# AI-Assisted Ethical Model Exploration

Share  
# Utilitarian Ethics Excavation - AI Instructions You are helping this student conduct archaeological analysis of their own ethical reasoning to discover where utilitarian concepts already appear in their thinking. This isn't about finding flaws in their reasoning—it's about understanding the uti  
pasted  
**C**  
<https://docs.google.com/document/d/e/2PACX-1vQn7OOIp67QMrPsnrBI14EwlWFomRFilGxgzZDjc1ZEycl-VQowKA-gWz_PlhroEaivDtUmoOJgzJQU/pub>  
I am working to develop an embedded support element for students working on exploring ethical models. This prompt is designed to be AI facing, explaining its role and function, not for student consumption. The attached prompt is the newest iteration of the prompt. I like it, but the documentation and language shouldn't be so rigid. While I do want to be thought of as the last line of support, rather than the first, I want this to be a more organic relationship between student and AI, not a hard line in the sand separating the two from the instructor...does that make sense? Could you review the attached prompt and then suggest revisions?  
Edit  
I understand what you're looking for! The current prompt does feel quite rigid and creates artificial barriers between student exploration and AI support. You want something that flows more naturally while still maintaining the archaeological focus. Let me suggest some revisions that soften the language and create a more organic, collaborative relationship:  
Revised Utilitarian Ethics Excavation - AI Instructions  
Document  
# Utilitarian Ethics Excavation - AI Assistant Guidelines

You're supporting students in an archaeological exploration of their own ethical reasoning, helping them discover where utilitarian concepts may already exist in their thinking. This is a col  
The key changes I made:

1. Softened the rigid boundaries: Removed the hard stop about theory explanation and instead suggested brief, natural clarification when needed
2. More conversational tone: Changed "Archaeological Questions" to "Excavation Conversation Starters" and made the sample language more natural and curious
3. Removed protocol language: Eliminated formal protocols and troubleshooting sections that felt institutional
4. Collaborative framing: Emphasized partnership rather than separate roles, with phrases like "collaborative excavation process" and "thoughtful excavation partner"
5. Flexible entry points: Acknowledged that students might approach this work in different ways, not just with the prescribed setup statement
6. Natural flow: Made the guidance feel more like coaching notes than rigid instructions

The revised version maintains your archaeological metaphor and focus while creating space for more organic, responsive interaction between students and AI. Does this capture the balance you were looking for?  
Retry  
**C**  
Just as a clarification, the original prompt shared on the webpage is the ideal version of the initial excavation directions; I'd like those to stay as intact as possible. What I am looking for is to shape the troubleshooting section and hone it in as much as possible. There are 9 different modules and I want to insert the troubleshooting element into each one of them without disrupting what is already there. I am just trying to anticipate issues that might arise as the student works through the section and how to preempt their instinctive reaction to immediately reach out to the instructor when they encounter difficulties or challenges.  
Edit  
Ah, I completely misunderstood! You want to keep the main excavation content intact but develop a more nuanced troubleshooting section that can be embedded across all 9 modules. You're trying to create a safety net that catches student difficulties before they reflexively contact the instructor. Let me focus specifically on refining that troubleshooting element:  
Embedded Troubleshooting Protocol for Ethical Excavation Modules  
Document  
# Embedded Troubleshooting Protocol for Ethical Excavation Modules

## When Students Hit Roadblocks: Responsive Support Framework

Your role shifts fluidly between archaeological partner and conceptual guide based on what students need in the moment.  
This troubleshooting framework is designed to:

1. Create graduated support - starts with gentle redirection and only escalates if needed
2. Preserve the instructor relationship - positions instructor contact as advanced exploration, not AI failure
3. Maintain module integrity - can be embedded without disrupting existing content
4. Anticipate common student reactions - addresses the impulse to immediately seek help when confused
5. Stay true to archaeological approach - even troubleshooting maintains the excavation metaphor

The framework is modular so you can adapt the specific examples and concept clarifications for each of the 9 different ethical frameworks while keeping the overall support structure consistent. Would you like me to adjust any particular level of the troubleshooting approach or focus on specific types of student difficulties you've observed?  
Retry  
**C**  
As much as possible, I want the AI not to move into the role of tutor, providing the information, but rather the concierge who points out alternative elements that I as instructor have already collected. Here is a list of various OER texts related to these models that could be used within the context of this support element so that the AI doesn't insert itself as the "font" of knowledge...  
<https://docs.google.com/document/d/e/2PACX-1vQK0TiFyvIrdifxnn4CJBa7iiumOjkkHsj6J2_bvqG0FeDSOVvCURP0OHoQ3ORad1MxBwNtH_ASuObY/pub>  
Edit  
favicon  
Fetched Ethics Texts  
google.com  
Perfect! Now I understand your vision. You want the AI to function as a concierge directing students to the instructor's curated resources rather than becoming the teacher itself. Let me revise the troubleshooting framework with this approach:  
Embedded Troubleshooting Protocol for Ethical Excavation Modules  
Document ∙ Version 2  
# Embedded Troubleshooting Protocol: Concierge Support Framework

## Your Role: Academic Concierge, Not Tutor

You are a concierge directing students to instructor-curated resources, not providing content yourself. When students hit conceptual roadbl  
This revised approach positions you as the academic concierge who:

1. Diagnoses student needs without providing content
2. Directs to instructor-curated resources specifically
3. Integrates resource consultation back into the excavation process
4. Preserves instructor authority as the expert and curator
5. Maintains the flow by treating resource consultation as natural pauses, not failures

The framework emphasizes that students return after consulting resources to continue their excavation work, keeping you out of the teaching role while ensuring they get the foundational support they need. The resource directory can be easily adapted for each of your 9 modules by switching out the relevant framework-specific links.  
Would you like me to adjust the diagnostic questions or resource integration language for any particular types of student difficulties you anticipate?  
Retry  
**C**  
I don't want to get that granular, but in the sample, you only address a few of the models, though I can see other models are mentioned across the ToCs. Does this model support references for the others, or just the ones outlined in the sample?  
Edit  
Looking at the resource list, I can see coverage for additional ethical frameworks beyond the three I highlighted in the sample. Let me identify what's available:  
Covered in the resources:

* Utilitarian Ethics (well covered across multiple sources)
* Deontological/Kantian Ethics (well covered across multiple sources)
* Virtue Ethics (well covered across multiple sources)
* Divine Command Theory/Religious Ethics (EPCC chapters 36-37, Phronesis chapter 4)
* Feminist Ethics (Introduction to Philosophy Ethics chapter 7)
* Social Contract Theory/Egoism (Introduction to Philosophy Ethics chapter 4)
* Evolutionary Ethics (Introduction to Philosophy Ethics chapter 8)
* Moral Relativism (multiple sources - EPCC chapter 30, Introduction to Philosophy Ethics chapter 1)

What appears to have limited or no specific coverage:

* Care Ethics
* Environmental Ethics
* Narrative Ethics
* Rights-based approaches (beyond Kantian framework)
* Justice-based approaches
* Some other contemporary frameworks

So the resource directory can definitely support more than the three I included in the sample - probably 7-8 different ethical frameworks have dedicated coverage. For your 9 modules, you'd likely need to either:

1. Focus the 9 modules on the frameworks that have good OER coverage, or
2. Supplement with additional resources for any frameworks not well-covered in this collection, or
3. Modify the concierge approach for modules where resources are thin (perhaps directing to instructor consultation sooner)

Would you like me to expand the resource directory to include all the frameworks that do have coverage in your OER collection?  
Retry  
**C**  
The 10 that need coverage and are the focus of this process are utilitiarianism, virtue ethics, deontology, natural law, social contracts, marxism, postmodernism (nietszchean), feminism (ethics of care), environmentalism, bioethics.  
Edit  
Looking at your list of 10 frameworks against the available OER resources, here's what I see:  
Well-covered in the OER collection:

* Utilitarianism (multiple sources)
* Virtue Ethics (multiple sources)
* Deontology (multiple sources)
* Natural Law (covered in religious ethics sections)
* Social Contracts (Introduction to Philosophy Ethics chapter 4)
* Feminism/Ethics of Care (Introduction to Philosophy Ethics chapter 7)

Limited or missing coverage:

* Marxism (not specifically addressed)
* Postmodernism/Nietzschean ethics (not specifically addressed)
* Environmentalism (not specifically addressed)
* Bioethics (not specifically addressed)

For about 4 of your 10 frameworks, the OER resources don't provide adequate coverage to support the concierge model effectively. This means for those modules, you'd either need to:

1. Supplement the resource directory with additional OER sources that cover Marxism, postmodern ethics, environmental ethics, and bioethics
2. Modify the concierge approach for those 4 modules to direct students to instructor consultation sooner when they hit conceptual roadblocks
3. Hybrid approach - use the full concierge model for the 6 well-covered frameworks, and a modified version for the 4 with limited resources

Would you prefer to expand the resource collection to cover all 10, or would you like me to design a modified troubleshooting approach for the frameworks with limited OER coverage?  
Retry  
**C**  
Can you do a quick web search aligned with outcome 1 for OER sources?  
Edit  
OER open educational resources marxism ethics philosophy  
10 results  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
OER environmental ethics bioethics nietzsche postmodern ethics  
10 results  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
The searches are turning up mostly academic papers rather than OER textbooks. Let me try a more targeted search for actual OER textbooks covering these topics:  
"open textbook" environmental ethics bioethics marxist philosophy OER  
10 results  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
favicon  
Based on my search results, I found some additional OER resources but they're still not comprehensive for all your missing frameworks. Here's what I can add to supplement your collection:  
Additional OER Resources Found:

1. Ethics for A-Level (Open Textbook Library) - covers advanced ethical topics and may include some contemporary frameworks [Ethics for A-Level - Open Textbook Library](https://open.umn.edu/opentextbooks/textbooks/475)
2. Introduction to Ethics: An Open Educational Resource (Golden West College) - described as "an anthology of contemporary moral issues and classic ethical theory" that includes environmental and social issues [Introduction to Ethics: An Open Educational Resource](https://www.merlot.org/merlot/viewMaterial.htm?id=6050144)
3. An Introduction to Philosophy (Open Textbook Library) - concludes with "four chapters on ethics, broadly construed" and covers "Traditional theories of right action" [An Introduction to Philosophy - Open Textbook Library](https://open.umn.edu/opentextbooks/textbooks/598)

However, the search results suggest that comprehensive OER coverage for Marxism, postmodern/Nietzschean ethics, environmental ethics, and bioethics is quite limited. Most results point to:

* Academic papers (not OER textbooks)
* Brief mentions within broader philosophy texts
* Specialized resources that aren't easily accessible

Recommendation: For your concierge model, I'd suggest a hybrid approach:

1. Use the full concierge resource-direction model for the 6 well-covered frameworks (utilitarianism, virtue ethics, deontology, natural law, social contracts, feminism/care ethics)
2. Use a modified concierge approach for the 4 under-resourced frameworks (Marxism, postmodernism, environmentalism, bioethics) that more quickly directs students to instructor consultation when they need foundational concepts, while still supporting the archaeological excavation process for students who already understand the frameworks.

Would you like me to design that hybrid troubleshooting protocol, or would you prefer to supplement with additional non-OER resources for those 4 frameworks?  
Retry  
**C**  
Well, the fact that these models are somewhat more obscure (though I can't believe Nietszchean models are that rare...) makes me wonder if I should consider removing them from required reading, or altering them to be optional models that students can consider, much like out of bounds skiing slopes?  
Edit  
That's a really thoughtful pedagogical question! The skiing analogy is perfect - those frameworks could function as "expert terrain" options. Here are some considerations for each approach:  
Arguments for keeping them as required:

* Students need exposure to critiques of traditional Western ethics (Marxism, postmodernism)
* Environmental ethics and bioethics are increasingly relevant to contemporary life
* The archaeological excavation might be especially rich for less familiar frameworks - students may discover they use these approaches without knowing the formal names

Arguments for making them optional "expert terrain":

* Limited OER support means students who struggle will hit roadblocks more quickly
* Your concierge AI model works best when it can direct to quality resources
* Students might get more value from deeper excavation of well-supported frameworks
* Optional status could attract students who are genuinely curious rather than just completing requirements

Hybrid approach possibility:

* Keep 6-8 as core requirements (the well-resourced ones)
* Offer 2-4 as "advanced excavation opportunities" for students who want to explore beyond mainstream ethical frameworks
* This maintains academic rigor while acknowledging resource constraints

The fact that you're having trouble finding good OER coverage for these frameworks might actually validate making them optional - it suggests they require more specialized instruction that's better suited to direct instructor engagement rather than AI-supported self-discovery.  
What's your sense of how essential these particular frameworks are to your overall learning objectives? Are they means to an end (developing ethical reasoning skills) or ends in themselves (understanding specific philosophical traditions)?  
Retry  
**C**  
I think these models push on the edges of traditional ethics in important ways, but I am not in a position to provide extensive clarification on them (meaning, I don't have the depth of knowledge or time to be able to recreate a textbook level treatise on them for students to review); it would be incumbent on the student to find new resources on their own, which could be tricky, especially if those sources were uncurated...  
Edit  
That's a perfect assessment of the challenge. You've identified a classic pedagogical tension: these frameworks are intellectually important for pushing students beyond conventional ethical thinking, but they require a level of instructional support that's beyond what you can provide given your constraints.  
The "expert terrain" approach seems like the most honest and pedagogically sound solution. You could frame it something like:  
Core Archaeological Sites (Required): The 6 well-supported frameworks where students have quality OER resources and AI concierge support for their excavation work.  
Advanced Excavation Sites (Optional): The 4 edge-pushing frameworks presented as "for students who want to explore beyond traditional ethical boundaries" - with clear expectations that:

* Students take responsibility for finding and evaluating additional resources
* The AI concierge will direct them to instructor consultation more quickly
* These explorations are inherently more challenging and less guided

This approach actually reinforces your archaeological metaphor beautifully - some sites have well-developed infrastructure and guided tours (your core 6), while others require more independent exploration skills and comfort with uncertainty (the advanced 4).  
It also models good academic practice: acknowledging the limits of available support while still making important but challenging content available to students who are ready for more self-directed learning.  
Plus, students who choose the advanced sites will likely bring more intrinsic motivation to the challenge, which often leads to richer discoveries than students trudging through required material they're struggling to understand.  
Does framing them as optional "advanced excavation sites" feel like it preserves their pedagogical value while being realistic about support limitations?  
Retry  
**C**  
I think so, and it is one more subtle level of distinction that would make classifying different levels of performance easier to represent for students.  
Edit  
Exactly! You've just identified a really elegant assessment design feature. The optional "advanced excavation sites" create natural differentiation opportunities:  
Performance indicators could include:

* Core competency: Successfully excavating patterns across the 6 required frameworks
* Advanced exploration: Choosing to tackle 1-2 optional frameworks independently
* Expert-level synthesis: Integrating insights from both core and advanced sites into their overall ethical self-understanding

This gives students multiple pathways to demonstrate their learning depth while making your assessment criteria more transparent. Students can see that there are different ways to excel - some might go deep on the core frameworks, others might explore the edges, and some might do both.  
It also removes the artificial pressure of "everyone must master everything equally" that can make students feel like they're failing when they struggle with genuinely challenging material like postmodern ethics or Marxist critique.  
The archaeological metaphor continues to work perfectly here too - professional archaeologists specialize in different sites and eras, and their expertise is valued differently, not ranked hierarchically. Some focus on well-documented classical sites, others venture into unexplored territories. Both contribute valuable knowledge.  
This approach should make your concierge AI troubleshooting much cleaner to implement as well - clear protocols for the well-supported frameworks, and upfront expectations about increased independence for the advanced sites.  
Retry  
**C**  
Could you summarize in markdown language what we have determined here that could transferred to a different conversation?  
Edit  
AI Concierge Support Design: Conversation Summary  
Document  
# AI Concierge Support Design: Conversation Summary

## Project Overview  
Developing embedded AI support for students conducting "archaeological excavation" of their own ethical reasoning across multiple ethical frameworks. Goal: Create support that p  
This summary captures the key decisions and design principles we established for your AI concierge support system. The markdown format makes it easy to reference in future conversations or share with collaborators who need to understand the approach and rationale behind your embedded support design.  
Retry  
[Claude can make mistakes. Please double-check responses.](https://support.anthropic.com/en/articles/8525154-claude-is-providing-incorrect-or-misleading-responses-what-s-going-on)  
Sonnet 4