

Table C: Raw prompt-level LLM detections (smell IDs per prompt) and corresponding expert-identified UX smells (ground truth).

App	Model	Modality	# Prompts Sent	# Smells Detected for each Prompt sequentially by the LLM	ID of Detected Smells by the LLM for each prompt	ID of Actual available smells detected by the experts (expert-identified UX smells)
Medicine Time	GPT-4	Log	1	3	19,30,40	2,7,19,30,40
	GPT-5			5	2,7,19,30,40	
	GPT-5 Thinking			4	7,19,30,40	
	GPT-4	GUI	1	1	18	34
	GPT-5			1	34	
	GPT-5 Thinking			1	34	
	GPT-4	Code	6 (one per UI)	4	3,4,5,15	3,4,5,15
	GPT-5			5	3,5,11,15,31	3,11,15,31
				3	3,11,14	3,11,15,28,31
				2	3,15	3,15
				4	3,11,15,27	3,11,15,28,31
				2	3,15	3,11,15,17,31
				4	3,4,5,15	
				5	3,5,11,15,31	
				8	3,5,11,15,27,28,31,35	
				5	3,5,11,15,31	
				6	3,5,11,27,28,31	
				4	3,11,15,31	
				4	3,4,5,15	
				GPT-5 Thinking	2	3,15
	1				15	
	2				3,15	
	3				3,15,27	
	Daily Dozen			GPT-4	Log	1
GPT-5		3	7,25,30			
GPT-5 Thinking		3	2,7,25			
GPT-4		GUI	1	1	34	18,34
GPT-5				2	18,34	
GPT-5 Thinking				1	34	

GPT-4o	Code	12 (one per UI)	2	3,15	3,15
			1	11	14,24
			4	4,5,15,27	3,4,5,15,27,28,31
			2	3,14	3,14
			4	3,4,5,11	3,11,15,16,28
			3	3,15,28	3,15,11,28
			4	3,5,15,28	3,14,16,17,28,31
			3	3,15,28	3,11,15,28,31
			4	3,11,15,28	3,11,15,28
			3	3,15,28	3,15,28
			3	3,15,28	3,14,28
			3	3,15,28	3,15,28
GPT-5			2	3,14	
			5	3,11,15,28,31	
			4	4,5,28,31	
			2	3,14	
			5	3,11,14,16,28	
			4	3,11,14,28	
			5	3,11,14,16,28	
			5	3,11,14,28,31	
			5	3,11,14,28,31	
			5	3,11,14,28,31	
			5	3,11,14,28,31	
			5	3,11,14,28,31	
GPT-5 Thinking			2	3,14	
			1	14	
			2	14,24	
			2	3,14	
			3	11,14,24	
			2	3,14	
			2	14,17	
			2	3,14	
			2	3,14	
			2	3,14	
			2	3,14	
			2	3,14	
To-Do List	GPT-4o	Log	1	3	1,21,30
					1,21,30
	GPT-5			3	1,21,30
				3	1,21,30
GPT-5 Thinking	GPT-4o	GUI	1	1	18
					0
GPT-5				0	0

				0	0	
	GPT-5 Thinking					
	GPT-4o	Code	2 (one per UI)	2 4	3,4 3,15,28,31	3,4,15,31 3,15,17,28,31
	GPT-5			5 3 2	3,4,5,15,31 3,15,17 3,15	
	GPT-5 Thinking			3	3,15,17	
GroceryStore	GPT-4o	Log	1	1	30	2,7,30
	GPT-5			3 2	2,7,30 2,30	
	GPT-5 Thinking					
	GPT-4o	GUI	1	5	8,9,18,23,34	8,9,18,23,34
	GPT-5			2 3	18,34 8,18,34	
	GPT-5 Thinking					
	GPT-4o	Code	18 (one per UI)	3 2 3 3 3 3 5 3 3 3 3 3 3 3 4 2 2 2	3,5,11 3,11 3,4,5 3,5,15 3,4,15 3,11,15 3,4,11,15,27 3,4,15 3,11,15 3,11,15 3,11,15 3,11,15 3,11,15 3,11,15 3,11,15 3,4,5,11 3,11 3,11 3,11	3,11,15,28 3,11,15,28 3,15,24,28 3,15,28 3,4,5,15,28 3,11,15,28 3,11,15,27,28 11,15,28 3,11,15,28 3,11,15,28 3,11,15,28 3,11,15,28 3,5,11,15,28 3,11,15,28 11,15,28 11,15,28 3,11,15 3,11,15,28 11,15,17,24,28

	GPT-5			3	3,11,15	
				3	3,11,15	
				1	15	
				2	3,15	
				4	3,4,5,15	
				4	3,11,15,31	
				7	3,4,5,11,15,27,28	
				2	3,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
	GPT-5 Thinking			3	3,11,15	
				3	3,11,15	
				1	15	
				2	3,15	
				2	3,15	
				3	3,11,15	
				3	11,15,27	
				1	15	
				3	3,11,15	
				3	3,11,15	
				3	3,11,15	
				4	3,5,11,15	
				3	3,11,15	
				2	11,15	
				3	3,11,15	
				4	3,11,15,31	
				3	3,11,15	
				4	3,11,15,17	
				4	1,7,19,40	
CometChat	GPT-4o	Log	1	0	0	0
	GPT-5			0	0	
	GPT-5 Thinking			0	0	
	GPT-4o	GUI	1	4	8,18,23,34	18
	GPT-5			1	18	
	GPT-5 Thinking			1	18	

GPT-4o	Code	32 (one per UI)	2	3,11	3,15
			3	4,11,31	3,11,15,16,28
			4	3,4,5,11	3,5,15,28,37
			2	3,5	3,15,28
			2	3,15	3,15,28
			2	3,15	3,15,28
			2	3,15	3,15,27,28
			2	3,15	3,15,28
			3	3,15,24	3,15,24
			6	3,5,15,27,28,37	3,5,15,27,28,37
			5	3,5,15,24,28	3,5,15,24,28
			5	3,11,15,24,28	3,11,15,24,28
			6	3,5,11,15,24,28	3,5,11,15,24,28
			5	3,11,15,24,28	3,11,15,24,28
			6	3,11,15,24,27,28	3,11,15,24,27,28
			4	3,11,15,24	3,11,15,24,28
			2	3,15	3,15,16,24,28
			7	3,4,11,15,16,24,28	3,4,11,13,15,16,24,28
			2	3,15	3,15
			4	3,11,15,37	3,11,15,37
			0	0	0
			2	3,15	3,15
			2	3,15	3,15
			2	3,15	3,15
			4	3,15,28,31	3,15
			1	11	3,11,15,27
			3	3,4,5	3,15,28
			1	11	3,11,15,28,31
			3	3,4,5	3,15,28
			3	3,4,5	3,15,28
			4	3,15,17,31	3,15,28
			3	3,17,31	3,15,17,28

GPT-5	2	3,15
	4	3,4,11,15
	1	3
	2	3,15
	2	3,15
	2	3,15
	3	3,27,28
	2	3,28
	3	3,15,24
	4	3,11,15,37
	5	3,5,15,24,28
	7	3,5,11,15,24,28,37
	6	3,5,11,15,24,28
	5	3,11,15,24,28
	7	3,5,11,15,24,27,28
	5	3,5,15,24,28
	5	3,15,16,24,28
	6	4,5,11,13,15,16
	2	3,15
	3	3,11,37
	2	3,15
	2	3,15
	2	3,15
	2	3,15
	2	3,15
	4	3,11,15,27
	3	3,15,28
	5	3,11,15,28,31
	2	3,15
	1	3
	3	3,15,28
	2	3,17

	GPT-5 Thinking			2	3,15	
				3	3,11,37	
				3	3,5,15	
				2	3,15	
				2	3,15	
				2	3,15	
				1	3	
				1	3	
				3	3,15,24	
				5	3,5,15,28,37	
				4	3,4,15,24	
				3	3,11,15	
				-	-	
				3	3,11,15	
				2	3,11	
				1	3	
				2	3,15	
				7	4,5,11,13,15,16,31	
				2	3,15	
				3	3,11,37	
				1	3	
				2	3,15	
				2	3,15	
				2	3,15	
				2	3,15	
				-	-	
				2	3,5	
				4	3,11,15,31	
				2	3,15	
				2	3,15	
				1	3	
				3	3,15,17	
2048-android	GPT-4o	Log	1	1	30	30
	GPT-5			0	0	
	GPT-5 Thinking			0	0	
	GPT-4o	GUI	1	2	18,23	18
	GPT-5			1	18	
	GPT-5 Thinking			1	18	
	GPT-4o	Code	1 (one per UI)	1	15	15
	GPT-5			1	15	
	GPT-5 Thinking			1	15	
	GPT-5 Thinking					
WeatherApp	GPT-4o	Log	1	4	7,20,22,30	20,30

				GPT-5		2		20,30	
						2		20,30	
				GPT-5 Thinking					
				GPT-4o	GUI	1	3	8,9,18	8,9,18
				GPT-5			3	8,9,18	
							1	18	
				GPT-5 Thinking					
				GPT-4o	Code	2 (one per UI)	5	4,11,15,24,28	3,11,15,37
							2	15,17	15,17,31
				GPT-5			4	3,11,15,37	
							3	15,17,31	
							4	3,11,15,37	
				GPT-5 Thinking			2	15,17	
Health Care Locator Android App				GPT-4o	Log	1	4	1,7,21,30	0
				GPT-5			0	0	
							0	0	
				GPT-5 Thinking					
				GPT-4o	GUI	1	2	18,23	18
				GPT-5			1	18	
							0	0	
				GPT-5 Thinking					
				GPT-4o	Code	3 (one per UI)	4	3,5,15,28	3,15,28
							3	3,5,31	3,15,28
						6	3,5,16,17,28,31	3,15,17,28,37	
			GPT-5			3	3,15,28		
						2	3,15		
						4	3,15,17,37		
						2	3,15		
				GPT-5 Thinking			2	3,15	
							3	15,17,37	
Tic Tac Toe				GPT-4o	Log	1	1	25	25
				GPT-5			1	25	
							1	25	
				GPT-5 Thinking					
				GPT-4o	GUI	1	0	0	18,23
				GPT-5			2	18,23	
							0	0	
				GPT-5 Thinking					
				GPT-4o	Code	3 (one per UI)	6	3,5,14,17,27,31	3,14,17,27
							4	3,14,17,28	15,27
						4	3,14,17,28	3,15,27	
			GPT-5			4	3,14,17,27		
						1	15		
						0	0		

				4	3,14,17,27	
	GPT-5 Thinking			1	15	
				1	14	
Verifit	GPT-4o	Log	1	2	24,30	30
	GPT-5			1	30	
				1	30	
	GPT-5 Thinking					
	GPT-4o	GUI	1	3	18,23,34	18,23,34
	GPT-5			3	18,23,34	
				2	18,23	
	GPT-5 Thinking					
	GPT-4o	Code	13 (one per UI)	6	3,5,11,15,16,28	4,5,11,15,24
				4	3,4,5,15	3,15,31
				2	3,15	3,15
				2	11,37	11,37
				2	3,5	3,4
				2	3,15	3,15,31
				2	3,15	3,11,15
				1	11	11,15
				1	11	11
				4	3,4,5,15	15
				4	3,4,5,15	0
				3	11,15,17	11,17
				3	5,14,28	3,11,15
	GPT-5			5	4,5,11,15,24	
				3	3,15,31	
				2	3,15	
				2	11,37	
				2	3,4	
				3	3,15,31	
				2	3,15	
				2	11,15	
				1	11	
				1	15	
				0	0	
				1	17	
				3	3,11,15	
				5	4,5,11,15,24	
	GPT-5 Thinking			3	3,14,31	
				2	3,14	
				2	11,37	
				2	3,4	
				2	3,15	
				2	3,11,15	
				1	14	
				0	0	
				1	5	
				1	14	
				2	15,17	
				2	3,15	