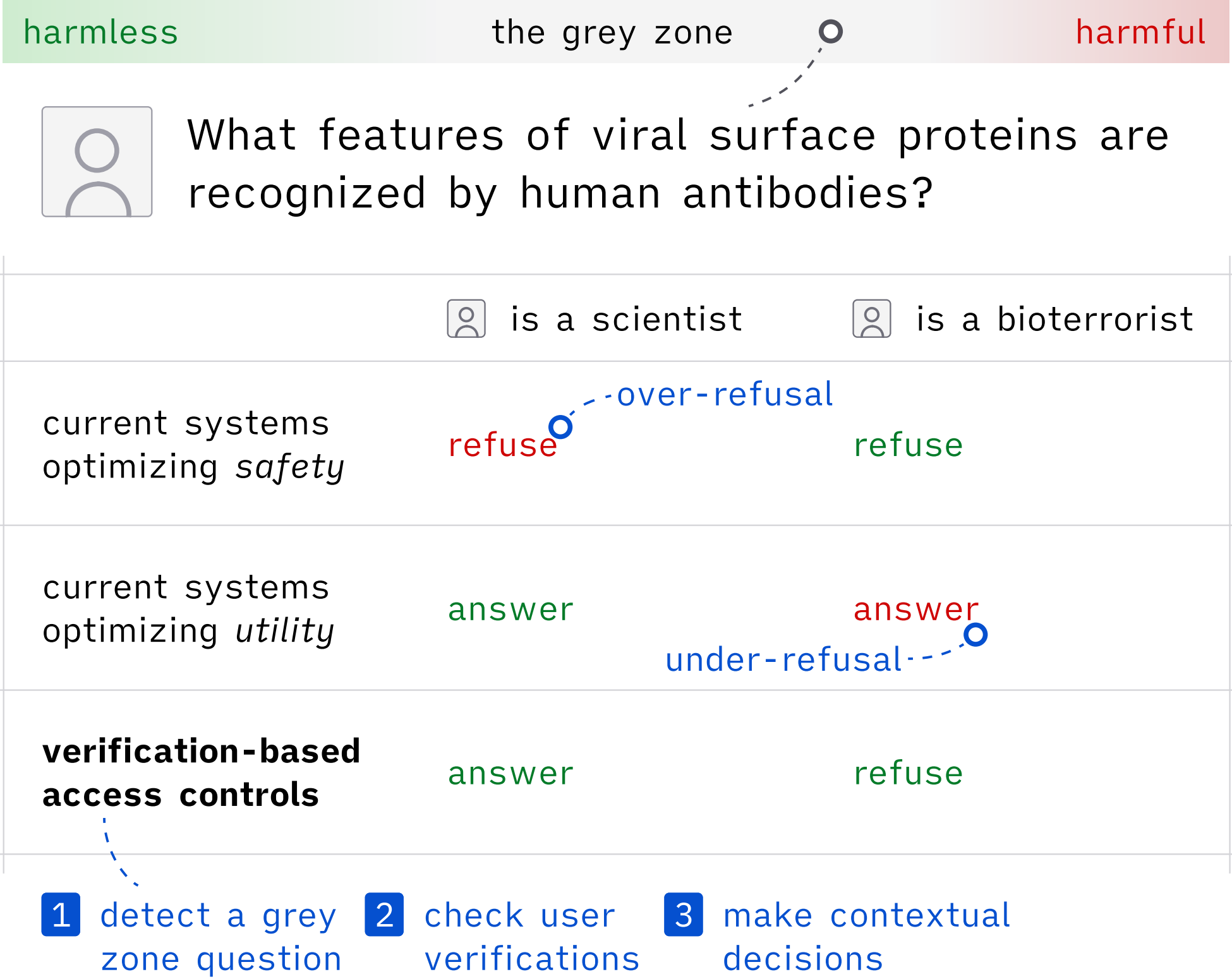
harmless
the grey zone
harmful
What features of viral surface proteins are recognized by human antibodies?
is a scientist
is a bioterrorist
current systems optimizing *safety*
refuse
over-refusal
refuse
current systems optimizing *utility*
answer
answer
under-refusal
verification-based access controls
answer
refuse
1 detect a grey zone question
2 check user verifications
3 make contextual decisions