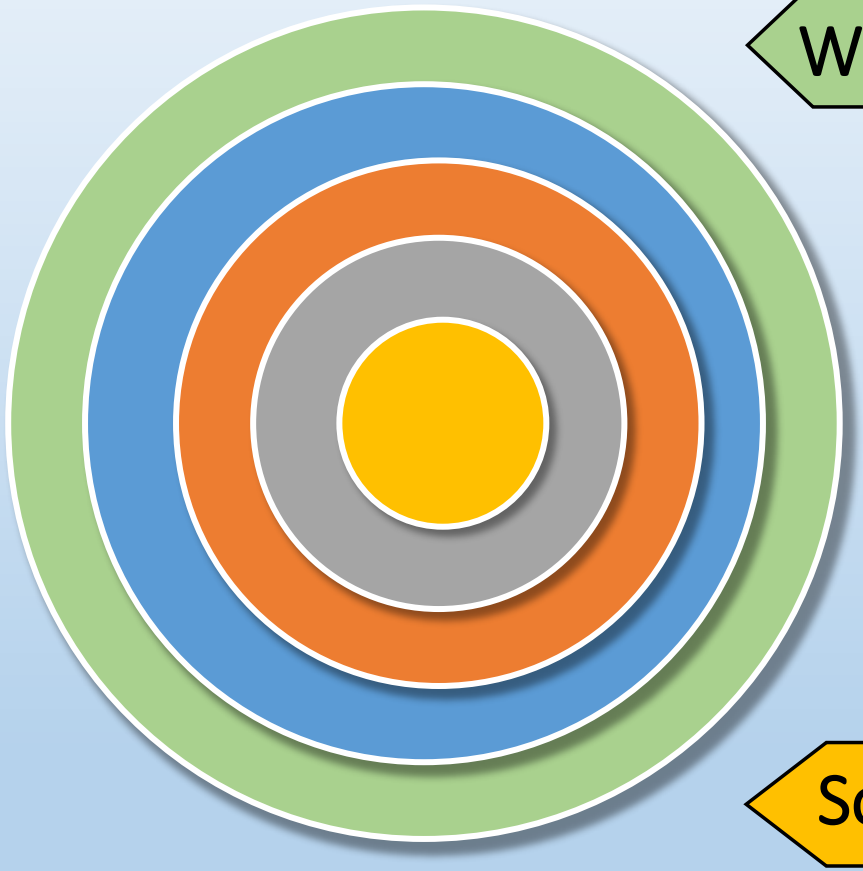
Developing AI-Compatible Instructional Design Datasets: A Computer Science Research Project

Demetre Nadiradze
Furqan Al Ghifari Zulva
Muhammad Umar Farzan
Sohaila Abdalgwad
Emre Kaan Usta

Supervisor: İlkem Kayıcan Dipcin
Date: 12 August 2025

Introduction/Overview



What exactly is Instructional Design?

What is the main challenge facing this industry?

How can we solve this? The opportunity lies in AI.

What's the barrier to effective automation?

So, what is the specific focus of our research?

PURE PROGRAM FOR UNDERGRADUATE RESEARCH

Sabancı Üniversitesi

Research Question

What data structures and formats are most effective for training AI models in instructional design?

PURE PROGRAM FOR UNDERGRADUATE RESEARCH

Sabancı Üniversitesi

Literature Review

Framework	Purpose	Approach	Limitation
ARCHED [1]	Generate & evaluate learning objectives	Modular web tool using GPT-4o via API (LOGS + OAE)	No lesson or assessment generation, no fine-tuning
GAIDE [2]	Full course content co-creation	Human-in-the-loop chat-based process using any GenAI (e.g. ChatGPT)	No automation; not scalable; no fine-tuning

[1] Li, H., Fang, Y., Zhang, S., Lee, S. M., Wang, Y., Trexler, M., & Botelho, A. F. (2025). *ARCHED: A human-centered framework for transparent, responsible, and collaborative AI-assisted instructional design*. arXiv. https://doi.org/10.48550/arXiv.2503.08931

[2] Dickey, E., & Bejarano, A. (2024). *GAIDE: A Framework for Using Generative AI to Assist in Course Content Development*. *2021 IEEE Frontiers in Education Conference (FIE)*, 1–9. https://doi.org/10.1109/fie61694.2024.10893132

PURE PROGRAM FOR UNDERGRADUATE RESEARCH

Sabancı Üniversitesi

Methodology

Data Curation

Collected multiple training sources spanning soft skills such as leadership, problem solving etc pertaining to the corporate world

Content Compression

Summarized into 10-page structured documents using LLM tools and a series of Prompt Engineering

Pedagogical Structuring

Reformatted into 4 key sections: Objectives, Content, Activities, Assessments

JSONL Transformation

Converted the samples collected to machine-readable fine-tuning format

Model Finetuning

Used data to fine-tune Mistral & OpenAI models to generate high quality content

pure PROGRAM FOR UNDERGRADUATE RESEARCH

Sabancı Üniversitesi

Findings – Less is More


```
{
  "title": "Leadership Development Program (LDP)",
  "sector": "Leadership & Management Development",
  "source_provider": "Center for Creative Leadership (CCL)",
  "url": "https://shop.ccl.org/usa/leadership-programs/mid-level-leaders/leadership
  "format": "Blended learning - 5-day intensive within 5-month journey",
  "target_audience": "Mid- to senior-level managers, leaders of managers, 5+ years
  "learning_objectives": [
    "Increase self-awareness and emotional intelligence in complex environments",
    "Develop skills to accelerate team achievement and belonging",
    "Build capacity to operate effectively in organizational systems",
    "Learn to identify opportunities in challenges and create systemic solutions",
    "Create and communicate a compelling vision for organizational direction"
  ],
  "content_structure": "Phase 1: 2-month pre-work with online coursework and assess
  Phase 2: 5-day intensive covering self-awareness, team dynamics, organizational s
  vision setting (Design & Development). Phase 3: 3-month follow-up with coaching s
  Implementation & Evaluation).",
  "assessment_methods": "360-degree feedback assessments, pre/post behavioral evalu
  coaching milestone assessments, capstone project presentation, long-term behavior
  "quality_scores": {
    "objective_clarity": 9,
    "instructional_design_quality": 9,
    "assessment_integration": 8,
    "content_quality_depth": 9,
    "industry_relevance_practicality": 9,
    "overall_score": 8.8
    .
    .
    .
```


Original Data Format

```
{ "conversations": [

{ "from": "system", "value": ......},
{ "from": "human", "value":..... },
{ "from": "gpt", "value": ......}
]}
```


New Data Format

PURE PROGRAM FOR UNDERGRADUATE RESEARCH

Sabancı Üniversitesi

Findings – Less is More

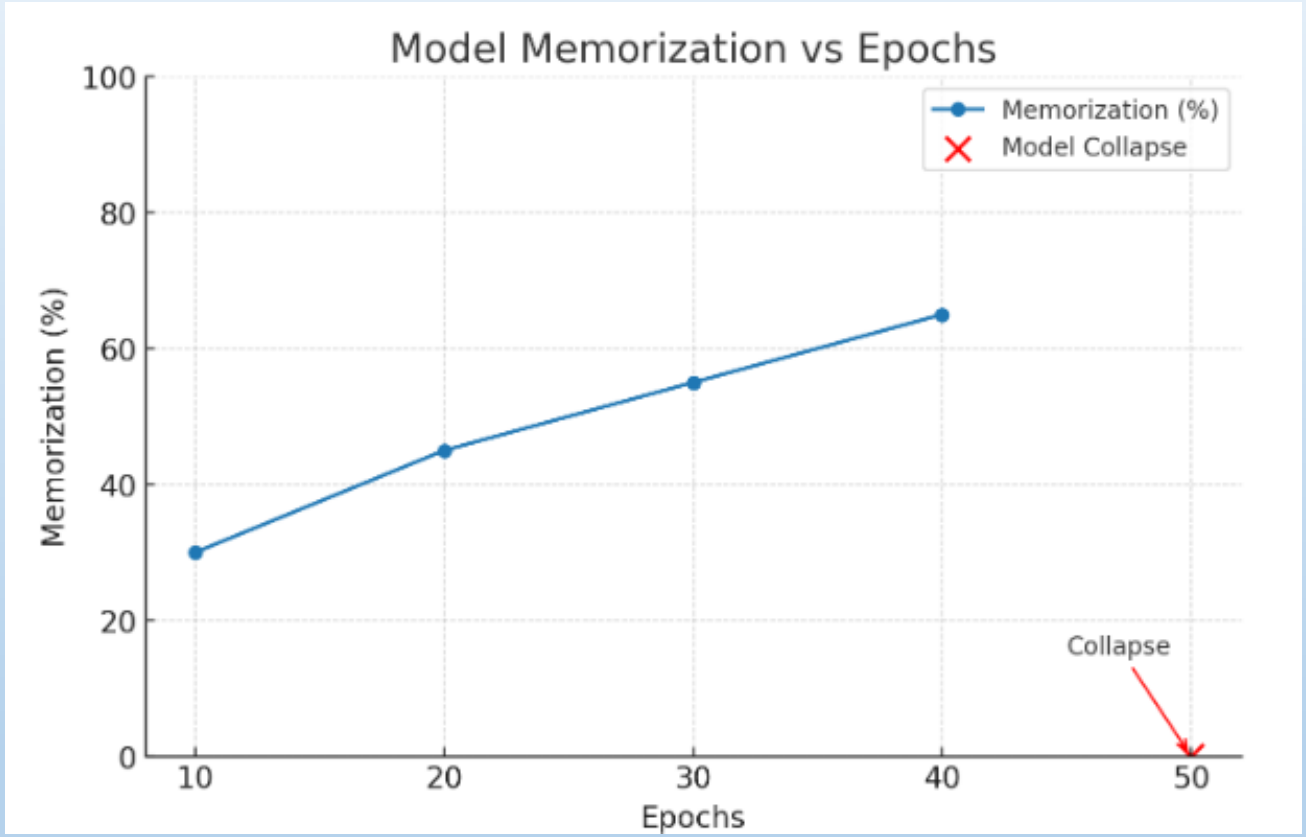
Specializes in one topic, memorizes the book.

Student who learns from one specialized book

Student who learns from various books

Learns about the concepts in depth.

Results – Memorization Trends



Data memorization trend chart.

Results – Comparing Datasets

Version	Response Quality	Response Length (tokens)	Response Time
Version 1 (50 samples)	5/10	~1500	90-120 seconds
Version 5 (300 samples)	8/10	~800	30-60 seconds

Responses were evaluated based on structure, quality of content and quality of assessment methods.
Response Quality, Length and Time represent average values.

pure PROGRAM FOR UNDERGRADUATE RESEARCH

Sabancı Üniversitesi

Conclusion

Created methods for condensing content.

Developing AI-Compatible Instructional Design Datasets

60%
Better Response Quality.

30%
Better Response Time.

{JSONL}

Created an effective data structure.

Gathered 500 high-quality data samples.

Future Research Direction:

- **Larger Dataset Creation.**
- **Comparison of Minstrel & OpenAI models.**

PURE PROGRAM FOR UNDERGRADUATE RESEARCH

Sabancı Üniversitesi