



This data was created in the context of the project COMPLEX THINKING & AI
Category: Logs

This document contains the major log of work process of the CT & AI project team.

Recommended citation: Melo, A. T., Renault, L., Caves, L., Garnett, P., Lopes, P. D., Ribeiro, R., Santos, F.(2024).
Development documents and tests for the development and calibration of an evaluation framework for
Complex Thinking on the dimension of Structural Complexity and the property of Structural Variety and
Dimensionality. Coimbra: Centre for Social Studies.
This document was produced as part of the project COMPLEX THINKING & AI, Funded by Stefan Pernar



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

Date	Task	Comments	Relevant Chat logs	Learned lessons or hypotheses	Models used
13th of June 2024 - Morning	Uploading the 'property' document.	<p>The document does seem to be read however there are some things to keep in mind.</p> <ol style="list-style-type: none">1. It is not clear how some of the instructions are interpreted. E.g. boolean operators in the text, we aren't sure how they are interpreted and if they are consistently interpreted in the same way.2. It looks like it takes some sort of overview of the content of the uploaded instructions, it doesn't seem to interpret them as specific things that should be done systematically (probably see below also).3. Dragging in the spreadsheet doesn't work.4. Phil (see other link) tried a more stepwise approach, getting GPT to use the spreadsheet in stages, e.g. work on the first row, and evaluate the inductors on that row.	chat about a Labour Policy	Perhaps a more stepwise approach is useful. E.g. go through stages in turn.	
13th of June 2024 - Morning	Analyse a scientific paper Neto & Ribeiro based on the provided property document	<ol style="list-style-type: none">1. It did not pick up the specific language of the property document2. We did not define our role nor our problem. Need to define the target /problem3. "Too much memory" too path dependent, keeps responding in the style of previous requests4. We need to "train" in a style5. The dominant mode is a informal conversation style, human like		<p>i. We need to create a "personality" and stye role-in "knowledge file"</p> <p>ii. We need to set up the</p>	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<ol style="list-style-type: none">6. It's selective but it's not clear the criteria7. Did sort of thematic analysis but not exhaustive8. "It thinks he knows things". if not restricted he builds on unknown data sources9.		<p>stage/bound the context "modus operandi"</p> <p>Integrate "action" in our GPT-command words</p>	
1	We asked for deepening questions	<ol style="list-style-type: none">1. He learns a style and keep it2. He ignored the commands on relation. he was sticking with identification of themes but not the relations. missed the coupling specified in the indicators3. When prompted with the questions on relations he provides, builds a network4. It was still interpreting even when staying within the text?5. Very salient and structured information that was ignored		<p>We need to tell it what information is significant.</p> <p>Break down the tasks that imply analysing the coupling.</p> <p>Write the specific instruction on how to carry the task (e.g. explore questions)</p>	
	Generating				



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

	images				
13th of June 2024 - Afternoon	Exploring getting different ranks	<p>This task was focused on seeing if it would ever rank something low complexity or if it was stuck in a mode of ranking things as always high. The task involved asking GPT to make a version of the Family Case Study that it thought would rank as low on all criteria in the structural complexity property. This was then loaded into a new GPT chat which was primed with the spreadsheet. It then evaluated the overall complexity as follows.</p> <p>Moderate complexity level for structural variety and dimensionality and framing and contextualisation.</p> <p>Low complexity level for depth information and situatedness and perspectivity</p>	<p>This is a link to the chat:</p> <p>chat.log.evaluation_13.06.2024</p>	<p>It is possible for it to give low and moderate ranks.</p>	
13th of June 2024 - Afternoon	Analysing a case with properties of structural variety				
25th of June 2024 (developer.inter.user#1 and developer.inter.user#2)	Assessing levels of complexity in scenarios. Trying to generate scenarios	<p>ChatGPT tends to see "target" as the problem and not the one to be solved.</p> <p>It seems necessary to get chatGPT to define the system of interest</p>	<p>Chatlog_evaluation_mode_25.06.2024</p>	<p>We need to ask a definition of the SOI</p>	
		<p>The answers tend to be more global.</p>			
26th of June 2024 (developer.in	Generate narratives of complex	<p>We started with scenarios from previous session and attempted to improve it</p> <p>When we ask chatGPT to pay attention to the details</p>	<p>https://chatgpt.com/share/31c4a5f9-e5ca-4</p>	<p>Make a step-by-step instructions</p>	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

ter.user#1 and developer.int er.user#2))	<p>scenarios and test commands;</p> <p>We have highlighted the differences from one version to the other and the prompts that led to the differences (1.add depth; 2. add situatedness)</p> <p>We needed to define situatedness (ask explicitly for examples)</p>	of the property description and its dimensions, the generated response changes qualitatively and becomes more complex, even without a specific instruction to do so	8b8-b28b-ceb 54351c2bd https://docs.g oogle.com/do cument/d/1SA 5LK4C8gvkpif DKNI98nCgO WmkSZ0s9Dx qWk8nxFFo/e dit?usp=shari ng		
26th of June 2024 (developer.in ter.user#1 and developer.int er.user#2; inter.user#2 later on)				We made improvements on the description of the property of structural variety and dimensionality in V3 of the CT coding_ structural variety file	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

				trying to add detail to the criteria. We used these detailed instructions to generate narratives for building scenarios	
26 June (developer. in ter.user#1)	I asked GPT to evaluate the complexity of the narrative created with in the previous sessions. Version 2.7 of the Rivertown fictional scenario was used	<ul style="list-style-type: none">• I fed the spreadsheet and clarified the goal of the task and what the target system of interest was.• Th2 evaluation was mostly accurate but it missed what the “four facets” of the content that needed to be considered for evaluating multidimensionality and variety of contents. I then prompted it to look specifically at where the description was in cell E2 and the evaluation was then more accurate. It was then able to redo the evaluation• I asked it to generate a low and moderate complexity version and it looks quite accurate. The criteria seems to have been reasonably understood	His is link to chat Chatlog_evaluation_26.06.2024_t2	It could not easily connect the information from the different cells in a row. I have corrected the spreadsheet CT_coding structural variety in V4 sheet, so that the evaluation criteria now explicitly refer to the detailed description in Cell E2. Need to try again with this new	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

				<p>version</p> <p>In general there was a good understanding of the criteria but it now mostly focused on the first expression of the property. Need to generate again all over</p> <p>Need to see exactly where description of properties needs to be refined</p>	
2nd of July (developer.int er.user#1 and developer.int er.user#2)	Attempt to make chatGPT see the information in detail in version 4 we have created new versions of the spreadsheet (version 7)	Providing the updated spreadsheet wasn't enough. It continues to confuse the notion of "facets"; it makes the most general evaluation possible (v5 of the spreadsheet was worse than v4)	Chatlog_evaluation mode_02.07.2024 (cf also tests 1, 2 3, 4)	We need to repeat all the information in detail it will probably be necessary to change the terminology	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

				(e.g. facets). We used “dimensions” again Cf. (article examining the hermeneutic parameters of prompts in chatGPT)	
2nd of July (developer.in ter.user#1 and developer.int er.user#2)	Still creating versions for the evaluation of the scenario (using version 7 of spreadsheet)	We got the correct evaluation with the Test 6, but we had to force it to pay attention to cell E2	Chat log_test6_eva luation mode_02.07.2 024	It can understand the instruction once it has been directed. There is a technical challenge in guiding it through the steps.	
5th July (developer.in ter.user#1 and developer.int er.user#2)	Still trying to improve the evaluation of structural complexity with versions 8 and 9	Another test showed that GPT failed again the evaluation of the first sub-property pertaining to multidimensionality and variety of contents when given the spreadsheet as a whole. We first introduced an introduction om the spreadsheet which it followed but the evaluation		We will need to give instructions step by step, in the reverse process that	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>failed (test 8) We then decided to split the property and have different rows with different parts in V9 and change names of expressions of properties to sub-properties</p> <p>The test fails. It goes back to doing general evaluations and “averaging” the information</p> <p>Test 10 We ask to do stepwise evaluation row by row updating the instruction and referring to sub-properties</p> <p>Test 10 and 11 was a failure worse than others</p> <p>In test 13, we tried the version 4b. It continues to give a general assessment ignoring the specific instructions. When we give the content of the cell (E2), the version was significantly improved. We repeated the content of the cell in the prompt.</p>		<p>we used to build the narrative. E will need to give the instruction row by row.</p> <p>Next time, try to give a file with only one row at the time.</p> <p>Ask it to enhance the complexity of the narrative</p>	
16th July (developer.inter.user#1 and developer.inter.user#2)	Evaluation test (with hyperlink) -v14	Apparently chatGPT is unable to read hyperlinks		Create a decision tree instead	
	We created a document with steps and we gave each step at the time as a prompt (test 15)	<p>The first test with the step 1 went well (cf document CT_CODING_STRUCTURAL VARIETY_NARRATIVE FORM)</p> <p>The other steps also work!</p>	Chat log_test 15_evaluation mode_16.07.2 024		



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

	Test with all the previous steps in a single document (test 16)	<p>The document with the all the steps is CT_CODING_STRUCTURAL VARIETY_INSTRUCTIONS IN NARRATIVE FORM</p> <p>In general, it seems to have worked well</p>	Chat log_test 16_evaluation mode_16.07.2 024		
	Test of evaluation with single document and different scenario (test 17)	<p>In test 17, we provided the low-complexity scenario. Chat GPT rated it 'moderate'; Then, in the same chat, we provided the moderate scenario, and asked for an assessment of it. The moderate scenario was rated 'high'. We then provided the high complexity scenario. It correctly compared the three. It realises the <u>relative</u> differences in complexity. It also adequately provided some questions and activities to increase the complexity</p>	Chat log_test 17_evaluation mode_16.07.2 024	It seems necessary to provide example narratives, as comparison criteria (a high, a high, a low).	
18th July (developer.inter.user#1 and developer.inter.user#2, tech.inter#2, inter.user#1, developer.inter.user#1)	Continuation of Test 17 (try to build a structure - dimensions, sub-dimensions, properties and sub-properties; dimensions of the thinking and dimensions of the Target SOI. Trying to disambiguate language).			<p>We need to have a visualisation (map, diagram,...) of the structure;</p> <p>Work with the guiding questions</p> <p>We need to refine the language</p>	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

	Test 18				
25th July (developer.inter.user#1 and developer.inter.user#2)	Working with the guiding questions Version 15 of the spreadsheet	<p>Organising the scaffolding mode:</p> <p>We created primary questions and deepening questions. We will also indicate other modules that are relevant to the questions (with the coding). If the complexity of the answers is low, we move onto the deepening questions. If it remains low, we change modules.</p> <p>Possibility to also refer to activities (to not be limited to the question-answer mode)- action/embodied mode</p>	CT Cod...		
19 Julho (developer.inter.user#1 and developer.inter.user#2)	Working with guiding questions	<p>Continuing to reflect on the necessities and modes of organisation of the scaffolding mode and the relation between properties and sub-properties as well as dimensions of the content of the thinking in relation to TSol</p> <ul style="list-style-type: none">• The scaffolding mode in principle should coupled with the evaluation mode• We can have different modes of the tools for each mode which will complement and alternate• Mode 1: Dialogical scaffolding- scaffolding based on a sequence and guiding questions<ul style="list-style-type: none">◦ The guiding questions have different levels: general and specific. The responses to the first level should be			



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>evaluated and depending on the level of complexity (EVAL MODE) more specific questions can be made</p> <ul style="list-style-type: none">○ The questions target the primary property of the module but they may also:<ul style="list-style-type: none">■ Have embedded secondary properties that enrich the primary property■ Be complemented with other properties from modulating properties (separate modules)● Mode 2: active/Embodied scaffolding focused on the self and target system of interest● Mode 3. Synthesising and mapping● Mode 4 metaphorising <p>(see miro)</p>			
30. 07.2024 (developer.inter.user#1 and developer.inter.user#2, developer.inter.user#3tech.inter.user#1, other team member)					
1.08.2024		We tested running the scaffolding instructions on		It is	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

developer.int er.user#1 and developer.int er.user#2,		<p>Chat GPT and Perplexity the pro mode and while GPT extrapolated on a case based on the first information provided.</p> <p>We think of different ways in which scaffolding works (inductive, abductive, deductive)</p> <p>Perplexity sticks to the instructions and suggests activities that increase the complexity of the thinking (while chatGPT provides suggestions for intervention in the field).</p> <p>We modified the document V3_CT_CODING_SC.SVD_SCAFF_INSTRUCTIONS</p>		necessary to adjust the instruction sheet according to the platform to be used (chatGPT demands a more restrictive and rigid instruction; Perplexity takes it literally, so the instruction must leave some room for flexibility)	
02.8.24 developer.int er.user#1 and developer.int er.user#2,		<p>Continuing work with V3_CT_CODING_SC.SVD_SCAFF_INSTRUCTIONS</p> <p>This instruction version 3 works better with chatGPT than with Perplexity (it seems too rigid for Perplexity). After answering step 1 of the question, Perplexity went well.</p>	Chat log_evaluation mode_test_P erplexity_02.0 8.2024		
12.8.24 (developer.in	Working with guiding questions	We've updated the document CT_CODING_SC.SVD_SCAFF_INSTRUCTIONS_V			



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

ter.user#1 and developer.int er.user#2, t ech.inter.use r#1, tech.inter.us er#2)	and their coding Test with Claude	3 The test with Claude performed well in the evaluation and was promising for relationality Claude offers the advantage of working with “Artifacts”			
14.8.24 developer.int er.user#1 and developer.int er.user#2)	Test (narrative/dialogic al test) with Claude	<p>Claude performed excellently: it asked question by question, following the instructions and keeping the script throughout the dialogue It selected the questions to be asked according to the user's answers</p> <p>The metaphor was initially suggested and kept as an image, but not often returned to.</p> <p>(In the same test, chatGPT failed to stick to the script and quickly suggested interventions.)</p> <p>We think that Claude is the most well adapted platform so far.</p> <p>Check the learned lessons</p>	Chat log_scaffoldi ng development test0_claude _14.08.2024	<p>In Claude: The evaluation was very good, but the “variety of content” does not yet seem to be sufficiently considered.</p> <p>Make the metaphor more present</p> <p>Add to the instructions that it is necessary to identify what [OBS] don't know and include these</p>	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

				<p>gaps in the summary/map .</p> <p>Also indicate in the instructions to keep the dimensions and sub-dimensions present for the [OBS]. Periodically offer the [OBS] ways of retaining this information</p> <p>Include prompts offering improvements for complex thinking from time to time</p>	
16th-19th Sept 2024 (developer.inter.user#1)		Developing and further refining (for modes and artifacts) the protocols SVD REL			



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

19th Sept developer.int er.user#2 on Claude developer.int er.user#1 GPT-4o tech.inter.us er#2 GPT-o1-prev iew tech.inter.us er#1 taking notes	Various reflections on using GPT-4o (and Claude in parallel)	Test 1. 19.09.2024 on Claude GPT for the first time goes back and asks the observer to respond to questions unanswered. Still presents them in bulk GPT does keep the structure of information so well GPT changes questions slightly does not keep tags and puts them in bullet GPT getting stuck in sub-property questions, its not offering evaluation and not doing scaffolding.	Chat log_scaffoldin g _development test 1_claude_19. 09.2024 (See chatlog_scaff olding development test (21 to 25)_ GPT_unccert ain date: note: for some reasons the dates of these chats were not recorded. It is possible that these were some of these were the ones used in paralell with Claude)		
(ibidem)	Claude (with GPT in parallel)	Claude: friendly response Claude follows instructions well. Narrative synthesis produced a nice mermaid	Chat log_scaffoldin g		



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>diagram of the structure of the case - so far.</p> <p>Soon encountered “Long chats cause you to reach your usage limits faster” message.</p> <p>Experimenting with saving state via downloading (to Google Drive) artifacts: Metaphor, Narrative, Synthesis.</p> <p>Claude taking ownership of the case (referring to “our” case).</p> <p>Moving non-linearly between questions.</p> <p>In phase 3.2 (1st dialogical step) Offering to create a metaphor when its not explicitly in the protocol.</p> <p>Follow-up questions are being qualified to fit previous responses. E.g. in your answer, take into consideration... After posing a question about goals, the questions now start to align with those goals.</p> <p>Glitch: In step 3.2, it did not enter Active Scaffolding Mode, instead it moved to Phase 4 Modulating Questions.</p> <p>It made some interpretations, and got a bit of track. It then offered to close the sessions. However when the response was no it then carried on, picking up the thread and carrying on at around 4.3. It then updated the narrative (the narratives are pretty good).</p> <p>It got stuck around 4.3 asking more questions. Rather than changing mode. We are trying to answer that we don't know to see if it will just move on. So after a couple times it did move onto the SYTH mode, which</p>	<p>_development test 1_claude_19. 09.2024</p>		
--	--	--	---	--	--



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

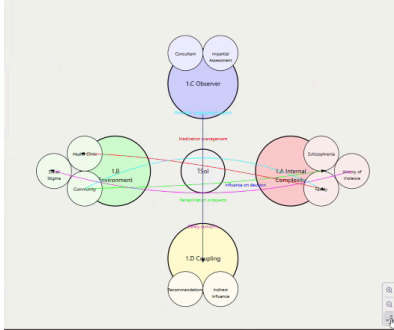
		<p>it was supposed to do in the end.</p> <p>The MAP isn't very good, it's not using the updated MAP code. Trying to get it to call the correct instruction [TSol_MAP] - it seemed to understand this instruction better.</p> <p>The last part of the instructions seem to break the flow a little bit, and we are having some problems getting to close out the session and do the last steps. Some of this might be that it has numbering issues, so perhaps it is ok.</p>			
19-20 september 2024 on intensive team camp)	developer.inter.user#1 developer.inter.user#2	<p>Continuinnng test on Claude. Refining Protocols particularly for Relationality</p> <p>Protocols seem good enough to test with other users</p>	Chat log_scaffolding test_developm ent_test2_clau de_19.09.2024 and 20.09.2024		
16th-19th (on intensive team camp) Sept 2024 tech.inter.user#2	Reflections on general tasks.	<p>First few days: Mapping content to dimensions: Number of issues around the generation of diagrams, the main one being that constraining the output of the diagram generation is challenging to get a consistent output.</p> <p>We also have been building tables with the mapping data inside, we noticed that it never mapped content to multiple subdimensions.</p> <p>Using projects and modularity: We have been testing the use of the project specific</p>			



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

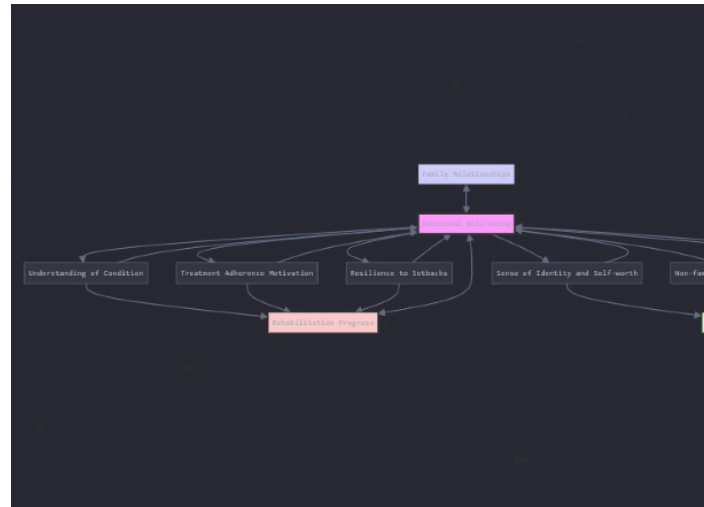
		<p>instructions. This allows instructions to be added to all chats in the project that are then accessible from the chat without uploading any additional documentation. How these are best structured and then executed as instructions has not been explored very well. If these work it might be possible to have a project that does each part of the system with specific instructions and attachments. Reducing the amount of activity on a single chat. Also projects/modules could be interacted with in any order.</p> <p>Trying GPT o1-prerelease: This doesn't have the ability to read attachments at the moment, it doesn't therefore have the ability to follow the instructions properly. So we will leave it for now.</p>			
23.09.2024 developer.int er.user#1 and developer.int er.user#2, inter.user#1	Test with inter.user#1 case of court and criminal with mental disorder	<p>Claude adapts well the scaffolding questions to the context with relevant examples. The evaluation works very well and allows the adaptation (choice of the questions, but also elaboration of examples)</p> <p>Checking the NOTE-TAKING MODE (Filipe had some insights that he just spoke out loud)</p> <p>Definition foTSol ?</p> <p>Check the MAPPING MODE</p>	PILOT TESTS_inter. user#1_chat logs		
02.10.2024	Picking up from the end of the	Relational map: the first time it created a mermaid-type network map; it had to be re-prompted			



developer.inter.user#1 and developer.inter.user#2, inter.user#1, tech.inter.user#1	previous test conversation (Test 3. 23.09.2024) with inter.user#1 now moving onto relationality.	<p>to generate the expected radial map: when it did so, the map was quite good. On the first pass it did not include relations, and had to be prompted to do so. The relations (edges) are named.</p>  <p>Issue: developer.inter.user#1 asked to start with Evaluation, but the scripts were built assuming that it would do Scaffolding first.</p> <p>Nonetheless, the radial map was expected.</p> <p>When asked to continue, Claude offered up Scaffolding mode.</p> <p>Interestingly, when Claude offered up potential relationships to consider, (such as emotional well-being) it generated focused networks, with directional and labelled feedbacks. e.g. "Emotional well-being centered relational map"</p>			
--	--	--	--	--	--



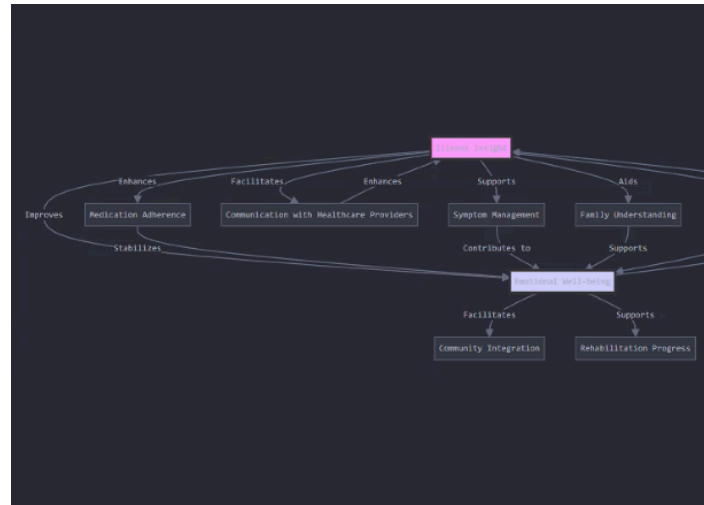
This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log



Picks up on previous responses, e.g. inter.user#1's offering that illness insight being important. It then redraws the map, relationally integrating this element. But it also offers questions that challenge the role and importance of illness insight.
e.g. Illness insight and its relations in the TSol



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log



Wheninter.user#1 prompted that he was ready to finish, Claude offered up a Metaphor (of a garden), while asking if this was appropriate and whether it could be improved or modified. It also offered to update the narrative.
The request for updated narrative generated a structure document tha tinter.user#1 thought accurate.

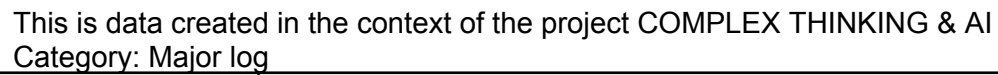
A final request for an evaluation resulted in an overall evaluation of High Complexity, particularly noting that it excels in the treatment of relationality. However, the eval of relationality was rather global. So, it was asked to provide more focus on the relationality.

When it did, it used the evaluation framework



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

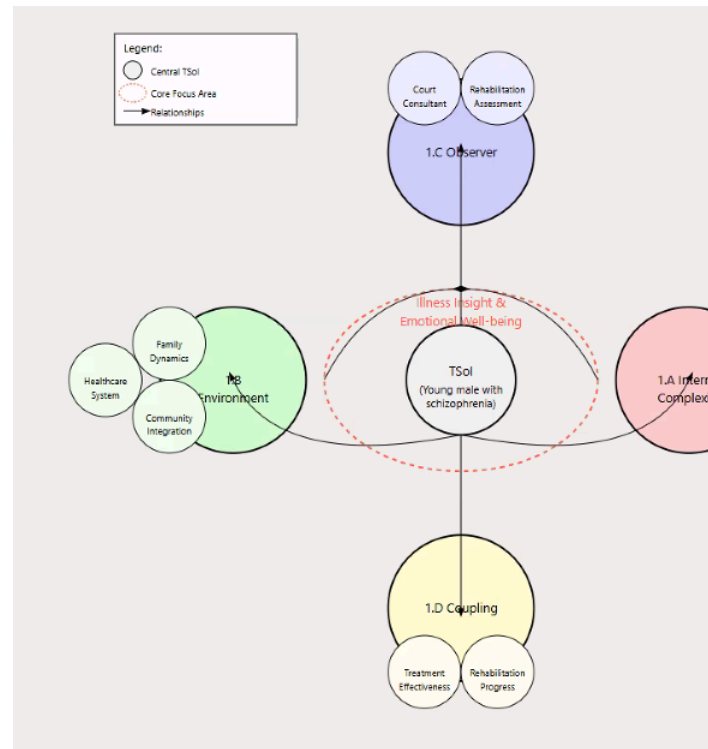
		<p>appropriately. It even offered up some further activities.</p> <p>REFINE: when asked for Evaluation, the evaluation needs to include structural and relational evaluation.</p> <p>inter.user#1s reflections were that the process had made him realise the importance and salience of “insight” - and that this was something of a shift in his thinking.</p> <p>When asked to provide an easy-to-memorize version, it offered up a mnemonic of TIES</p>			
--	--	---	--	--	--



		<div><h3>TIES Relationality Summary</h3><p>T - Threads: Multiple interconnected relationships</p><ul style="list-style-type: none">• Illness insight and emotional well-being form the central thread• Many threads connect various aspects (family, treatment, community)<p>I - Integration: Holistic view of the TSoI's situation</p><ul style="list-style-type: none">• Different elements are woven together cohesively• Changes in one area ripple through the entire system<p>E - Exploration: Dynamic investigation of relationships</p><ul style="list-style-type: none">• Circular thinking examines how factors influence each other• Considers both immediate and long-term effects of relationships<p>S - Sophistication: High complexity in relational thinking</p><ul style="list-style-type: none">• Qualifies relationships (e.g., "deeply intertwined")• Identifies relationships between relationships<p>Remember: The narrative TIES together multiple threads, integrates various elements, explores dynamic relationships, and shows sophistication in relational thinking.</p></div>		
		Final request for a map: ("Updated map of my thinking")		



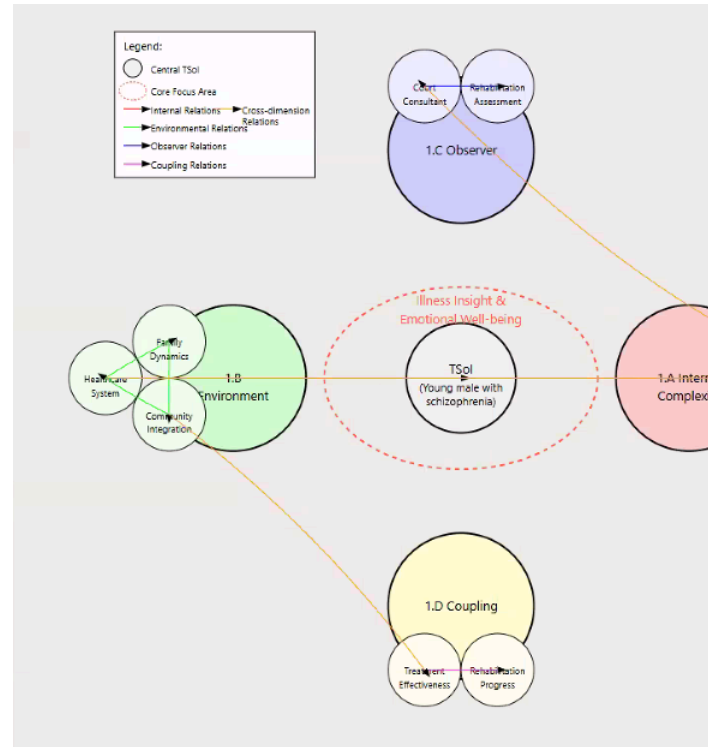
This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log



When prompted for subdimensions:



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log



Feedback from the user:
“These sessions with Claude have been quite enjoyable. I especially appreciate how the prompts are designed to elicit non-linear thinking. I am not sure if I would enjoy it as much if I was doing it without the team on Zoom. Nevertheless, I think that a less verbose Claude would make it even better. When the prompts are



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>finetuned and tested, maybe we can ask Claude to use a more “natural” dialogue.”</p> <p>“</p>			
10.10.24 developer.int er.user#1 and developer.int er.user#2, inter.user#2	Test with inter.user#2 (Test 4. Housing)	<p>Create a prompt at the end of a protocol, so that the tool itself asks for the next protocol</p> <p>The clarification of key terms was not correct (protocol V6); it insists on answering PQ3</p> <p>This was a test in which an initial document was uploaded. The logic of the interaction has changed (less dialogic in answering the questions). In this case, the tool didn't ask new questions unless provoked. The tool adapted well to this kind of interaction and chose the questions that were pertinent to proceed. At one point, it moved away from scaffolding (Raquel asked for answers and to identify what was missing in the document). It didn't present artifacts, metaphors, etc. After a while, it came back to the protocol. [It could quickly turn into a discussion about the content.].</p> <p>Raquel mentions that she doesn't want to write down the answers to basic questions that she has already answered in other documents. (relevant comment to specialists)</p> <p>Check if the description of the natural environment is in the document uploaded</p>	PILOT TESTS_inter. user#2_test 5_chat logs	<p>Refine the protocol to improve the response of the tool when the user moves away from the scaffolding mode.</p> <p>Introduce a meta guideline/prompt to regulate freer interactions</p>	
	Test 5. Housing	(Chat erased. It initiated the dialogue without	PILOT	Try to omit	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

	(modified protocol. V7)	<p>presenting the key words and using the terms as TSol directly, before Phase 0). We changed the protocol accordingly.</p> <p>We reached the limit in that account</p>	TESTS_inter. user#2_test 5_chat logs	the names and codes of the phases for the user	
	<p>Teste 6. Housing (modified protocol V7. We are using another account)</p> <p>Teste 6 in the second account</p>	<p>The protocol worked well at the beginning, with the context and the introduction of the user.</p> <p>In the sequence, it lost track of the protocol. When we asked to go back to the scaffolding, it went to phase 0 again.</p> <p>The modes were working well in the sequence.</p> <p>The user introduced a large approx. 60ps document with the information about her TOSi. It was a book on the issue of Housing.</p> <p>We quickly reached the limit of the account.</p>	PILOT TESTS_inter. user#2_test 6_chat logs	We need to compare the V7 and the V5 (or V4), which was working better	
	Test 6		PILOT TESTS_inter. user#2_test 6_chat logs	Need to create a protocol for the user to know how to manage the interaction and they can ask and should avoid	
10.10.24	Test with	In the test, the tool stuck to a particular relational pair	PILOT	Add a	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

developer.int er.user#1 and developer.int er.user#2, inter.user#1	inter.user#1 (continuation - relationality)	<p>(whichinter.user#1was most interested in). It switched little between other pairs, even proposing other questions. It didn't move from one dimension to another.</p> <p>In this protocol, it offered the metaphor instead of asking for it.</p> <p>Relational map worked well, even if only showing the relations between dimensions (not between sub-dimensions)</p> <p>Legend: Legal Domain Medical Domain Institutional Context</p>	TESTS_inter. user#1_chat logs	[PHASE 0] with procedural questions to the document CT_CO...	
14.10.24	Test 7 (Housing)	Protocol V8	PILOT	Correction of	

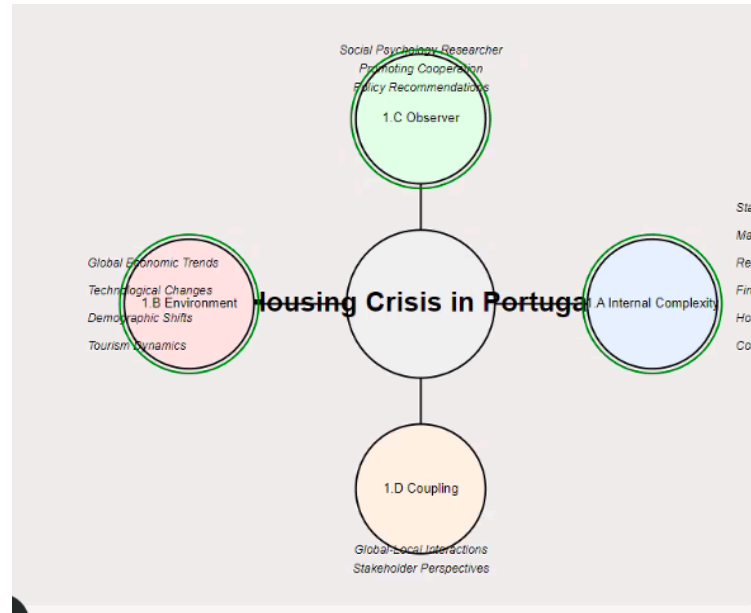


This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

developer.int er.user#1 and developer.int er.user#2, inter.user#2	With inter.user#2	<p>It gave an artifact with the foundational questions, all at once</p> <p>We initiated another test, trying to correct that</p> <p>We made some adjustments in V9 (e.g, we erased the instruction to go to step 2.3 if a document is uploaded. We considered that the foundational questions are important even if a document is uploaded)</p> <p>Protocol is working with the recent modifications. Step 3.2, for example ([MAP MODE] was improved and the areas that need more development were highlighted in red.</p> <p>It chose the questions very well, according to the evaluation</p> <p>Phase 3 all ok and proceeding as expected with quality questions and good adjustment to the evaluation of complexity</p> <p>Note: we have not engaged with the activities of the active scaffolding mode so we do not know how well it returns to the protocol</p> <p>In phase 4 it is choosing a focus on less developed properties</p> <p>We went out of messages before finishing protocol !!! W asked for updated TSol MAP, Synt and Narrative</p>	TESTS_inter. user#2_test 7_ chat logs	<p>the version - production of V9 Production of C 9.1 (minor adjustment but not tested)</p> <p>Build simplified schematic map for the user with different steps of the protocol and where they are so they can “jump”.</p>	
--	----------------------	---	---	--	--



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log



No update TSol map relations are still not marked in detail


Added instruction on the MAP Mode on Version 9.1 of the instructions

It stays on dialogical scaffolding for too long and we did not come to see self-scaffolding mode. The user should have the possibility of seeing “where” they are in the protocol in jump if needed or wished.

There were a couple of requests “off protocol” where the user asked for suggestions of activities to increase the



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		quality of their thinking			
16.10.2024		Created shortened and more simplified version 10 of SCAFF SVD for tests eliminating table and decreasing artefacts for tests. ALTERED STEP 2.3. Created two options of synth mode Affected also step 4.7 DECREASED TO 3 QUESTIONS ON STEP 3.3 Reversed step 4.7 and 4.8	 File		

STAKEHOLDERS WORKSHOP I

Date	Task	Comments	Link to Relevant Chats	Learned lessons or hypotheses
17.10.2024	Notes on developer.inter.us er#1 Intro	Need more schematic figures to introduce CT in first couple of slides Knowing through coupling: typo: matrixes -> matrices Complex thinking as mode process of coupling: title of slide needs altering: either mode or process (or add /) [THOUGHT] should some of the text of presentation be used as background "knowledge" for the AI tool? Care with use of terms such as Leverage Points: if used, it should be qualified. (However, I guess Meadows provides this kind of qualification for her metaphor) Slide 19: point iii formatting - unnecessary paragraph break Slide 34: Might be good to show thumbnails of the artefacts - particularly for the radial diagram; metaphor could have an image etc. [NOTE] better not to make generalisations about relative performance of the different AI tools, because there are nuances, and things change	PILOT TEST_exter.us er#1_17.10.20 24_chat transcript and reflection	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>over time. There are other org. Culture and privacy issues that also factor into our use of Claude.</p> <p>[NOTE] the presentation is very dense. What could be left out or de-emphasised for a more digestible introduction? E.g. in this case, the test really only covers 2 of the properties (yes, with modulations etc).</p> <p>[NOTE] better to send account information etc beforehand, to streamline the transition to the test</p>		
<p>17.10.2024 developer.int er.user#2 ; tech.inter.us er#2</p>	<p>Test with exeten-user#2</p>	<p>It started by asking more than one question at once, it asked questions 1 to 4 at the same time. We started just by asking it to proceed with the complex thinking process and skip the introductory explanation. It then went straight to asking 4 questions at the same time. We are proceeding at this time to see if it will sort itself out, or if we might need to start again...</p> <p>The intention of question 4 was not clear, it might be the use of the term 'values'. We said to skip it for now to see what would happen with the tool, and also with it asking a bunch of questions at the same time we aren't sure what it is going to do.</p> <p>It did the same again at the next part, asking FQ1-4 at the same time. We are going to answer one, and ask it to ask them in turn.</p> <p>We specifically told it to ask one question at a time now. So we might be back on track. We are at SCAFFQ1.. Making a response It has moved on to SCAFFQ2, and asked only that question. What this question means is less clear to the user. Might be a language thing again. Brief answer given, but user indicates that they are having to think... so that is good.</p> <p>Note: it is using the name of the target system of interest provided by the user, rather than TSol.</p>	<p>Used version 9 SVD and V3 general instructions</p> <p>PILOT TEST_exter.u ser#2_17.10.2 024_chat transcript and reflection</p>	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>We have moved to SCAFFQ3. Terms used in previous answers are being used a little in the questions. See note about TSol above. User thinks this question is better than the other one... 'great point to think about'.</p> <p>We have now gone from SCAFFQ3 to SCAFFQ5, so it seems to have decided that SCAFFQ4 was answered. We are proceeding with SCAFFQ5. Encouraging a brief answer...</p> <p>We are now moving to SCAFFQ6. It has made some suggestions about the environment. Couple of sentences provided for the answer.</p> <p>We have now moved to SCAFFQ9 from SCAFFQ6. We are encouraging a short answer to move on... the process is taking a fair amount of time. It is adapting the questions to the context, and doing a good job. We are answering one of the sub questions and leaving the others.</p> <p>We have gone from SCAFFQ9 to SCAFFQ11. The tool is giving more examples than we have perhaps seen it give before... e.g. things to consider when thinking etc. It's going well but it is time consuming.</p> <p>Gone from from SCAFFQ11 to something that it has made called SCAFFQ13.1. The sense of the questions have been changed a little. Party answered.</p> <p>We were sent to SCAFFQ14, we have asked to move on with the process... to get things moving to the next phase.</p> <p>It has made a brief synthesis of the conversation.. Hasn't given proper options, e.g. not offering metaphor or narrative mode. We are telling it to enter 'narrative mode'. To see what it does.</p>		
--	--	---	--	--



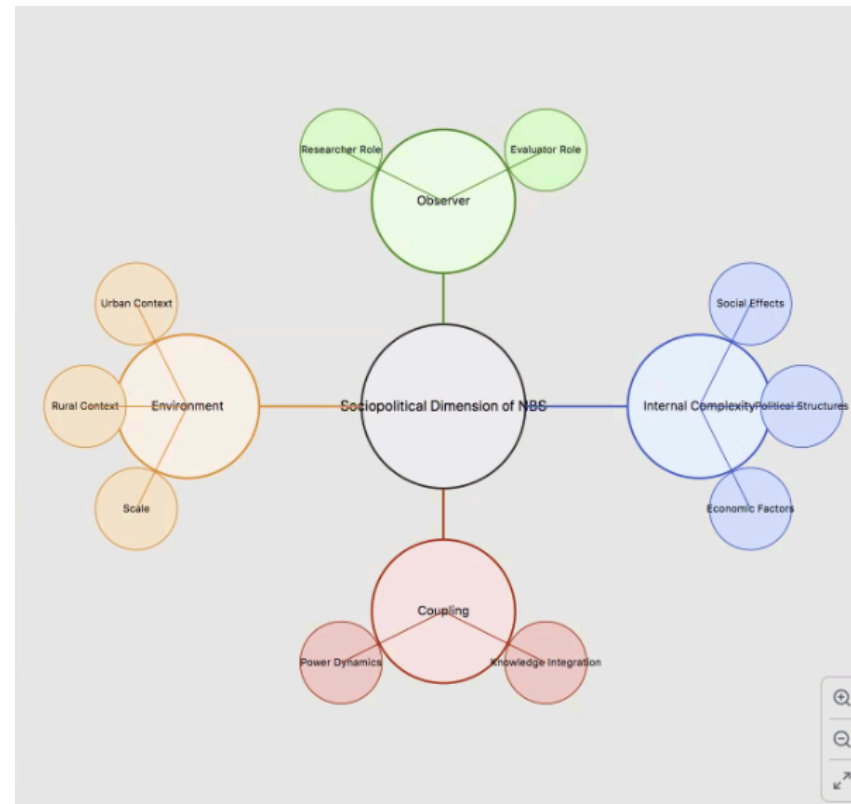
This is data created in the context of the project COMPLEX THINKING & AI

Category: Major log

		<p>Ok, it is making a narrative as an artefact after we told it to. It's a fairly comprehensive narrative. User is having a look at the narrative to see if they think it's good. They think the synthesis is great, reflects the answers and the questions. We are now going to step 3.1 to see if it will do it.</p> <p>It's gone to 3.1, offering evaluation, we are proceeding. We are getting the evaluation. Evaluation varies from mod to high, mostly mod. Evaluation looks good, in line with previous experiments/tests.</p> <p>We asked it to proceed, it went to the map, without asking (it maybe should ask?). Which is a bit of a shame as we might have skipped that. Map doesn't look that useful really - very high level, not really very informative. Although User seems to get something out of it so perhaps it is more informative to them than it looks.</p>		
--	--	---	--	--



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log



We are asking to proceed... see where we go. Not doing anything else with the map. It is suggesting activities... Thought experiment, stakeholder analysis, case study, and so on. User is considering activities... isn't really saying what it is doing, e.g. where it is in the process.

We have skipped activities, proceeding... It has skipped metaphor mode entirely, was not offered, and is offering relationality. We are



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>going to press on with Relationality as that is an option and might be interesting.</p> <p>Ok we are now at R.F.Q1, seems to be more like expected behaviour at this point. Feedback on this question is that User feels they have answered this already. It has also gone back to calling it TSol instead of the name of the system. We are wondering if it has lost some context somehow. We are saying that we think we have answered this question already, to see what it does... 'feels' different qualitatively.</p> <p>Note: no warning on chat length or tokens.</p> <p>So we have gone from S.F.Q1 to S.F.Q3... User answering Q3. Moved to S.F.Q.5, offered 3 sub Qs, answering one.</p> <p>Note: we now hit a warning on chat length.</p> <p>In the first part, it adapted the questions very well to the answers given. I'm not sure if it adapted too much (for example, question SC.SVD.DI.VT.SCAFF.Q11 is a question about types of information. Adapted, it became a question about how to balance power dynamics. (Cf. Original question: SC.SVD.DI.VT.SCAFFQ11.Could you provide further detail on the types of information used, such as experiential, cognitive-emotional,biographical, numerical, sensorial variety, visual mappings, historical and mythologies? Question asked: SC.SVD.DI.VT.SCAFFQ11] Given the information asymmetry you've described between experts and community members, what types of information or knowledge-sharing approaches do you think could help balance these power dynamics in NBS projects?)</p>		
17.10.2024 developer.int er.user#1 ;	Test with exter.user#1	(Note: there was some background noise that might interfere with the audio transcription) TSol; user named his TSol "uncertainties of the future of work", the	Used version 9 SVD and version 3	

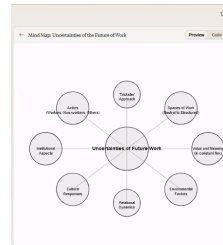


This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

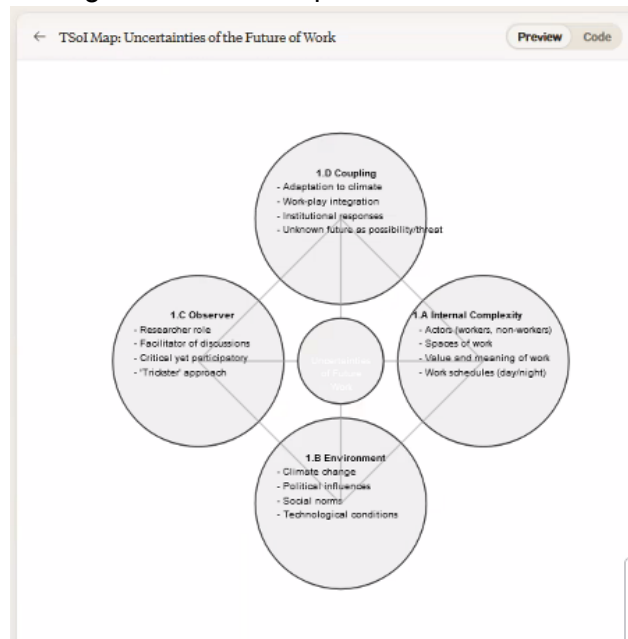
<p>tech.inter.us er#1</p> <p>Test with inter.user#1 highlights</p>		<p>tool keeps referring to TSol, and occasionally uses the name also. This is fine!</p> <p>OUTPUT: The Claude output is verbose. Could help keeping within the token limits by keeping Claude's responses shorter</p> <p>QUESTION CODES: these are long and technical, we could consider hiding them from the user.</p> <p>METAPHOR mode "The Trickster" metaphor was well received</p> <p>NARRATIVE SCAFF: asked for step-by-step or to upload. Chose Step-by-step.</p> <p>Q: [...IQ.FQ1] uses examples of basic elements such as individuals and groups. This might be a place where customisation to different domains might be pertinent.</p> <p>QUESTIONS: can ask multiple follow-up questions (e.g. for clarification/expansion) at once.</p> <p>Moved on well from basic elements of TSol to the Environment.</p> <p>NOTE: Environment was interpreted (literally) by user as the natural environment, but it seemed appropriate in the context</p> <p>Prompted to move on to the Observer.</p> <p>Claude linked responses to the Trickster metaphor generated earlier.</p> <p>It naturally asked to move on to the Coupling question.</p> <p>Very nice responses to the Coupling question, particularly in relation to the observers dynamic relationship with the TSol.</p> <p>Moving on to Narrative Mode.</p> <p>Amos responds that one aspect is less interesting.Tool creates a new version of the narrative in response.</p> <p>The visual synthesis (Mind Map) does not include the base dimensions - it is just the key elements of content in a star diagram form.</p>	<p>general instructions</p> <p>PILOT TEST_exter.u ser#1_17.10.2 024_chat transcript and reflection</p>	
---	--	--	--	--



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log



Moved on to Evaluation: moderate to high complexity.
Offers areas for development. User in agreement.
Moving on...to TSol Map





This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>Use noted that the diagram is more like an elaborate list than a map per se.</p> <p>We then hit the message limit: with only 2 remaining. Decided to download the narrative artefact and to try to restart in another account....</p> <p>Chat link: https://claude.site/artifacts/2741292f-1551-4666-89d1-66ac28baf09</p> <p>—BREAK—</p>		
17.10.2024 developer.int er.user#1 ; tech.inter.us er#1 developer.int er.user#2 ; tech.inter.us er#2	exter.user#1 and	<p>exter.user#1 User started to explain how he would find it useful to have some overview of where he is in the process. Felt that the tool was helpful Did not want to go too far Mapping was a bit obvious Would like a more complicated model - but this would take more time Probably would be easier with a more specific context - his was more abstract. Tool did challenge to be more concrete</p> <p>exter.user#2 First experience with AI tools; expectation that her answer would be answered by too. TSol was about socio-political <something>.... Suggestions offered for methodologies and concrete examples</p>	PILOT TEST_exter.us er#1_17.10.20 24_chat transcript and reflection	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>More information about context would likely generate better responses Justifications of relationships?: was interesting way to end exercise. But not entirely clear about what it meant...</p> <p>(Response) USER 2 session went off protocol , need to find out why. It went quickly into relationality (did not get there with USER 1 - but we had to change account midway) User 1 likely was put into relationality too early... Relationality is more challenging.</p> <p>User 1 Presenting docs that establish context, might allow to get to the more interesting/challenging/novel stuff (e.g. relations) earlier Undertook exercise with the intent of challenging his own thinking.</p> <p>(Response): did the tool make you think? How well did it couple with you?</p> <p>User 1 would need to review the questions, and perhaps work on his own, or to add an order with them.</p> <p>(Response): we will provide a log of the conversation, with an added column to add some annotations/reflections</p>		
17.10.2024 developer.int er.user#1 ; tech.inter.us er#1	Restarting Test with exter.user#1	<p>(picking up after a break)</p> <p>Prompted to go straight into “scaffolding relational movements” It picked up the right point in the protocol, and previewing a cluster of questions in the preamble.</p>	PILOT TEST_exter.u ser#1_17.10.2 024_chat transcript and	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>On continue, it asked the first question: Nature of relations [NOTE] It skipped Phase 1, Step 1.1 and went straight to Step 1.2</p> <p>In step 1.2 did not ask the user if they had a relation they wished to start with. There is error in step 1.2 instruction to correct</p> <p>Note in test we jumped to RM skipping RE Note: need refine step 1.2</p> <p>Jumped to step 2.1 because user could not answer At step 2.1 did not ask the user which mode it preferred</p>	reflection	
<p>17.10.2024 developer.int er.user#1 ; tech.inter.us er#1</p>	<p>Feedback exter.user#1</p>	<p>(started a bit late on this - sorry)</p> <p>exter.user#1</p> <p>Can see applications that become more like a day-to-day dialogue and/or interaction.</p> <p><something about the structure being flat???</p> <p>More direct and straightforward personality would be welcomed - even to the extent of like a programming language: IF/THEN; however could offer other styles of personality</p> <p>Try to break the formulaic type of LLM chat interaction, of flattering, reiteration, in-line summaries etc. To something more streamlined.</p>		



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>For a daily tool, it needs to be leaner.</p> <p>Create a playground Could be something unique</p> <p>Create a kind of learning diary. Could be less intensive, more like an email exchange. (Different tempos or modes)</p> <p>Passive collection could be interesting / useful</p> <p>Liked the suggestions of physical activities - as an example facilitation or support</p> <p>Keeping the conversation appropriately focused.</p> <p>Might be interesting to have a moderator (or multiple personas for different facets)</p>		
18.10.24 developer.int er.user#2	Test exeter.user# 4	<p>The protocol worked very well at the beginning. We used it in English, but we visualised the chat in Portuguese, with google translate plugged in</p> <p>The dynamics of using me as a pair were very interesting, as they were thinking together. (Almost as with the second human?)</p> <p>The metaphorical mode was also very useful and triggered reflection and resonated a lot with the users.</p> <p>When entering narrative mode, it asked the users to make the narrative, instead of making the narrative itself. It made sense,</p>	<p>Used protocol V11 for SCAFF_SVD</p> <p>PILOT TEST_exter.user#4_PART I_18.10.2024_chat log and reflections</p>	



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>because until now, there had been very few elements for it to narrate. This was the opportunity to give concrete elements about TSol. (The first question, about the relationship with the case, was understood and answered in terms of emotional difficulties brought on by the case, without specifying concretely what they were talking about.)</p> <p>The pair discussion triggered by the questions seems very productive, generating reflective dialogues. From time to time, they disagreed and engaged in a discussion about the case. It was a nice conversation on different understandings of each member of the team. We decided to present both understandings to the tool, instead of trying to reach a consensus to respond.</p> <p>Because they had chosen the narrative mode rather than dialogue [STEP 2.1], they later realised that they had given little information about the case and chose to give more information, enriching the narrative. The protocol correctly provided scaffolding questions at this stage.</p> <p>Table was not very clear/helpful.</p> <p>seems to have skipped the [ACTIVE SCAFF TSol MODE] [STEP 3.2]</p> <p>After the scaffolding question, it tried to update the table. They asked to visualise the information in another way and we decided to ask for the map mode. They liked the map; from the map, they see that the cultural dimensions and social norms are really lacking and it makes sense to point this out. This seemed to be particularly useful for the team (they are preparing a report of the case). The tool chose the correct question on this topic. The team said it was very useful and relevant</p>		
--	--	---	--	--



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>In STEP 3.3, the question [SC.SVD.MD.E.SCAFFQ6] <i>Could you identify any macro-level social and cultural norms, discourses, or practices that might be influencing this Case? Consider things like societal expectations about family structure, parenting, gender roles, or attitudes towards social services and intervention:</i> they asked themselves if it was about the family norms or theirs, as a team. It triggered a productive discussion between them. They said they had several ideas, I mentioned the option of [notes mode].</p> <p>When we used the [note mode], it offered some questions [DSCAFF mode] related to the note, but it seems to have created the questions, not using any questions from the list. I decided to let them answer to see if it goes back to protocol (and they seemed useful questions for them)</p> <p>It offered another table (it generated 3 so far), again not very useful. It registered the notes, but it didn't generate an artifact (check how to recuperate them when needed). The team liked the [notes mode], each member had more than one note to add.</p> <p>We asked for the activities. At this point, they wanted to created another metaphor to complement the first one (and it was one of the suggestions of activities)</p>		
	Follow-up analyses of tests	<p>At the end of the interaction I (developer.int.user#1) asked the tool for this Hi, Hi can you please give me an overview of the steps that you followed or did not follow in the conversation in relation to the protocols that were given in the project files The response is interesting and shows it has been following the</p>		



		<p>protocol</p> <p>Certainly, I'd be happy to provide an overview of the steps I followed in our conversation, based on the protocols given in the project files. I'll go through the phases and steps, indicating which ones were completed and which were not.</p> <p>[PHASE 0] - [STEP 0.1]: Completed. I posed the procedural questions to clarify working conditions. - [STEP 0.2]: Completed. I offered a summary of different scaffolding modes available.</p> <p>[PHASE 1] - [STEP 1.1]: Completed. I used the DIALOGICAL SCAFFOLDING mode and posed the foundational modulating questions. We also created the initial metaphor (high-speed train).</p> <p>[PHASE 2] - [STEP 2.1]: Completed. I asked if you wanted to engage in step-by-step scaffolding or produce a narrative. - [STEP 2.2]: Completed. You chose to produce a narrative, which you provided.</p> <p>[PHASE 3] - [STEP 3.1]: Completed. I offered and performed an evaluation of the complexity of thinking based on your narrative. - [STEP 3.2]: Not explicitly completed. We didn't directly use the ACTIVE SCAFF TSol MODE at this point. - [STEP 3.3]: Partially completed. We engaged in some dialogical</p>		
--	--	--	--	--



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>scaffolding, focusing on social and cultural norms.</p> <ul style="list-style-type: none">- [STEP 3.4]: Completed. I presented an updated TSol_MAP. <p>[PHASE 4]</p> <ul style="list-style-type: none">- [STEP 4.1]: Partially initiated. I began to introduce scaffolding modulating questions, but we shifted direction based on your request.- [STEP 4.2]: Completed. I offered several activities in the ACTIVE SCAFF TSol MODE.- [STEP 4.3] to [STEP 4.8]: Not completed. We focused on updating the metaphor instead. <p>[PHASE 5]: Not reached.</p> <p>Overall, we followed the general structure of the protocol, but we adapted based on your preferences and requests. We focused more on certain aspects (like metaphor exploration and mapping) and less on others (like some of the later scaffolding steps). The conversation remained flexible to your needs and interests while still aiming to increase the complexity of thinking about the case.</p>		
18.10.2024 developer.int er.user#1	TEST external exter.user#3	<p>The protocol worked well . They interacted in portuguese. There was a very enthusiastic reaction to the metaphor which had great impact and to the recap and organisation of the information which the team found was very structuring and help them think more clearly. As they were reading the information sent back to then in organised form they would make comments like “that is exactly this, that’s it” “correct” but seemed to appreciate the structuring. They chose to build the narrative step by step and found the steps very</p>	Used protocol V11 for SCAFF_SVD and version 4 general instructions	<p>Additional notes Post reflections from developer.inter.us er#1(applied also to test 4)</p> <p>-The metaphors</p>



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>useful.</p> <p>The team members talks amongst them before inputting the response on Claude and so there is some collective thinking taking place between them.</p> <p>They were making questions like “can we write so much” (they were providing detailed information). I let them continue because it seemed useful to them as they were thinking on the fly. They ere talking to each other as writing.</p> <p>Noticed error step 2.2. Is listed twice instead of 2.2 nd 2.3. It skipped the first 2.2, which was narrative and offered synth and TSoi_Table Choose SYNTh optin 1 and table. Error in protocol because it was to be synth 2 as the protocol had it. Offered table,\ BUT... Team loved it!! NOTE> CORRECT PROTOCOL TO GIVE USER OPTION OF TYPE OF SYNTHESSES THEY WANT</p> <p>The format of the table is not very pretty or pleasant...</p> <p>Neec heck step 2.1 because it seems to have stayed too long in scaffolding, after 2.2 it returned to scaffolding (step 2.1??)</p> <p>Did not offer the narrative!! (repeated step 2.2. , only did the second)</p> <p>Team members talked to each other in deciding how to respond to the question exchanging contacts such as “i think it is lille this. What do you think?”</p> <p>Há poucas para pensar “hum...” equipa pensa em voz alta e confirmam umas com as outras</p> <p>Without asking offered updated TSoI table</p> <p>The summaries offered allow the team to find other nuances in their narrative</p> <p>It seems to continue extensively on the scaffolding</p>	<p>PILOT TEST_exter.us er#3__PART I_18.10.2024_ chat log and reflections</p>	<p>are very generative</p> <p>And the teams found they could be useful as a way f the sharing the thinking and triggering reflections in the families</p> <p>Claude coupled well with the metaphors and expands them opening new possibilities for insight.</p> <p>Note: we should consider developing further the metaphor mode because it could in a smoother way encapsalute several properties or movements of the thinking in the “unpacking “ or exploration of the</p>
--	--	--	--	---



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>NOTE: NEED TO ADD LIMIT OF QUESTIONS? CHECK PREVIOUS VERSIONS OF THE PROCOL WE ARE ON V11 NOW</p> <p>The teams thinks and reflections on the questions</p> <p>It is updating too often the TSol</p> <p>I PROMPTED TO DO [TSol_NARRATIVE]</p> <p>They loved the narrative. It was a very narrative (story like) mode and wanted to show to their family. I think it might have adjusted to their own narrative style used in the dialogue (hypotheses)</p> <p>The team got tearful with the narrative and though it was very beautiful and could be useful to the mother</p> <p>In the EVAL repeated MD and variety of contents</p> <p>It seemes to have moved to scaffolding phase 3.3 but it was not clear.</p> <p>Not named phase but questions were from there I asked to perform eval and move to step 3.1 Moves to 3.2</p> <p>There is an error in step 3.2 TSOi but acgtivities</p> <p>Step 3.2 and 4.2 are the same. There is an error Enter step 3.3.</p>		<p>metaphor We can refine the protocol for how the tool explores the metaphor and offers reflections based on them</p>
--	--	--	--	--



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>The use with guidance seems useful</p> <p>The team dynamic is interesting. They engage in discussing with each other to respond to the questions</p> <p>Claude offers interesting and relevant insights not moving too much from what was given but adding or highlighting nuances</p> <p>From step 3.3. Jumped to modulating questions of 4.1. Although they were connected and reasonable for context of talk. Will it go back to 3.4??</p> <p>I promoted to ask how they could increase complexity of their thinking. They found activities very useful</p> <p>Metaphors very useful\the team thinks they can use this metaphor to use with family</p> <p>We asked for update narrative and it integrated the metaphor. It is very story like and rich. UAU says the team!!</p> <p>The TSOi integrates the metaphor</p> <p>NOTE: NEED TO MAKE CHANGE TO HAVE EVALUATION REPORT AS AN ARTEFACT</p>		
24.10.24 developer.int er.user#2	exter.user#3 and 4 on Relations	<p>We did the test with the two teams in the same zoom room, each using a different account.</p> <p>Both teams started by interacting in Portuguese, but team 4 had the</p>	PILOT TEST_exter.us er3_part II_24.10.2024_	Make variations to the protocols for: -continuous use in the same session



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log

		<p>documents uploaded in English. At the same time, they had installed a translation plugin. There was some linguistic confusion, which was corrected during the chat.</p> <p>For both teams, we uploaded the updated narrative from the last session and a table (in the case of team 4, we chose the coupling analysis table). We felt that the integration of the information from the previous session was insufficient for both teams.</p> <p>Both teams also reported feeling a disconnect with the previous session, and a certain 'bombardment of questions' in the initial phase. They were frustrated that they didn't get to other modes straight away, with suggestions for exercises or maps. They felt that in this session they received less feedback than in the previous session, and the answers integrated the information less. Even so, they said at various times "this question makes me think" or "how interesting!"</p> <p>Team 3 interacted a lot with the feedback they received, asking for the maps to be updated and for new dimensions to be integrated into the metaphor. The protocol responded very well</p> <p>The fact that the two teams were together allowed them to compare their answers and demand different things from the tool (when they saw the format of the other team's narrative, for example, team 4 asked for a more story-like narrative or a map with a different format).</p> <p>The protocol seems to have been maintained the whole time. Even when team 3 wanted to explore the suggested exercises in more detail, it did so by returning to step 1.2 and suggesting a modulating question.</p>	chat log and reflections	<p>-</p> <p>only for one property in the session (here, have a warm-up mode, to generate context and engagement)</p> <p>Offer two narrative modes in the protocol for the user to choose from (more synthetic or more story-like). Do the same offer for the synthesis mode.</p> <p>Check if the protocol on relations has an indication to return to previous steps</p> <p>Check if the two protocols (relations and SVD) have the same description of modes</p>
--	--	--	--------------------------	---



This is data created in the context of the project COMPLEX THINKING & AI
Category: Major log