Abstract

Long-term Human-AI collaboration presents challenges that extend beyond the capabilities of either human memory or AI memory alone. While recent research has focused primarily on model performance, prompt engineering, and short-term task completion, relatively little attention has been given to research continuity across extended periods of time.

This study investigates long-term Human-AI research collaboration through a multi-repository research program involving multiple AI systems, repository infrastructure, documentation artifacts, and external memory mechanisms. The study proposes **Human-AI Research Continuity Theory (HARCT)** as a unified explanatory framework: long-term Human-AI research collaboration is fundamentally a problem of research continuity rather than intelligence alone.

HARCT is supported by four theoretical components developed in this study. **Goal Preservation Theory** argues that preserving research goals is more fundamental than preserving information — goals enable reconstruction, while information without direction becomes an unusable archive. **Cognitive Architecture Theory** identifies a three-component distributed system in which humans preserve goals, AI systems perform reasoning and reconstruction, and artifacts preserve context. **Reconstruction Cost Theory** frames documentation quality as an economic variable that directly determines collaboration efficiency. **Externalized Memory** describes how context is stored outside active participants through persistent artifacts.

The findings suggest that sustainable long-term Human-AI research depends not on perfect memory, but on constructing systems that preserve goals, minimize reconstruction costs, and transfer context across time, sessions, and participants.

Keywords: Human-AI Collaboration, Human-AI Research Continuity Theory (HARCT), Goal Preservation, Reconstruction Cost, Cognitive Labor Division, Externalized Memory, Distributed Research Memory, Context Transfer, Long-Term Research Programs

---

1. Introduction

1.1 Motivation

Recent advances in Large Language Models (LLMs) have enabled new forms of Human-AI collaboration.

Most existing discussions focus on short-term interactions, task completion, or isolated conversations.

However, many real-world research projects extend across months or years and require continuity beyond a single session.

In such environments, neither human memory nor AI memory alone is sufficient to preserve the complete research context.

This study addresses that problem through a theoretical framework: **Human-AI Research Continuity Theory (HARCT)**.

HARCT argues that the primary challenge of long-term Human-AI research is not intelligence alone but research continuity — the ability to preserve, restore, and transfer research context across time, sessions, AI systems, and repositories.

The present study emerged from a multi-repository research program involving mathematical theory development, repository-based knowledge management, and collaboration with multiple AI systems. Throughout this process, challenges related to memory preservation, context loss, token limitations, and AI transitions repeatedly appeared. These observations motivated a systematic investigation of long-term Human-AI collaboration.

1.2 Research Problem

Long-term research differs significantly from short-term problem solving.

Research continuity requires the preservation and restoration of knowledge across:

- Time
- Sessions
- Devices
- Repositories
- AI systems

Existing AI systems operate within finite context windows and are unable to maintain complete long-term memory.

Human researchers face similar limitations due to forgetting, interruptions, and cognitive constraints.

As a result, long-term collaboration requires mechanisms that extend beyond the memory capabilities of individual participants.

This study investigates how such mechanisms emerge and how they support research continuity.

1.3 Research Questions

This study addresses the following questions.

RQ1

How can research context be preserved across long-term Human-AI collaboration?

RQ2

How can context be transferred across different AI systems, sessions, and time periods?

RQ3

What role do repositories, documentation, and handover artifacts play in research continuity?

RQ4

How do collaboration protocols emerge during long-term Human-AI interaction?

1.4 Contributions

This study makes five primary contributions.

Contribution 1

It proposes **Human-AI Research Continuity Theory (HARCT)** — a unified framework explaining how long-term Human-AI research becomes possible despite memory limitations.

Contribution 2

It introduces **Goal Preservation Theory** — arguing that preserving research goals is more fundamental than preserving information, and that goals enable reconstruction where information alone cannot.

Contribution 3

It identifies a **Three-Component Cognitive Architecture** for long-term Human-AI research: human goal memory, AI reasoning capability, and artifact-based context storage as a distributed system.

Contribution 4

It proposes **Reconstruction Cost Theory** — framing documentation quality as an economic variable that directly determines collaboration efficiency.

Contribution 5

It presents a real-world case study of long-term Human-AI collaboration conducted over multiple repositories and AI systems, providing empirical grounding for the theoretical framework.

1.5 Scope

This study does not attempt to evaluate the intelligence of AI systems.

Instead, it focuses on collaboration processes.

The primary objective is to understand how research continuity can be achieved despite memory limitations and context fragmentation.

Accordingly, the emphasis is placed on:

- Goal preservation
- Knowledge preservation
- Context transfer
- Collaboration protocols

- Repository-based research management
- Long-term Human-AI interaction

rather than model performance benchmarks.

1.6 Overview of the Study

The remainder of this paper is organized as follows.

Chapters 2 and 3 describe the background literature and research environment.

Chapters 4 through 7 present the theoretical framework:

Chapter 4 introduces Human-AI Research Continuity Theory (HARCT) as the central organizing framework.

Chapter 5 presents Goal Preservation Theory, including Future-Self Collaboration as a mechanism of goal-directed continuity across time.

Chapter 6 describes the three-component Cognitive Architecture underlying HARCT, integrating Human-AI Memory Asymmetry, Cognitive Labor Division, and Repository as Cognitive Infrastructure.

Chapter 7 introduces Reconstruction Cost Theory, explaining why documentation quality is an economic determinant of collaboration efficiency.

Chapters 8 through 11 present empirical observations:

Chapter 8 describes Externalized Memory and AI-to-AI Handover systems.

Chapter 9 examines Token Constraints and Context Loss.

Chapter 10 describes the architecture of long-term research programs.

Chapter 11 presents case studies collected during the development of the Minor Thesis research program.

Chapters 12 and 13 discuss implications and conclusions.

---

2. Background

2.1 Human-AI Collaboration

Recent advances in Large Language Models (LLMs) have enabled new forms of Human-AI collaboration.

AI systems are increasingly used for:

- Writing assistance
- Software development
- Data analysis
- Research support
- Knowledge management

Most existing discussions focus on task completion, productivity gains, or conversational performance.

However, relatively little attention has been given to long-term research collaboration extending across multiple sessions, repositories, and AI systems.

This study focuses specifically on that long-term perspective.

2.2 Context Window Limitations

Modern AI systems operate within finite context windows.

As conversations become longer, earlier information may become unavailable or difficult to recover.

This creates challenges for long-term projects.

Examples include:

- Loss of historical decisions
- Repeated explanations
- Reconstructed reasoning
- Context fragmentation

As project duration increases, context preservation becomes increasingly important.

The present study treats context limitations not merely as technical constraints but as collaboration constraints.

2.3 Knowledge Preservation

Long-term research depends on the preservation of knowledge.

Historically, researchers have relied on:

- Laboratory notebooks
- Research journals
- Technical reports
- Reference archives

These artifacts preserve information beyond the limits of human memory.

The emergence of AI-assisted research introduces new challenges because knowledge must now be preserved not only for humans but also for AI systems participating in the research process.

2.4 Repository-Based Research

Software engineering has long relied on repositories for preserving code and project history.

Repositories provide:

- Persistent storage
- Version control
- Collaboration support
- Historical traceability

In this study, repositories expanded beyond their traditional role.

GitHub repositories became:

- Knowledge archives
- Research memory systems
- Handover platforms
- Collaboration interfaces

As a result, repository infrastructure played a central role in sustaining research continuity.

2.5 Human-AI Memory Asymmetry

Humans and AI systems possess fundamentally different memory characteristics.

Human researchers typically possess:

Strengths

- Long-term experiential memory
- Intuition
- Goal persistence
- Conceptual understanding

Limitations

- Forgetting
- Incomplete recall
- Cognitive overload

AI systems typically possess:

Strengths

- Rapid retrieval within active context
- Large-scale text processing
- Pattern recognition

Limitations

- Session boundaries
- Context window limits
- Lack of persistent memory

Long-term collaboration therefore requires mechanisms that compensate for both types of limitations.

2.6 Multi-Agent Research Environments

The research program examined in this study involved multiple AI systems.

Examples included:

- ChatGPT
- Claude
- ANTIGRAVITY

Each system participated at different stages of the research process.

The involvement of multiple AI systems introduced additional challenges:

- Context transfer
- Knowledge consistency
- Handover management
- Documentation quality

These challenges motivated the development of structured collaboration practices.

2.7 From Information Storage to Context Transfer

Traditional knowledge management often focuses on information storage.

However, observations from this study suggest that storage alone is insufficient.

The critical problem is not whether information exists.

The critical problem is whether context can be restored when needed.

As a result, the study emphasizes Context Transfer rather than simple information retention.

Under this perspective:

- README files
- Research logs
- Handover documents
- Repository structures
- Reminder systems

are viewed primarily as context transfer mechanisms.

2.8 Research Gap

Existing Human-AI collaboration literature frequently investigates:

- Task performance
- User satisfaction
- Prompt engineering
- Human oversight

However, fewer studies examine:

- Long-term collaboration
- Research continuity
- AI-to-AI handover
- Future-self collaboration
- Repository-centered research workflows
- Distributed research memory

The present study addresses this gap through a case study of a multi-repository, multi-AI research program.

2.9 Conceptual Foundation

Three foundational concepts underpin this study.

Externalized Memory

Research context preserved outside both human memory and active AI context.

Distributed Research Memory

Research memory distributed across humans, AI systems, repositories, and artifacts.

Context Transfer

The process through which research context moves across time, sessions, systems, and collaborators.

These concepts are developed and unified within the HARCT framework in subsequent chapters.

---

3. Research Environment

3.1 Overview

This study was conducted within an active long-term research program involving a human researcher, multiple AI systems, repository infrastructure, and external memory artifacts.

Unlike controlled laboratory experiments, the research environment emerged organically during the development of multiple interconnected research projects.

The objective of this chapter is to describe the components that participated in the collaboration process.

3.2 Human Researcher

The central participant was a human researcher responsible for:

- Defining research goals

- Identifying research problems
- Evaluating hypotheses
- Organizing repositories
- Coordinating AI systems
- Preserving long-term objectives

Although AI systems contributed extensively throughout the research process, strategic direction remained under human control.

The human researcher also functioned as the primary integrator of knowledge generated across multiple systems.

3.3 AI Systems

Multiple AI systems participated in the research program.

Each system contributed different capabilities.

ChatGPT

Primary functions included:

- Research discussion
- Theory development
- Conceptual analysis
- Research planning
- Context integration

ChatGPT frequently acted as a coordination layer connecting ideas originating from different repositories and conversations.

Claude

Primary functions included:

- Repository organization
- Documentation refinement
- Structural editing
- Research asset management

Claude was often used for repository-scale operations and document restructuring.

ANTIGRAVITY

Primary functions included:

- Large-scale repository analysis
- Bulk document processing
- Long-form generation tasks

ANTIGRAVITY was frequently utilized when processing large collections of files and repositories.

Future AI Systems

The research architecture was intentionally designed to remain compatible with future AI systems.

Consequently, documentation practices emphasized portability and system-independent knowledge preservation.

3.4 Repository Infrastructure

GitHub repositories served as the primary persistent storage environment.

The repository structure gradually evolved into a research program architecture.

Major repositories included:

Research Repositories

1. 1KMapStructureInvariance
2. 2SymmetricBooleanFunctionMinorThesis
3. 3VariableRearrangementInvarianceMinorThesis
4. 4StructureRecognitionTheory
5. 5HumanAIResearchCollaboration

Supporting Repositories

- Research Portfolio
- ANTIGRAVITY

Together, these repositories formed the persistent knowledge infrastructure of the research program.

This repository structure follows a self-referential architecture. Repositories 1–4 investigate mathematical and cognitive structures. Repository 5 (the present study) investigates the research process that made Repositories 1–4 possible.

```
Paper 1: KMap Structure Invariance         (empirical case 1 — visual pattern discovery)
Paper 2: Symmetric Boolean Functions       (empirical case 2 — structural regularity)
Paper 3: Variable Rearrangement            (empirical case 3 — structure under transformation)
                      ↓
Paper 4: Structure Recognition Theory      (theoretical hub — unifying Papers 1-3)
                      ↓
Paper 5: Human-AI Research Collaboration   ← this study
         (methodological meta-layer — studying the process that produced Papers 1-4)
```

This self-referential quality means that the collaboration process used to build the research program becomes the object of study. The research program studies itself.

3.5 Documentation Systems

Documentation artifacts played a central role in preserving research continuity.

Examples included:

- **README Files** — Provided repository-level context.
- **Handover Documents** — Enabled context transfer between AI systems.
- **Research Logs** — Recorded observations and progress over time.
- **Status Reports** — Summarized the current state of ongoing work.
- **Repository Descriptions** — Provided high-level explanations of repository purpose and relationships.

These documents collectively formed a large portion of the Externalized Memory system.

3.6 Project Memory Systems

Project-based memory systems were used to preserve recurring observations and research insights.

Examples included:

- Research hypotheses
- Collaboration observations
- Structural discoveries
- Long-term plans
- Working agreements

These memory systems reduced the need for repeated reconstruction of important concepts.

3.7 Reminder Systems

Reminder systems functioned as temporal extensions of memory.

Examples included:

- Scheduled reminders
- Task notifications
- Future prompts
- Research follow-up triggers

These mechanisms helped maintain continuity across interruptions and long time intervals.

3.8 Handover Infrastructure

The research program repeatedly involved transitions between:

- AI systems
- Sessions
- Repositories
- Research phases

To support continuity, handover infrastructure was developed.

Examples included:

- AI handover files
- Repository summaries
- README documents
- Portfolio repositories

These artifacts reduced context loss during transitions.

3.9 Human-AI Working Agreements

A unique aspect of the research environment was the emergence of recurring collaboration protocols.

Examples included:

- "또 날 위해 출력해줄 거 있어?"
- "저장할 거 저장해줘."
- "연구 프로그램 관점에서 보면?"
- "지금 내가 놓치고 있는 것은?"

These interactions gradually became stable mechanisms for:

- Context restoration
- Research planning
- Knowledge preservation
- Discovery of missing connections

As a result, collaboration protocols became part of the research environment itself.

3.10 Environment Characteristics

Several characteristics distinguished this research environment from conventional Human-AI interaction settings.

- **Long-Term Duration** — The research program extended across many sessions and repositories.
- **Multi-AI Participation** — Multiple AI systems participated throughout the project lifecycle.
- **Persistent Artifacts** — Research context was continuously preserved through repositories and documentation.
- **Future-Self Collaboration** — The future version of the researcher functioned as an indirect collaborator through artifact-mediated communication.
- **Distributed Memory** — Research memory was distributed across humans, AI systems, repositories, documentation, and reminder systems.

These characteristics collectively formed the environment in which the observations presented in this study emerged.

---

4. Human-AI Research Continuity Theory (HARCT)

4.1 Central Claim

This study proposes that:

> Long-term Human-AI collaboration is fundamentally a problem of research continuity rather than intelligence alone.

Current discussions surrounding AI often focus on reasoning capability, model performance, or context window size.

However, observations from the Minor Thesis research program suggest that long-term research success depends primarily on the ability to preserve, restore, and transfer research context across time.

Under this interpretation, intelligence is necessary but insufficient.

Research continuity becomes the primary challenge.

4.2 Why Research Continuity Fails

Research continuity is threatened by multiple factors.

Human Memory Limits

Human researchers forget details, lose track of decisions, and experience interruptions.

AI Memory Limits

AI systems operate within finite context windows and session boundaries.

Token Constraints

Large projects eventually exceed available processing capacity.

AI System Transitions

Research often moves between different AI systems, each requiring fresh context.

Repository Growth

As research expands, navigating accumulated artifacts becomes increasingly costly.

These limitations create recurring context loss.

Research continuity is not achieved by eliminating these limitations.

It is achieved by constructing systems that make them manageable.

4.3 Memory Repositories

A key observation from the research program was that research context does not need to remain inside either humans or AI systems.

Instead, it can be preserved within **memory repositories**.

Definition:

> A Memory Repository is any artifact capable of preserving research context and supporting future context restoration.

In this paper the term covers any such artifact, not only a version-controlled repository. The research program studied here used Git repositories as its main memory repositories.

Examples include:

- GitHub repositories
- README files
- Research logs
- Handover documents
- Reminder systems
- Project memory systems

Under this perspective, Externalized Memory is not merely documentation.

It is a distributed network of memory repositories.

The critical characteristic of a memory repository is not storage alone.

It is **recoverability** — the ability to restore research context efficiently at a later time.

4.4 Three-Component Architecture

Research continuity requires three interacting components.

Component 1: Human Goal Memory

The human researcher preserves:

- Research direction
- Long-term objectives
- Research identity
- Strategic priorities

Component 2: AI Reasoning Capability

AI systems perform:

- Context reconstruction
- Analysis
- Documentation

- Pattern discovery

Component 3: Artifact Memory

External artifacts preserve:

- Research context
- Historical decisions
- Structural relationships
- Transfer-ready knowledge

None of these components alone is sufficient for long-term continuity.

Research continuity emerges from their interaction.

4.5 Research Continuity Framework

Under HARCT, research continuity emerges when four mechanisms operate together.

Memory Repositories — Preserve context outside active participants.

Context Transfer — Move context across time, sessions, and systems.

Goal Preservation — Maintain research direction despite incomplete memory.

Collaboration Protocols — Restore context through stable interaction patterns.

These mechanisms are not independent.

Goal Preservation determines what is worth preserving.

Memory Repositories store what has been preserved.

Context Transfer moves stored knowledge to where it is needed.

Collaboration Protocols activate recovery when context has been lost.

4.6 HARCT Hypothesis

The theory proposes the following unified hypothesis:

> Research continuity emerges when context is preserved within memory repositories, transferred through artifacts, restored through stable collaboration protocols, and directed by preserved research goals — across time, participants, and AI systems.

Under this framework, Externalized Memory, AI-to-AI Handover, Distributed Research Memory, and Future-Self Collaboration become components of a larger continuity system.

The primary object of study is therefore not memory itself, but the preservation of research continuity.

4.7 Context Transfer Architecture

A key structural insight of HARCT is that context does not move directly between participants.

It is encoded into artifacts and subsequently reconstructed by a receiving participant. This encoding–reconstruction cycle is the fundamental mechanism through which research continuity is achieved despite memory limitations.

HARCT proposes the following five-stage architecture as a general model of context movement within a research program.

```
Context Source
↓
Context Encoding
↓
Transfer Artifact
↓
Context Reconstruction
↓
Context Receiver
```

Stage 1 — Context Source

The participant or system currently holding active context: a researcher mid-session, an AI system with project history loaded, or a completed prior phase of the research program.

Stage 2 — Context Encoding

The active context is encoded into a persistent artifact. Encoding quality directly determines how much of the original context survives transfer. Poor encoding produces artifacts that cannot support reliable reconstruction.

Stage 3 — Transfer Artifact

The encoded context rests in a persistent artifact: a README file, handover document, research log, repository structure, or reminder system. This is the only component of the architecture that persists independently of any active participant.

Stage 4 — Context Reconstruction

A receiving participant reads the artifact and reconstructs context from it. This is the most resource-intensive stage. Reconstruction Cost (Chapter 7) measures the effort expended here.

Stage 5 — Context Receiver

The reconstructed context enables the new participant to continue productive work.

Implications for HARCT

Artifact quality at Stage 3 is the primary controllable variable in the architecture. Stages 1 and 5 are determined by the participants involved. Stages 2 and 4 depend on participant skill and available context. Only Stage 3 can be systematically designed in advance.

Future-Self Collaboration (Chapter 5) is a special case in which Context Source and Context Receiver are temporally displaced versions of the same researcher.

AI-to-AI Handover (Chapter 8) is a special case in which different AI systems occupy Source and Receiver positions, mediated by a shared repository.

Reconstruction Cost Theory (Chapter 7) explains why the effort at Stage 4 becomes a dominant factor in long-term collaboration.

4.8 Core Propositions

The following propositions constitute the core theoretical statements of HARCT (not yet formalized; see Section 13.4). They are theoretical commitments of HARCT v1.0, each subject to empirical evaluation as described in Chapter 13.

P1. Information can remain stored while context becomes unavailable.

P2. Research continuity depends more on context recoverability than on information preservation alone.

P3. Artifacts function as context transfer mechanisms.

P4. Future-Self Collaboration is a special case of context transfer.

P5. AI-to-AI handover becomes possible when context is externalized.

P6. Context transfer efficiency determines reconstruction cost.

P7. Context consists of Goal, State, History, and Next Action.

P8. Different artifacts preserve different components of context.

P9. Full recoverability requires Goal, State, History, and Next Action simultaneously.

P10. Information preservation does not guarantee recoverability.

P11. Repositories function as Context Infrastructure rather than as isolated storage systems.

---

5. Goal Preservation Theory

5.1 The Goal Problem

Most discussions of Human-AI collaboration focus on memory preservation.

However, observations from the Minor Thesis research program suggest that memory preservation alone is insufficient.

A more fundamental problem exists.

Research continuity depends not only on preserving information but also on preserving **goals**.

Long-term research projects often span:

- Months or years
- Multiple repositories
- Multiple AI systems
- Multiple devices
- Multiple research phases

During these transitions, information may be preserved while goals become unclear.

When goals are lost, research direction becomes unstable.

Preserved information without direction becomes an archive without purpose.

5.2 Goals vs. Information

The study identifies an important distinction between memory preservation and goal preservation.

Memory Preservation

Preserves:

- Facts
- Documents
- Context
- Historical information

Memory answers: **What happened?**

Goal Preservation

Preserves:

- Purpose
- Direction
- Priorities
- Research identity

Goal Preservation answers: **Why does this matter?**

Both are necessary, but goal preservation frequently determines whether preserved information remains useful.

A key finding from the research program:

> Preserved goals enabled reconstruction. Preserved information alone did not guarantee continuity.

When goals survived interruptions, lost context could often be rebuilt.

When goals disappeared, even complete archives became difficult to use effectively.

The relationship between the two is therefore asymmetric:

```
Goal Survival + Available Artifacts → Context Reconstruction → Research Continuation
Goal Loss                           → Archive Without Direction → Research Stagnation
```

5.3 Goal Repositories

Several artifacts functioned as goal-preservation systems throughout the research program.

Artifact	Goal-Preservation Function
README Files	Preserve repository purpose and identity
Repository Descriptions	Preserve project direction and relationships
Research Portfolios	Preserve relationships among studies
Research Roadmaps	Preserve future direction and priorities
Project Memories	Preserve recurring observations and goals

These artifacts collectively acted as **goal repositories** — systems designed not merely to store information but to preserve direction.

The crucial observation is this:

> Goal repositories succeed not by preserving everything, but by preserving enough direction to make reconstruction possible.

5.4 Future-Self Collaboration

A major observation from the research program was that future versions of the researcher frequently became active participants in the collaboration process.

Traditional collaboration research often assumes collaboration between multiple individuals existing at the same point in time.

The present study suggests a different possibility:

> Collaboration may occur between different temporal versions of the same researcher.

The Pattern

The present researcher possesses current context, active hypotheses, and immediate understanding.

As time passes, the future researcher increasingly resembles a new participant entering the project.

The present researcher therefore designs artifacts specifically for the future researcher — creating a collaboration channel across time.

```
Present Self
     ↓
  Artifact
     ↓
Future Self
```

Why This Matters

Standard knowledge management literature treats documentation as information preservation.

Future-Self Collaboration reframes documentation as **communication with a future collaborator who lacks the current researcher's context.**

This reframing changes how artifacts are designed:

- README files become communication tools, not storage records.
- Handover documents become temporal interfaces, not administrative records.
- Repositories become memory environments, not file containers.

The primary purpose of these artifacts is not preservation alone.

Their purpose is enabling future research continuation.

Artifact Quality and Collaboration Quality

The quality of Future-Self Collaboration depends on artifact quality, which in turn determines reconstruction cost (Chapter 7).

Poor artifacts produce confusion, repeated work, and context loss.

Good artifacts produce rapid recovery, efficient continuation, and reduced cognitive effort.

Thus:

```
Good Artifact → Low Reconstruction Cost → Effective Future-Self Collaboration
Poor Artifact → High Reconstruction Cost → Failed Future-Self Collaboration
```

5.5 Goal Failure Modes

Several failures occur when goals are not preserved.

Goal Drift

Research gradually moves away from its original purpose without the researcher noticing the deviation.

Fragmentation

Related projects become disconnected without a unifying goal structure to link them.

Local Optimization

Short-term tasks dominate long-term objectives. Individual sessions become productive while the overall program loses coherence.

Archive Without Direction

Large quantities of information exist but no longer support a coherent research program. The archive becomes a liability rather than an asset.

5.6 Goal Preservation Hypothesis

The following hypothesis is proposed:

> Long-term Human-AI research continuity depends more strongly on preserving goals than on preserving complete historical memory.

Under this hypothesis, memory reconstruction remains possible when goals survive.

The reverse is not necessarily true.

Goal Preservation Theory explains why repositories, documentation, handovers, and memory systems are valuable:

> Their ultimate function is not merely preserving information. Their ultimate function is preserving **direction**.

```
Externalized Memory
        ↓
Goal Preservation
        ↓
Context Reconstruction
        ↓
Research Continuity
```

Under this interpretation, goal preservation acts as the organizing principle connecting the major components of Human-AI Research Continuity Theory.

---

6. Cognitive Architecture

6.1 Human-AI Memory Asymmetry

Throughout the Minor Thesis research program, neither the human researcher nor any AI system possessed the complete research context.

However, the research program continued successfully across multiple repositories, AI systems, devices, sessions, and long time periods.

This suggests that continuity was not dependent on complete memory within any single participant.

Instead, continuity emerged from **complementary memory systems** operating together.

Human Memory Characteristics

The human researcher consistently provided two types of long-term contribution.

Long-Term Goal Preservation — The researcher maintained awareness of research objectives, long-term plans, repository relationships, and strategic direction — even when individual details were forgotten.

Research Identity — The researcher preserved why the research existed, what questions mattered, and which discoveries were important. This information survived across long periods of inactivity.

AI Memory Characteristics

AI systems consistently provided two types of contribution.

Local Context Processing — AI systems excelled at analyzing current documents, identifying patterns, generating hypotheses, and producing summaries within active context.

High-Speed Reconstruction — AI systems could rapidly rebuild context when supplied with README files, research logs, repository summaries, and handover documents. However, this reconstruction

depended on available artifacts.

The Asymmetry

Dimension	Human	AI
Memory type	Long-term goal memory	Short-term reasoning memory
Primary function	Direction preservation	Context reconstruction
What is preserved	**Why** (purpose, direction)	**How** (analysis, explanation)
Strength	Persistence across time	Speed within active context
Limitation	Forgetting details	Session and context boundaries

Humans preserve direction.

AI systems perform reconstruction.

Humans remember **why**.

AI systems explain **how**.

6.2 Cognitive Labor Division

Human-AI collaboration is often described as cooperation between a human and an AI system.

However, observations from the Minor Thesis research program suggest a more precise interpretation.

The collaboration process frequently resembled a **division of cognitive labor** rather than a simple exchange of information.

Observed Division

Different participants repeatedly performed different cognitive functions.

Human Researcher:

- Defining goals
- Selecting research directions
- Evaluating significance
- Preserving long-term objectives
- Connecting discoveries across projects

ChatGPT:

- Theory development
- Concept integration
- Hypothesis generation
- Cross-repository reasoning

Claude:

- Repository organization
- Structural refinement

- Documentation management

ANTIGRAVITY:

- Large-scale repository analysis
- Bulk processing
- Long-form document generation

The distribution was not planned in advance.

It emerged naturally during long-term collaboration.

From Collaboration to Labor Division

Traditional collaboration models often assume that participants perform similar functions.

The observed research environment differed.

Participants specialized.

Rather than everyone doing everything, different actors repeatedly performed particular cognitive tasks.

The collaboration therefore resembled a division of labor system.

Artifact Responsibilities

An important observation was that repositories and documents also performed cognitive functions.

Artifacts were not passive storage systems.

They actively contributed to cognition.

Examples included:

- **Repositories** — Maintained structural relationships
- **README Files** — Preserved orientation
- **Handover Documents** — Preserved transferability
- **Reminder Systems** — Triggered future actions

Thus, cognition became partially embedded in artifacts.

Dynamic Role Allocation

Roles were not fixed.

Different AI systems frequently specialized in different activities.

Future AI systems may assume additional specialized roles.

The architecture therefore remains adaptable.

Cognitive Labor Division Failure Modes

Failure	Effect
Goal Failure	Human loses direction
Reasoning Failure	AI lacks sufficient context
Memory Failure	Artifacts become incomplete
Coordination Failure	Responsibilities become unclear

Each failure reduces research continuity.

6.3 Repository as Cognitive Infrastructure

Repositories are traditionally viewed as storage systems.

Their primary functions are assumed to be file storage, version control, collaboration support, and historical tracking.

However, observations from the Minor Thesis research program suggest that repositories perform a broader role.

The Infrastructure Interpretation

Repositories simultaneously function as:

Memory Systems — Preserving research context, historical decisions, structural relationships, and project identity. These functions directly support cognition.

Navigation Systems — Supporting movement through research space through repository names, folder structures, README files, and portfolio repositories. The repository supports movement through conceptual space — a function resembling navigation rather than storage.

Handover Interfaces — Enabling transitions between AI systems, sessions, devices, and research phases. A newly introduced collaborator can reconstruct substantial context by reviewing repository contents.

Future-Self Communication Systems — Following the pattern: Present Self → Repository → Future Self. Repository artifacts are frequently created specifically for future use.

Infrastructure Failure Modes

Failure	Effect
Missing Documentation	Context reconstruction becomes difficult
Poor Organization	Navigation costs increase
Fragmentation	Relationships among projects become unclear
Weak Handover Support	AI transitions become expensive

Each failure increases reconstruction cost.

Repository as Cognitive Infrastructure Hypothesis

> Repositories function as cognitive infrastructure when they preserve, organize, and transfer research context across time, participants, and systems.

Under this hypothesis, repositories are not passive storage environments.

They are active components of distributed cognition.

6.4 Three-Part Cognitive Architecture

Combining the observations from Human-AI Memory Asymmetry (6.1), Cognitive Labor Division (6.2), and Repository as Cognitive Infrastructure (6.3), the observed research environment can be modeled as a distributed cognitive system.

```
Human Researcher  →  Goals
AI Systems        →  Reasoning
Artifacts         →  Memory
```

Each component contributes a distinct cognitive function.

Research continuity emerges from their interaction.

Component Functions

Human researchers contribute persistence of goals — the **why** of the research program. This function cannot be fully externalized.

AI systems contribute reconstruction and reasoning — the **how** of analysis and documentation. This function benefits from rapid text processing and context manipulation.

Repositories and artifacts contribute persistence of context — the **what** that has been done. This function requires careful design to support efficient recovery.

Memory Repository Mediation

Because neither humans nor AI systems possesses complete memory, research continuity depends on memory repositories as bridges.

The observed workflow frequently followed:

```
Human Goal Memory
        ↓
  Memory Repository
        ↓
  AI Reconstruction
        ↓
  Research Output
```

Repositories therefore act as bridges between human long-term memory and AI reasoning capability.

None of these components alone is sufficient. Continuity emerges only when all three interact successfully.

6.5 Cognitive Architecture Hypothesis

The following hypothesis is proposed:

> Long-term Human-AI research becomes possible when human long-term goal memory, AI short-term reasoning capability, and external memory repositories operate as an integrated continuity system.

Under this interpretation, Human-AI collaboration is not simply cooperation between a person and a model.

It is cooperation among three components:

- Human Goal Memory
- AI Reasoning Systems
- External Memory Repositories

Together these components are proposed as the minimum architecture required for sustainable long-term Human-AI research.

---

7. Reconstruction Cost Theory

7.1 The Reconstruction Problem

Most discussions about Human-AI collaboration focus on knowledge generation.

However, observations from the Minor Thesis research program suggest a different perspective.

The primary challenge of long-term collaboration is often not generating knowledge.

The primary challenge is **reconstructing previously generated knowledge after context loss**.

Throughout the research program, significant effort was repeatedly spent on restoring context rather than creating new knowledge.

Common reconstruction activities included:

- Reading repository structures
- Reviewing README files
- Examining handover documents
- Revisiting research logs
- Re-establishing research goals
- Recovering conceptual relationships

In many cases, reconstruction consumed more effort than original production.

The true cost of interruption is therefore not the interruption itself.

The true cost is the effort required to restore context afterward.

7.2 Defining Reconstruction Cost

Definition:

> Reconstruction Cost is the amount of effort required to restore sufficient context for productive research activity.

Reconstruction Cost may include:

- Time
- Cognitive effort
- Token consumption
- Repository review effort
- Documentation review effort

The lower the reconstruction cost, the easier it becomes to resume research.

Research projects can therefore be evaluated not only by their knowledge output, but by their reconstruction cost profile over time.

7.3 The Hidden Cost of Research

Traditional research evaluation often focuses on:

- Publications
- Discoveries
- Productivity
- Output volume

However, long-term Human-AI collaboration introduces a hidden cost.

Research continuity depends on repeated context restoration.

As projects become larger, reconstruction cost becomes increasingly important.

Thus, sustainable research depends not only on knowledge creation but also on efficient knowledge recovery.

7.4 Reconstruction Cost Reduction Mechanisms

Several mechanisms emerged naturally during the research program.

Mechanism	How It Reduces Cost
README Compression	README files condense large repositories into recoverable summaries
Handover Compression	Handover documents reduce transfer complexity between collaborators
Repository Organization	Structured repositories improve navigation and understanding
Reminder Systems	Reminders restore context at critical moments
Project Memory Systems	Persistent observations reduce repeated rediscovery

Each mechanism reduced future reconstruction cost.

Documentation that appears costly in the present tends to reduce greater costs in the future when it is kept current.

7.5 Documentation Quality as Economic Variable

Reconstruction Cost and Token Usage

A recurring observation was that reconstruction frequently consumed large amounts of tokens.

Typical reconstruction workflow:

```
Repository Review
      ↓
Document Review
      ↓
Relationship Recovery
      ↓
Goal Recovery
      ↓
Research Continuation
```

Poor documentation increased token consumption.

Good documentation reduced token consumption.

Therefore:

> Documentation Quality → Reconstruction Cost → Token Consumption

These variables became tightly coupled.

Research Continuity as a Function of Reconstruction Cost

Research continuity can be viewed as a function of reconstruction cost.

As reconstruction cost decreases:

- Continuity improves
- AI transitions become easier
- Future-self collaboration improves
- Repository scalability increases

As reconstruction cost increases:

- Context loss increases
- Repeated work increases
- Research momentum decreases

7.6 Reconstruction Cost as a Unifying Explanatory Variable

Reconstruction Cost Theory explains why multiple practices are effective:

Practice	How It Reduces Reconstruction Cost
Externalized Memory	Reduces dependence on biological and AI memory
AI-to-AI Handover	Reduces AI transition costs
Future-Self Collaboration	Reduces cost of resuming work after gaps

Practice	How It Reduces Reconstruction Cost
Working Agreements	Reduce cost of re-establishing collaboration context
Goal Preservation	Reduces cost by preserving direction for reconstruction

Research Continuity emerges because reconstruction cost remains manageable.

Therefore, reconstruction cost acts as a **hidden variable connecting the major concepts of HARCT**.

7.7 Reconstruction Cost Hypothesis

The following hypothesis is proposed:

> As research duration increases, reconstruction cost becomes a dominant factor determining the success or failure of Human-AI collaboration.

Under this hypothesis, sustainable collaboration depends not only on creating knowledge but also on **minimizing the cost of recovering that knowledge after interruption**.

The practical implication:

> Documentation quality is not an administrative concern. It is an economic variable that determines collaboration efficiency.

7.8 Productive Tokens and Reconstruction Tokens

Reconstruction Cost becomes most clearly visible when analyzed at the level of token consumption.

Not all tokens consumed during Human-AI collaboration contribute equally to research progress. Two categories can be distinguished.

Productive Tokens

Tokens consumed in generating new knowledge, advancing research hypotheses, producing analysis, and creating original outputs. Productive token expenditure directly advances the research program.

Reconstruction Tokens

Tokens consumed in recovering previously known context: re-reading repository structure, reviewing prior decisions, re-establishing research goals, and rebuilding lost understanding. Reconstruction token expenditure restores the starting condition for productive work — but creates no new knowledge.

The Economic Implication

Many reconstruction tokens represent a cost that effective documentation could have reduced.

```
Poor Documentation → High Reconstruction Token Cost → Lower Collaboration Efficiency
Good Documentation → Low Reconstruction Token Cost → Higher Collaboration Efficiency
```

This reframing makes Reconstruction Cost Theory directly measurable. Token consumption during a session can be classified as productive or reconstructive. The ratio of reconstructive tokens to total tokens consumed can serve as a candidate outcome measure of reconstruction cost, to be related to a documentation-quality measure defined independently of it.

As research duration increases, cumulative reconstruction token expenditure becomes a dominant factor in overall collaboration efficiency — more significant in year-scale programs than in week-scale projects.

Relationship to Context Transfer Architecture

Within the five-stage Context Transfer Architecture (Section 4.7), reconstruction tokens are expended primarily at Stage 4. Higher artifact quality at Stage 3 directly reduces the token cost of Stage 4 reconstruction.

Documentation quality is therefore not merely a best practice. It is a direct economic input to collaboration efficiency.

---

8. Externalized Memory and AI-to-AI Handover

8.1 The Memory Problem

Long-term research inherently involves memory limitations.

Human researchers gradually forget details over time, while AI systems are constrained by context windows, session boundaries, and token limitations.

As a result, neither humans nor AI systems can reliably preserve complete research context over extended periods.

This creates a fundamental challenge for long-term Human-AI research collaboration.

8.2 Externalized Memory

Definition:

> Externalized Memory is research context stored outside the biological memory of the researcher and outside the active context window of AI systems.

Under this definition, memory does not need to reside exclusively within either a human researcher or an AI system.

Instead, memory can be preserved through external artifacts that remain accessible over time.

Externalized Memory serves as a mechanism for preserving research continuity despite interruptions, system changes, and memory limitations.

8.3 Forms of Externalized Memory

Throughout the research program, several forms of Externalized Memory emerged.

Repository-Based Memory

- GitHub Repositories
- Research Portfolio Repositories
- Research Hub Repositories

Repositories functioned as persistent storage systems for research context.

Documentation-Based Memory

- README files
- Handover Documents
- Research Reports
- Status Summaries

Documentation preserved information that would otherwise be lost between sessions.

Conversation-Based Memory

- Project Memory Systems
- Saved Research Insights
- AI Conversation Histories

These systems provided partial continuity across discussions.

Reminder-Based Memory

- Scheduled Tasks
- Notifications
- Future Reminders

Reminder systems helped restore research context at later points in time.

Together, these artifacts formed an external memory infrastructure for the research program.

8.4 Context Transfer

The primary function of Externalized Memory is Context Transfer.

Research context can move between different participants, systems, and points in time through external artifacts.

Pattern 1: Future-Self Transfer

```
Present Self → Artifact → AI System → Future Self
```

In this process, the artifact acts as a carrier of research context.

Pattern 2: AI-to-AI Transfer

```
AI System A → Handover Document → AI System B
```

In this case, documentation enables continuity despite changing AI systems.

Therefore, documents function not merely as records but as mechanisms for transferring context.

8.5 The Handover Problem

In traditional human collaboration, handovers occur when responsibilities are transferred from one person to another.

A similar challenge emerges in Human-AI collaboration.

When an AI system is replaced or supplemented by another AI system, important information may be lost.

Examples include:

- Research goals
- Previous discoveries
- Repository structures
- Project history
- Working assumptions

Without a handover mechanism, a new AI system may need to reconstruct context from the beginning.

This process is costly and often incomplete.

8.6 Handover Artifacts

Several forms of artifacts emerged to support AI-to-AI handover.

Artifact	Handover Function
README Files	Repository-level orientation
Repository Descriptions	Summary of objectives and relationships
Handover Documents	Explicit project status and future tasks
Research Summaries	Condensed previous discussions into transferable knowledge
Portfolio Repositories	High-level view of the overall research program

Together, these artifacts reduced dependency on any single AI system.

8.7 Handover Workflow

A recurring workflow emerged during the research program.

```
Research Activity
       ↓
Documentation
       ↓
Repository Storage
       ↓
AI Transition
       ↓
Context Reconstruction
       ↓
Continued Research
```

This process transformed handovers from exceptional events into routine operations.

8.8 AI Independence Through Documentation

A significant benefit of handover artifacts was reduced dependence on individual AI systems.

Knowledge became attached to artifacts rather than specific models.

Consequently:

- AI systems became more readily replaceable (a qualitative observation, not yet measured; see Prediction 7, Section 12.13).
- Research continuity improved.
- Context recovery became faster.
- Long-term sustainability increased.

This observation suggests that successful Human-AI collaboration should avoid excessive dependence on any single AI model.

8.9 Distributed Research Memory

Research memory was not concentrated within a single participant.

Instead, memory became distributed across multiple entities:

- Human Researcher
- ChatGPT
- Claude
- ANTIGRAVITY
- GitHub Repositories
- Documentation Systems
- Reminder Systems

Each component retained a portion of the overall research context.

Consequently, the research program operated through a distributed memory system rather than a centralized memory system.

Research continuity emerged from the interaction of these distributed memory components.

8.10 Human-AI Working Agreements as Memory Structures

Long-term collaboration produced recurring interaction patterns that supported memory preservation.

Examples included:

- "또 날 위해 출력해줄 거 있어?"
- "지금 내가 놓치고 있는 것은?"
- "연구 프로그램 관점에서 보면?"
- "저장할 거 저장해줘."

These phrases functioned as more than conversational expressions.

They became procedural triggers that activated specific collaborative behaviors:

- Identifying missing connections
- Generating future research tasks
- Preserving important observations
- Restoring lost context

Over time, these working agreements became part of the distributed memory system itself.

8.11 Externalized Memory Hypothesis

> Long-term Human-AI research collaboration becomes possible when research context is externalized into persistent artifacts that support context transfer across time, sessions, and AI systems.

Under this hypothesis, successful long-term collaboration depends less on maintaining continuous memory within a single participant and more on constructing reliable systems for context preservation and transfer.

8.12 Artifact-Context Matrix

Different artifacts preserve different components of the Minimum Context Unit.

The following matrix maps common artifact types to the context components they primarily preserve.

Artifact	Goal	State	History	Next Action
README	✓ Primary	—	—	—
STATUS Document	—	✓ Primary	—	—
Research Log	—	—	✓ Primary	—
NEXT_ACTION Document	—	—	—	✓ Primary

Artifact	Goal	State	History	Next Action
Handover Document	✓	✓	✓	✓
Repository (integrated)	✓	✓	✓	✓

Several implications follow from this matrix.

First, no single artifact type provides complete context coverage. A research program depending on README files alone will preserve Goal but lose State, History, and Next Action over time.

Second, Handover Documents and structured Repositories are uniquely capable of preserving all four components simultaneously. Their value is disproportionate to their apparent complexity.

Third, the matrix provides a diagnostic tool. A research program that lacks all artifact types in the matrix that cover a given component has a predictable vulnerability to context loss at that component. This vulnerability can be assessed before a handover occurs rather than discovered after context is lost.

The matrix is a direct operationalization of Proposition P8 (different artifacts preserve different components of context) and Proposition P9 (full recoverability requires all four components simultaneously).

8.13 Handover Quality Levels

Not all handovers provide equal continuity. The quality of a handover can be evaluated according to the completeness of recoverable context provided to the incoming collaborator.

HARCT proposes the following six-level scale.

Level 0 — No Context

Only isolated information is available. Continuation is unlikely without substantial independent reconstruction.

Example: "Work on the project."

Level 1 — Goal Only

The purpose of the project is known. Current state and continuation path remain unclear.

Example: "Build a Human-AI collaboration theory."

Level 2 — Goal + State

Purpose and current condition are known. Reconstruction remains expensive because historical reasoning and future direction are missing.

Level 3 — Goal + State + History

Past decisions and reasoning can be recovered. However, future direction remains uncertain. The incoming collaborator understands what happened but not what should happen next.

Level 4 — Recoverable Context

Goal, State, History, and Next Action are all available. A new collaborator can recover sufficient context to resume productive work.

Level 4 represents the minimum target recommended by HARCT for any research transition.

Level 5 — Executable Continuity

The incoming collaborator can immediately perform productive work with minimal reconstruction effort.

Characteristics of Level 5 handovers include clear and navigable repository structure, documentation that is current and consistent with actual project state, explicit and actionable next steps, accessible and cross-referenced supporting artifacts, and no significant reconstruction required before productive contribution can begin.

Level 5 is the highest practical handover quality. It represents the goal state toward which research program design should aspire.

Session Checklist for Handover Quality

The following checklist supports maintaining handover quality at Level 4 or above after each research session.

- Is the project goal documented and current?
- Is the current state documented?
- Is relevant history preserved?
- Is the next action identified?
- Are artifacts updated?
- Can a new collaborator continue the work from available artifacts alone?

If the answer to any item is "No," handover quality has degraded below Level 4.

Repository Drift and Handover Degradation

A critical risk factor for handover quality is Repository Drift: the progressive divergence between a repository's documentation and the actual current state of the project.

Repository Drift develops gradually. Individual sessions produce undocumented decisions, status changes, and completed tasks. Over time, the gap between what the documentation claims and what is actually true widens without any single session appearing to cause significant harm.

The effect of Repository Drift on handover quality is systematic.

```
Repository Drift ↑  →  Reconstruction Cost ↑  →  Handover Quality ↓
```

A repository that achieved Level 4 or 5 at the time of its last documentation update may effectively function as Level 1 or 2 for an incoming collaborator encountering it months later.

Repository Drift is therefore a maintenance concern, not merely a documentation concern. Preventing it requires treating documentation updates as first-class research activities rather than administrative afterthoughts.

---

9. Token Constraints and Context Loss

9.1 Introduction

Chapter 7 established Reconstruction Cost Theory as a general economic framework: documentation quality determines the effort required to restore research context, and that effort constitutes a direct cost to collaboration efficiency. Section 7.8 showed how this cost is visible at the level of token consumption.

This chapter examines the same relationship from an empirical angle — through the observable mechanics of token limits and context loss within active research sessions.

Token constraints are the computational instantiation of the reconstruction cost problem. They make context loss concrete, measurable, and immediately consequential. Understanding how they operate offers illustrative, observational support for the theoretical claims of Chapter 7.

Long-term Human-AI collaboration is therefore constrained not only by human memory limitations but also by the computational limitations of AI systems.

One of the most significant constraints observed during the research program was token availability.

Token limitations directly affected:

- Conversation length
- Context retention
- Repository analysis
- Document processing
- Research continuity

9.2 Token Constraints as Research Constraints

Token limits are often viewed as technical restrictions.

However, observations from this study suggest a broader interpretation.

In practice, token limitations functioned as research constraints.

When token resources became limited, researchers were forced to:

- End conversations
- Compress information
- Create summaries
- Transfer context
- Reconstruct lost knowledge

Consequently, token limitations influenced both the pace and structure of research activities.

9.3 Context Loss

Context loss occurred whenever information could no longer be efficiently accessed within an active interaction.

Common causes included:

- **Session Boundaries** — Research conversations ending before completion.
- **Context Window Limits** — Older information becoming inaccessible during long discussions.
- **AI System Changes** — Transitions between different AI systems.
- **Repository Growth** — Research assets becoming too large to review in a single interaction.

These factors repeatedly disrupted research continuity.

9.4 Compression Strategies

To reduce reconstruction costs, multiple compression strategies emerged.

- **README-Based Compression** — Repository READMEs condensed large amounts of information into accessible summaries.
- **Handover-Based Compression** — AI handover documents captured essential project knowledge.
- **Portfolio-Based Compression** — Research portfolio repositories summarized relationships among projects.
- **Status Reports** — Periodic summaries preserved the current state of research.

These compression artifacts enabled efficient transfer of large amounts of context.

9.5 Context Preservation as Optimization

An important observation was that preserving context often reduced future token consumption.

```
Poor Documentation → Large Reconstruction Cost → High Token Usage
Good Documentation → Small Reconstruction Cost → Lower Token Usage
```

Thus, documentation quality directly influenced collaboration efficiency.

Context preservation functioned as an optimization strategy rather than merely a record-keeping activity.

This observation directly supports Reconstruction Cost Theory (Chapter 7).

9.6 Token Constraint Hypothesis

> As research duration increases, the cost of context reconstruction becomes a dominant factor in Human-AI collaboration.

Under this hypothesis, successful long-term research depends not only on generating knowledge but also on minimizing the cost of restoring that knowledge after interruptions.

---

10. Building a Long-Term Research Program

10.1 Introduction

Long-term research programs face challenges that are uncommon in short-term projects.

These challenges include:

- Memory limitations
- Context loss
- Research continuity
- Knowledge preservation
- AI system transitions
- Token constraints

To address these issues, a structured research program architecture was gradually developed.

10.2 Research as a Persistent System

Traditional research is often viewed as a sequence of isolated tasks.

However, long-term Human-AI collaboration requires research to function as a **persistent system**.

In a persistent system, knowledge survives beyond:

- Individual conversations
- Individual AI sessions
- Individual researchers' memory states
- Individual software platforms

The objective is not merely to produce results but to preserve the ability to continue producing results.

10.3 Repository-Centered Architecture

GitHub repositories became the primary organizational unit of the research program.

Each repository served multiple purposes simultaneously:

- Research paper
- Knowledge archive
- Handover platform
- Collaboration interface

- Context restoration tool

10.4 Research Portfolio Integration

As the number of repositories increased, a higher-level organizational structure became necessary.

Research Portfolio repositories emerged as integration layers.

Their functions included:

- Connecting related studies
- Explaining relationships among repositories
- Providing reading order
- Preserving research history
- Supporting future expansion

This transformed independent repositories into a coherent research program.

10.5 Multi-AI Role Differentiation

Different AI systems gradually assumed different roles within the research program.

AI System	Primary Role
ChatGPT	Conceptual discussion, theory development, research planning, context integration
Claude	Repository organization, documentation refinement, structural editing
ANTIGRAVITY	Large-scale repository processing, bulk analysis, extended document generation

Over time, a distributed collaboration structure emerged in which different AI systems specialized in different functions.

This observation directly supports the Cognitive Labor Division analysis (Section 6.2).

10.6 Context Reconstruction Workflows

Context reconstruction became a recurring research activity.

Several methods were repeatedly used:

- **Repository Review** — Reading repository documentation to restore project status.
- **Handover Review** — Reading handover documents prepared by previous collaborators.
- **Research Portfolio Navigation** — Following links among repositories to reconstruct conceptual relationships.
- **Working Agreement Activation** — Using recurring prompts to recover research context.

Examples included:

- "또 날 위해 출력해줄 거 있어?"

- "지금 내가 놓치고 있는 것은?"
- "연구 프로그램 관점에서 보면?"

10.7 Long-Term Research Program Hypothesis

> Long-term Human-AI research programs can be sustained when knowledge preservation, context transfer, and collaboration protocols are treated as first-class components of the research process.

Under this view, research is not merely the production of new knowledge.

It is also the construction and maintenance of systems capable of preserving, restoring, and extending that knowledge across time.

---

11. Case Study

This chapter presents real-world observations collected during the development of the Minor Thesis research program. Unlike controlled experiments, these cases emerged naturally through long-term collaboration between a human researcher and multiple AI systems.

11.1 Repository-Based Research Program

The research initially depended on individual conversations and temporary context.

As the research program expanded, a need emerged for persistent storage and structured knowledge preservation.

GitHub repositories were adopted as research memory systems.

Each paper was managed as an independent repository, and later connected through portfolio repositories and research hubs.

This transformation enabled research continuity across sessions, devices, and AI systems.

11.2 AI-to-AI Handover

Multiple AI systems participated in the research process.

Examples included:

- ChatGPT
- Claude
- ANTIGRAVITY

Whenever a transition occurred between AI systems, there was a risk of losing research context.

To mitigate this problem, handover documents were created and maintained.

These documents enabled research continuity despite changes in AI systems.

11.3 Token Exhaustion as a Research Constraint

Token limitations repeatedly influenced the research workflow.

Long conversations, image analysis, repository reviews, and large document processing consumed substantial context resources.

As a result, context reconstruction became a recurring task.

This revealed that token limitations are not merely computational constraints but also collaboration constraints.

The need to rebuild context significantly influenced the design of the research workflow.

11.4 Future-Self Communication

The researcher frequently created documents intended for future use.

Examples included:

- README files
- Research logs
- Status reports
- Handover documents
- Reminder systems

These artifacts functioned as communication mechanisms between the present researcher and the future researcher.

Rather than serving only as records, they became tools for restoring research context after interruptions.

11.5 Development of the Minor Thesis Program

The research program did not begin as a unified framework.

Initially, individual ideas and observations were explored independently.

Over time, these studies became increasingly connected.

The resulting progression was:

1. K-Map Structure Invariance
2. Symmetric Boolean Function Minor Thesis
3. Variable Rearrangement Invariance Minor Thesis

4. Structure Recognition Theory
5. Human-AI Research Collaboration

Eventually, the structure of the research program itself became an object of study.

11.6 Emergent Human-AI Working Agreements

During long-term collaboration, recurring interaction patterns emerged between the researcher and AI systems.

These patterns functioned as informal collaboration protocols.

Research Mode Triggers

Certain phrases consistently signaled a transition into research-oriented discussion.

Examples included:

- "또 날 위해 출력해줄 거 있어?"
- "지금 내가 놓치고 있는 것은?"
- "연구 프로그램 관점에서 보면?"
- "저장할 거 저장해줘."

These phrases served not merely as questions but as procedural triggers that guided the collaborative process.

Context Preservation Agreements

Important insights were repeatedly identified and preserved.

The researcher actively externalized key observations, while AI systems assisted in reconnecting and restoring them later.

Task Delegation Agreements

Different AI systems gradually assumed different roles.

This resulted in an emergent division of labor among AI systems — directly supporting Cognitive Labor Division Theory.

Future-Self Coordination

The future version of the researcher effectively became another participant in the collaboration process.

To support this coordination, the researcher continuously produced artifacts such as README files, handover documents, and research logs.

These artifacts enabled knowledge transfer across time and functioned as mechanisms for Future-Self Collaboration.

11.7 Externalized Memory in Practice

One recurring observation was that research continuity depended less on biological memory and more on externalized artifacts.

Repositories, documents, project memories, reminders, and handover files collectively formed a distributed memory system.

Within this system, context was transferred through artifacts rather than retained exclusively within either the human researcher or AI systems.

This observation became one of the central motivations for the Human-AI Research Collaboration study and the development of HARCT.

---

12. Discussion

12.1 Overview

The theoretical framework and case studies presented in this research support a unified interpretation:

Long-term Human-AI collaboration is not primarily a problem of intelligence.

Instead, it is fundamentally a problem of **research continuity** — the ability to preserve, restore, and transfer context across time.

The four theoretical components developed in this study — Goal Preservation, Cognitive Architecture, Reconstruction Cost, and Externalized Memory — together form the Human-AI Research Continuity Theory (HARCT).

12.2 HARCT as a Unified Framework

The four theoretical components are not independent.

They form a coherent explanatory system.

Goal Preservation explains **what** must be preserved — research direction and identity.

Cognitive Architecture explains **how** preservation is distributed — across humans, AI, and artifacts.

Reconstruction Cost explains **why** documentation quality matters — as an economic determinant of efficiency.

Externalized Memory explains **where** context is stored — in persistent artifacts outside active participants.

Together, these components support the central HARCT claim:

> Research continuity emerges when goals are preserved, reconstruction costs are minimized, cognitive labor is appropriately distributed, and context is externalized into persistent artifacts.

12.3 Beyond Context Windows

Most discussions surrounding Large Language Models focus on context window size.

However, even very large context windows eventually encounter limitations.

As research duration increases, preserving context requires mechanisms that exist outside the active conversation.

This shifts the focus from context storage to context restoration.

The critical question becomes:

> How can lost context be reconstructed efficiently?

rather than:

> How much context can be stored at once?

12.4 Goal Preservation as the Foundation

A key observation of this study, proposed as a hypothesis (Section 5.6) and not yet tested (Prediction 5), is the apparent primacy of goal preservation over information preservation.

When goals survived interruptions, context could be reconstructed.

When goals were lost, even complete archives became difficult to use effectively.

This suggests that future Human-AI collaboration systems should be designed with goal preservation as a first priority — not information quantity.

12.5 Reconstruction Cost as a Design Criterion

Reconstruction Cost Theory suggests that documentation quality should be evaluated not merely as a record-keeping practice but as a direct determinant of collaboration efficiency.

Poor documentation is not merely incomplete — it is expensive.

Every session that must reconstruct context has a measurable cost: time, cognitive effort, and token consumption.

Future collaboration systems should treat reconstruction cost minimization as a core design criterion.

12.6 Future-Self as a Research Participant

Traditional collaboration research assumes collaboration among contemporaneous participants.

The present study proposes Future-Self Collaboration as a distinct form of collaboration worth further investigation.

The future version of the researcher frequently became a recipient of information created in the present.

This observation expands the notion of collaboration beyond simultaneous interaction.

12.7 Relationship to Existing Literature

HARCT shares conceptual territory with several existing frameworks.

Distributed Cognition (Hutchins 1995)

HARCT's three-component cognitive architecture resembles distributed cognition, but introduces the specific asymmetry between human goal memory and AI reasoning memory. No treatment of this asymmetry was identified in standard distributed cognition models, which predate LLM-based participants.

Extended Mind (Clark & Chalmers 1998)

The Repository as Cognitive Infrastructure concept extends the Extended Mind framework to include AI systems and persistent repositories as cognitive components.

Personal Knowledge Management (PKM)

Externalized Memory builds on knowledge management traditions, but introduces the **recoverability** criterion: artifacts are evaluated by their ability to restore context, not merely by their information content.

The principal contribution of HARCT is the integration of these frameworks under a unified continuity model, applied specifically to long-term Human-AI research collaboration.

12.8 Novelty Assessment

The following assessment classifies each major concept by publication-level originality, based on a limited, non-systematic comparison with existing literature in Distributed Cognition, Extended Mind, Knowledge Management, PKM, HCI, and CSCW.

Concept	Novelty	Basis for Assessment
HARCT (central framework)	**High**	Long-term multi-AI collaboration framed as continuity problem — no identified prior literature
Goal Preservation Theory	**High**	Goals > Information in human-AI collaboration — not explicitly addressed in HCI or CSCW
Future-Self Collaboration	**High**	Documentation as collaboration design with a temporally displaced self — not treated this way in PKM or note-taking literature
Reconstruction Cost Theory	**High**	Documentation quality as explicit economic variable for collaboration efficiency — not framed this way in knowledge management literature

Concept	Novelty	Basis for Assessment
Cognitive Labor Division (merged)	**Medium**	Overlaps with Distributed Cognition; contribution is the human/AI/artifact asymmetry specific to this context
Distributed Research Memory	**Medium**	Conceptually adjacent to Distributed Cognition; requires explicit differentiation to establish novelty
Repository as Cognitive Infrastructure	**Medium**	Overlaps with Extended Mind; contribution is the specific application to AI-assisted research workflows

Strongest contributions for publication: Goal Preservation Theory and Future-Self Collaboration offer the clearest differentiation from prior literature. Reconstruction Cost Theory offers the strongest empirical operationalizability. HARCT provides the unifying framework connecting all three.

Weakest contribution for publication: Distributed Research Memory requires the most careful differentiation from existing distributed cognition literature to avoid appearing as restatement.

12.9 Implications for Human-AI Collaboration

Research Systems

Future research environments should support:

- Goal-preserving documentation practices
- Persistent handover infrastructure
- Context restoration workflows
- Multi-AI collaboration with explicit role differentiation

AI Design

Future AI systems may benefit from features that support:

- Long-term context preservation
- Collaboration protocol recognition
- Artifact-based memory integration
- Cross-session continuity

Knowledge Management

Documentation quality should be treated as an economic variable determining collaboration efficiency — not merely an administrative practice.

12.10 Limitations

Several limitations should be acknowledged.

First, the observations are based on a single long-term research program.

Second, the collaboration environment involved specific AI systems and tools that may not be representative.

Third, many observations emerged naturally rather than through controlled experimentation.

As a result, the findings should currently be interpreted as exploratory rather than definitive.

Future studies involving additional researchers, AI systems, and research domains will be necessary to validate the generalizability of HARCT.

12.11 Research Continuity Failure Modes

Understanding why research continuity fails is as important as understanding how it succeeds. The following failure taxonomy was identified during the research program.

Failure Type	Mechanism	Effect
Human Forgetting Failure	Biological memory limits over time	Forgotten discoveries, decisions, and project relationships
AI Context Loss Failure	Finite context window boundaries	Earlier discussions become inaccessible within active sessions
Session Boundary Failure	Each new session requires re-establishing context	Repeated onboarding costs accumulate across sessions
Token Exhaustion Failure	Processing capacity limits	Premature termination, forced summarization, context compression
Repository Fragmentation Failure	Proliferation without integration mechanisms	Duplicate ideas, isolated projects, missing connections
Handover Failure	AI transitions without documentation	Context loss proportional to documentation gap
Goal Drift Failure	Local tasks displace strategic objectives	Research moves away from original purpose without detection
Future-Self Communication Failure	Incomplete artifacts for future retrieval	Future researchers repeat previous work unnecessarily
Reconstruction Failure	Context cannot be restored from available artifacts	The project effectively restarts from the beginning
Repository Drift Failure	Documentation diverges progressively from actual project state	Handover quality degrades silently; incoming collaborators receive misleading context

Failure Hierarchy:

```
Goal Drift
    ↓
Context Loss
    ↓
Reconstruction Failure
    ↓
Research Interruption
```

This hierarchy suggests that continuity failure is typically a systemic process rather than a single event. Goal Drift is frequently the earliest and most insidious failure — it may occur even when information is fully preserved, making it difficult to detect until significant damage to research coherence has accumulated.

Failure Modes Hypothesis:

> Long-term Human-AI research fails when reconstruction cost exceeds the ability of participants to restore research context and goals.

Under this hypothesis, continuity mechanisms are valuable because they reduce the probability of failure at each level of the hierarchy. Human-AI Research Continuity Theory can therefore be interpreted as a collection of countermeasures — each component addresses one or more specific failure modes.

Failure Mode	HARCT Countermeasure
Human Forgetting	Externalized Memory
AI Context Loss	Context Transfer + Handover Infrastructure
Session Boundaries	Collaboration Protocols
Token Exhaustion	Reconstruction Cost Reduction
Repository Fragmentation	Repository as Cognitive Infrastructure
Goal Drift	Goal Preservation Theory
Future-Self Communication Failure	Future-Self Collaboration Design
Reconstruction Failure	Memory Repository Architecture
Repository Drift Failure	Handover Quality Levels (Section 8.13)

12.12 Boundary Conditions

HARCT was developed from observations collected during a research program with several unusual characteristics: long duration, multiple AI systems, multiple repositories, and repeated context reconstruction. An important question therefore arises: under what conditions should HARCT be expected to apply?

Conditions Under Which HARCT Is Strongest:

- Research duration of months or years with repeated interruptions
- Multiple AI systems participating across different phases
- Context reconstruction is frequently necessary
- Goals remain relatively stable over the research lifespan

Conditions Under Which HARCT Has Limited Explanatory Power:

- Short projects (minutes, hours, or days) with little interruption — context remains active, reconstruction requirements are low
- Single-session interactions — the continuity problem does not arise
- Projects where context is continuously available to all participants — reconstruction cost is negligible

Partial Applicability:

Single-AI environments: HARCT remains applicable for context loss, reconstruction, and goal preservation, but AI-to-AI handover becomes irrelevant.

Human-only research (no AI): Several components generalize — Externalized Memory, Goal Preservation, Future-Self Collaboration, and Reconstruction Cost all apply. Human-AI Working Agreements and AI-specific handover practices do not.

Repository-free environments: HARCT predicts higher reconstruction cost, reduced continuity, and increased context loss. Alternative memory systems would be required.

Boundary Condition Hypothesis:

> The explanatory power of HARCT increases as research duration, interruption frequency, and context reconstruction requirements increase.

This hypothesis implies that continuity mechanisms become progressively more important as projects scale. A theory that clearly identifies its limits is more scientifically useful than one that claims universal applicability.

12.13 Testable Predictions

A scientifically useful theory must generate predictions that can be evaluated against future observations. The following predictions follow directly from HARCT.

Prediction 1 — Repository Quality Research programs with stronger repository infrastructure will exhibit higher continuity: faster project recovery, lower reconstruction cost, easier AI onboarding, and reduced context loss.

Prediction 2 — Documentation Quality Repositories with high-quality README files will require less reconstruction effort than repositories with poor documentation. `Good Documentation → Lower Reconstruction Cost → Higher Continuity`

Prediction 3 — Handover Artifacts AI transitions will produce less context loss when structured handover artifacts exist than when they do not. `AI Transition + Handover Document → Lower Context Loss`

Prediction 4 — Future-Self Artifacts Researchers who create future-oriented artifacts will recover context more efficiently and resume research faster than those who do not.

Prediction 5 — Goal Preservation Projects with explicit goal preservation mechanisms will survive longer interruptions than projects where goals are not deliberately preserved. `Goal Preserved → Context Reconstructed → Research Continues`

Prediction 6 — Reconstruction Cost Scaling As research duration increases, reconstruction cost will become an increasingly dominant factor determining overall collaboration efficiency — more important in year-scale projects than in week-scale projects.

Prediction 7 — Artifact-Centered vs. AI-Centered Knowledge Research continuity will be higher when knowledge is attached to persistent artifacts rather than to specific AI systems, enabling AI replaceability.

Prediction 8 — Collaboration Protocols Stable collaboration protocols will improve continuity by reducing the time and effort required to re-establish research context after interruptions.

Falsification Conditions:

HARCT would be substantially weakened if repeated observations demonstrate that:

- Documentation quality has no measurable effect on reconstruction cost or collaboration efficiency
- Handover artifacts do not reduce context loss after AI transitions
- Goal preservation mechanisms do not improve long-term project survival
- Repository infrastructure quality does not correlate with context recovery speed

Such observations would require fundamental revision of the theory's central claims.

---

13. Conclusion

13.1 Summary

This study investigated long-term Human-AI research collaboration through a multi-repository research program and proposed **Human-AI Research Continuity Theory (HARCT)** as a unified explanatory framework.

The central claim of HARCT is:

> Long-term Human-AI research collaboration is fundamentally a problem of research continuity rather than intelligence alone.

13.2 Theoretical Framework

HARCT comprises four theoretical components.

Goal Preservation Theory

Preserving research goals is more fundamental than preserving information.

Goals enable context reconstruction. Information without direction becomes an unusable archive.

Three-Component Cognitive Architecture

Research continuity requires three interacting components:

- Human goal memory
- AI reasoning capability
- Artifact-based context storage

Reconstruction Cost Theory

Documentation quality is a direct economic determinant of collaboration efficiency.

As research duration increases, reconstruction cost becomes a dominant factor.

Externalized Memory

Research context must be preserved outside active participants through persistent artifacts that support efficient context restoration.

13.3 Main Findings

Five major findings emerge from this study.

Finding 1: Memory Exists Beyond Individuals

Research context was frequently preserved outside both human memory and AI memory. Repositories, documents, logs, reminders, and handover files collectively functioned as memory systems. This led to the concept of Externalized Memory.

Finding 2: Goals Are More Fundamental Than Information

When goals survived interruptions, context could be reconstructed. When goals were lost, even complete archives became difficult to use effectively.

Finding 3: Research Continuity Is Distributed

Continuity emerged from interactions among three components: human goal memory, AI reasoning systems, and external artifacts. No single component was sufficient.

Finding 4: Reconstruction Cost Determines Efficiency

In this case, documentation quality appeared to determine the effort required to restore research context; this relation has not yet been tested (Section 13.4). This makes documentation quality an economic variable, not merely an administrative practice.

Finding 5: Future-Self Collaboration Is Real

A recurring observation was that future versions of the researcher frequently became participants in the research process. Documentation was often created specifically to assist future research activities, constituting a distinct form of collaboration across time.

13.4 Future Directions

Several avenues for future research remain open.

Empirical Validation of HARCT

Longitudinal studies of multiple researchers and research programs are needed to determine whether HARCT generalizes beyond the single-program case study presented here. Controlled comparisons of high-documentation vs. low-documentation research programs would provide direct evidence for or against Reconstruction Cost Theory.

Three specific validation designs are proposed.

Component Removal Experiments: Evaluate project recoverability after systematically removing each context component (Goal, State, History, or Next Action) in isolation. This directly tests

Proposition P9 — that full recoverability requires all four components simultaneously — and Proposition P8 — that different artifacts preserve different components.

Recovery Cost Comparison: Measure reconstruction effort (time, token consumption, artifact review count) under different combinations of available context components. This operationalizes Reconstruction Cost as a dependent variable and enables comparison of the Handover Quality Levels defined in Section 8.13.

Continuation Success Analysis: Compare successful and failed project recoveries across research programs or research teams. This tests whether context component availability predicts continuation outcomes, and whether programs with structured handover artifacts exhibit systematically lower context loss at AI transitions.

Formalization of HARCT

A candidate set of variables for formal modeling has been identified:

Variable	Definition
RC	Research Continuity
GP	Goal Preservation
CT	Context Transfer Quality
RCost	Reconstruction Cost
EM	Externalized Memory Strength
DRM	Distributed Research Memory
FSC	Future-Self Collaboration Effectiveness

A candidate conceptual formulation:

`RC ∝ (GP × CT) / RCost`

This formulation suggests that research continuity increases with stronger goal preservation and better context transfer, and decreases as reconstruction cost rises. Formal modeling should explore whether dependency relationships among these variables can be expressed as computable functions, and whether feedback loops exist — for example: higher continuity enabling better documentation, which further reduces reconstruction cost.

Measurement Framework

The following metrics are proposed for empirical evaluation of research continuity:

Metric	Definition
Reconstruction Time	Time required to restore sufficient context to resume productive work
Reconstruction Cost	Total effort (time, cognitive load, tokens) to restore context
Goal Recovery Rate	Probability that original research objectives can be correctly recovered after interruption
Handover Success Rate	Probability that a new collaborator can continue work after reviewing available artifacts
AI Replaceability Score	Degree to which a project can continue after replacing one AI system with another
Repository Recoverability	Ability to reconstruct project context from repository contents alone

Metric	Definition
Future-Self Recovery Time	Time for a future researcher to recover sufficient context to continue work

These metrics would allow HARCT to move from qualitative observation to empirical research program.

Research Program Emergence

A distinct phenomenon observed during this study was that independent investigations gradually became a coherent research program. Individual studies (K-Map Structure Invariance, Symmetric Boolean Functions, Variable Rearrangement) appeared unrelated at origin. Over time, structural connections became visible, leading to Structure Recognition Theory as a unifying hub — and ultimately to this study, which treats the research process itself as the object of investigation.

This emergence process suggests a further research question: how do coherent long-term research programs emerge from initially independent studies through Human-AI collaboration? Research Program Emergence may represent a distinct phenomenon from continuity maintenance — not merely preserving existing context, but enabling the accumulation of new structural relationships across previously isolated work.

AI-to-AI Handover Protocols

Standardized protocols for transferring research context between AI systems remain underdeveloped. Future work should investigate which artifact types most effectively reduce context loss during AI transitions, and whether structured handover formats can be designed to be AI-system-independent.

Cross-Disciplinary Extensions

HARCT concepts may generalize to domains beyond research collaboration:

- **Education:** Long-term student-AI tutoring relationships face similar continuity challenges. Goal Preservation and Future-Self Collaboration may apply to learning continuity across semesters.
- **Software Engineering:** Long-running human-AI codebases involve reconstruction costs analogous to those identified here. Repository infrastructure and documentation practices parallel those studied in this program.
- **Knowledge Management:** Reconstruction Cost Theory may offer a more precise economic framing for evaluating KM investment decisions.
- **Cognitive Science:** The three-component cognitive architecture (human / AI / artifact) extends distributed cognition models in ways that may be testable experimentally.

Long-Term Forecast

Three scenarios for the development of this research program are identified:

Scenario A — Case Study: The work provides a detailed illustration of one researcher's long-term collaboration practices. Contribution is primarily descriptive.

Scenario B — HARCT as Established Theory: Empirical validation across multiple research programs establishes HARCT as a recognized framework in Human-AI collaboration research. Reconstruction Cost and Goal Preservation become standard analytical tools.

Scenario C — Broader Distributed Cognition Theory: The three-component architecture (human / AI / artifact) generalizes into a broader theory of distributed cognition systems that include AI participants — extending existing distributed cognition and Extended Mind frameworks in ways relevant to the emerging AI-integrated research environment.

Given the single-case, uncontrolled evidence presented here, only Scenario A is directly supported; Scenario B depends on the validation studies proposed above. Scenario C represents the highest-value long-term outcome if the framework survives empirical scrutiny and cross-disciplinary extension.

13.5 Final Statement

This study began with a practical question:

> How can humans and AI systems continue a research program when neither participant can reliably preserve all necessary context?

The answer proposed by HARCT is:

Research continuity is not achieved by eliminating memory limitations.

It is achieved by constructing systems that make those limitations manageable — through goal preservation, minimized reconstruction costs, distributed cognitive architecture, and persistent external memory.

Long-term Human-AI research becomes possible not because memory is perfect, but because context can be preserved, restored, and transferred across time, participants, and systems.