

Woman Leadership 10 years: Topic modeling and Computational Grounded Theory Approach

Final analysis completed

Chungil Chae

Wed, 1 April 2026

Table of contents

	9
초록	9
연구 어젠다	9
연구 계획	10
일정	10
진행 현황 (5/18 완료)	10
저자 정보	11
저자 기여	11
저자	11
조연주 (Yonjoo Cho)	11
채충일 (Chungil Chae), Ph.D., M.S.	11
유지은 (Jieun You)	13
이수미 (Sumi Lee)	13
선언	13
IRB	13
재원	13
AI 활용	13
Research Logs	14
	15
2026-03-24	15
2026-03-23	15
2026-03-20	16
2024-06-28	16
2024-06-23	16
2024-05	16
2024-03-29	16
2023-09-29	17
2023	17
2015--2022	17
	18
2026-03-24 --- 연구방향 결정 세션 (채충일 + Claude Code)	18

2024-03-29 --- 팀 미팅	18
	19
GT 코딩 --- 1차 코딩	19
Version 2.0 (2026-03-23) 완료	19
Version 1.0 (deprecated)	19
STM 분석 --- 계산적 분석	19
Analysis8 (최신)	19
Analysis0~7 (이전 반복)	20
GT 코딩 --- 2차 코딩 (ATLAS.ti 범주화 빌드)	20
Phase 5N --- 네트워크 뷰 (2026-03-24) 완료	20
Phase 5 --- 최종 구조 (2026-03-24) 완료	20
Phase 4 --- 초점 코드 (2026-03-24) 완료	20
Phase 3 --- 코드 통합 (2026-03-24) 완료	21
Phase 2 --- 범주화 (2026-03-24) 완료	21
예정 분석	21
STM 재분석 (예정 --- 논문 A)	21
GT 2차 코딩 심화 (예정 --- 논문 C-2)	21
	22
논문 A --- 계산적 STM 분석	22
Ver 0.4 (2024-06-23) --- 조연주 교수 수정본	22
Ver 0.3 --- 채충일 수정본	22
Ver 0.2 (구글드라이브)	22
Ver 0.1 (구글드라이브)	22
논문 B --- 개념 모델 + 질적 분석	23
논문 C-1 --- AI-CGT 방법론 페이퍼	23
논문 C-2 --- 여성 리더십 GT 발견 결과	23
Research Note	24
	25
2026-03-24	25
논문 A --- 이중 STM 비교 아이디어	25
논문 B --- 개념 모델 접근	25
논문 C-1 --- AI-CGT 방법론 기여 포인트	26
논문 C-2 --- 잠정 이론: ``장벽의 자원화''	26
Q&A	27
2026-03-24	27
Q: 토픽 모델링에서 전문 vs. 코드+인용문의 장단점은?	27
Q: 기존 STM 분석(Analysis0~8)을 바로 활용할 수 있나?	27

Q: 팀 협업이 활성화 상태인가?	28
Q: 2026-04-30 마감은 무엇을 의미하는가?	28
Q: GT에서 AI는 어디까지 할 수 있는가?	28
	29
전체 연구 절차	29
GT 코딩 절차	30
1차 코딩 (완료)	30
2차 코딩 --- ATLAS.ti 빌드 체인 (완료)	30
다음 단계	31
논문 C-2: GT 2차 코딩 심화	31
논문 A: STM 분석	31
Theoretical Framework	32
	33
1차 코딩 기반 이론 도메인 강도 (RD-001, 감정)	33
핵심 이론 요약	33
역할 일치 이론 (Role Congruity Theory)	33
리더 정체성 전환 (Leader Identity Work)	34
Claiming-Granting 프레임워크	34
심리적 자본 (Psychological Capital)	34
한국 고유 개념: 어머니 리더십	34
Theoretical Relationship	35
	36
현재 상태	36
Paper A --- 이론적 배경	36
Paper B --- 이론적 배경	36
Paper C-2 --- 감정 이론 모형	37
논문별 이론 포지셔닝 요약	37
Methodology and Method	39
	40
연구 방법론 개요	40
Grounded Theory --- Strauss & Corbin (1998)	40
Grounded Theory --- Charmaz (2014, 구성주의 GT)	41
AI-Computational Grounded Theory (AI-CGT)	41
Structural Topic Modeling (STM)	42

Method	43
Data Collection	44
	45
데이터 개요	45
기반 질적 연구 8편	45
참여자 구성	46
GT 1차 코딩 결과 (2026-03-23 완료)	46
주요 테마 분포 (잠정, RD-001 기준)	46
데이터 접근 및 윤리	47
Literature Review	48
Searching and Inclusion & Exclusion	49
Round 1	50
Search keywords and category	50
keywords combination	50
Round 2	51
Search keywords and category	51
keywords combination	51
Round 3	52
Search keywords and category	52
keywords combination	52
PRISMA	53
Reference List	54
Round 1	54
Round 2	54
Round 3	54
Additional (during and after writing)	54
Reserach Problems	55
Key References	56
Quotes and Paraphrases	57
Fri Jul 21 02:27:35 KST 2022	58
일반적인 질문 (생각나는 대로)	58

HRD master class에서의 질문	58
TWS세션에서의 질문	60
## 리터리처들에서 도출된 질문들	60
Products	61
References	62

List of Figures

List of Tables

지난 10여 년간 연구팀은 일-가정 균형, 경력 및 리더십 개발, 문화적 맥락을 중심으로 한국 여성 리더에 대한 8개의 질적 연구를 수행하였다. HRD 연구에서 계산적 분석이 부상함에 따라, 질적 연구의 주관성 문제를 보완하기 위해 대규모 텍스트를 분석하는 계산적 기법인 토픽 모델링을 활용하여 200개의 내러티브를 재분석할 필요성을 느꼈다. 본 연구의 목적은 토픽 모델링을 통해 도출된 연구 토픽을 확인하고, 이를 기존 8개 질적 연구의 테마와 비교하며, HRD 연구와 실천에 대한 함의를 제시하는 것이다.

키워드: 여성 리더, 계산적 분석, 토픽 모델링, 근거 이론, 한국, HRD

본 프로젝트는 200개 심층 인터뷰(2015--2022)라는 단일 종단 데이터셋에서 4편의 논문을 산출한다.

논문	유형	핵심 기여	타겟 저널
A	계산적 분석	이중 STM 분석: 전문 vs. 코드+인용문	HRDQ, EJHRD
B	이론 적용	문헌 기반 개념 모델 → GT 코딩 데이터 적용	HRD 주제 저널
C-1	방법론 개발	AI-human-in-loop GT: 절차적 지식, 장단점, 한계	ORM, QRMJ 등
C-2	실질 발견	한국 여성 리더십: S&C + 차마즈 + 최신 방법 적용 결과	리더십·젠더 저널

마일스톤	목표일
1차 코딩 완료	2026-03-23
GT Phase 2-5 분석 빌드 완료	2026-03-24
HITL 시스템 구현	2026-03-24
AI-Work-5 v2 네트워크 시각화	2026-03-24
ATLAS.ti Phase 5 네트워크 확정	2026-03-25
연구방향 확정 (A, B, C)	2026-03-24
C-1 논문 --- 방법론 논문 작성	진행 중
C-2 논문 --- HITL → 핵심 범주 추출	진행 중
A 논문 --- STM 이중 분석	대기 (Analysis8 R 스크립트 검토 후)
B 논문 --- 개념 모델 선택	대기 (문헌 리뷰 후)
A 논문 --- 초고 완성	2026-04-30

(5/18)

- 200개 인터뷰 데이터 수집 (2015--2022)
- 1차 코딩 완료: 1,018 코드 / 5,707 인용문 / 25,204 링크
- ATLAS.ti Build 1~5 검증 완료
- 연구방향 A, B, C 확정 (2026-03-24)
- GT Phase 2-5 분석 빌드 완료 (57 카테고리, 40 포커스 코드)
- HITL(Human-in-the-Loop) 시스템 구현
- Phase 5 네트워크 분석 확정
- C-1 논문 방법론 기술 (ai-work-change 경험 문서화)
- C-2 논문 HITL 체크포인트 → 핵심 범주 추출
- HITL 연구자 편집 (3개 JSON 파일 완료 대기)
- A 논문 STM 재분석 (전문 + 코드·인용문)
- A 논문 Findings 작성
- B 논문 개념 모델 선택 및 문헌 리뷰
- C-2 2차 코딩 (SK-01 범주화~)

- 지도·감독: 조연주
- 행정관리: 조연주
- 개념화: 채충일
- 데이터 수집: 조연주
- 데이터 정제: 채충일, 이수미
- 형식 분석: 채충일
- 조사: 조연주, 유지은, 이수미
- 방법론: 채충일
- 자원: 조연주
- 검증: 조연주, 유지은, 이수미, 채충일
- 시각화: 채충일
- 원고 작성: 채충일, 조연주, 유지은, 이수미
- 검토·편집: 조연주, 유지은, 이수미, 채충일

(Yonjoo Cho)

University of Texas at Tyler

(Chungil Chae), Ph.D., M.S.

- chadchae@gmail.com
- orcid: [000-0002-7364-1525](https://orcid.org/000-0002-7364-1525)
- [Google Scholar](#)



채충일 박사는 중국 원저우-킨대학교 비즈니스 애널리틱스 조교수로, 데이터 분석, HRD, 다국적 기업 내 지식 공유를 전문으로 한다.

(Jieun You)

Valdosta State University

(Sumi Lee)

University of Georgia

IRB

2차 데이터 분석 --- IRB 면제(Category 4). 원 8개 연구는 Indiana University IRB 승인.

미정

AI

본 프로젝트는 AI 지원 질적 코딩 및 계산적 분석에 Claude Code(Anthropic)를 활용한다. 모든 AI 생성 결과물은 연구팀의 검토 및 검증을 거쳤으며, AHRD AI 책임 활용 가이드라인을 준수한다.

Research Logs

2026-03-24

- 기존문서 전수 검토 완료 (draft 1~4, 조연주 교수 2024-06 수정본, AHRD 포스터, IRB 서류, 문헌 통합 분석, CFP)
- 연구방향 A, B, C 논문 확정
- A 논문 확정: 이중 STM 분석 --- 전문 + 코드·인용문
- B 논문 확정: 문헌 기반 개념 모델 → 질적 분석 (ai-work-change 방식 적용)
- C-1 논문 확정: AI-human-in-loop GT 방법론 페이퍼 (방법론 저널)
- C-2 논문 확정: 한국 여성 리더십 GT 발견 결과 (주제 저널)
- README, _project.yaml, public Quarto 북 업데이트 완료
- 연구방향 문서 작성: RD-001 ~ RD-004
- GT 2차 코딩 ATLAS.ti 빌드 전체 완료:
 - Phase 2: 57개 범주 TagGroup + 범주별 메모 58개
 - Phase 3: freq 2-9 코드 통합 (276매핑 + 162 미분류)
 - Phase 4: 초점코드 40개 TagGroup + 메모 42개
 - Phase 5: 타입폴더 3개 (in-vivo/gerund/meaning) + 메모 3개, 누적 109개
 - Phase 5N: 네트워크 뷰 5개 + tagTagLink 21개 + memoNetworkLink/memoTagLink 연결
- HITL (Human-in-the-Loop) 시스템 구현 완료:
 - extract_researcher_edits.py: .atlasti diff → JSON
 - apply_researcher_edits.py: JSON 편집 → Phase 빌드 자동 반영
 - Phase 2~5 빌드 스크립트 모두 HITL 통합
- AI-Work-Change v2 네트워크 빌드 (AI-Work-5- _v2.atlasti):
 - 5 NV + 47 tagTagLink + 5 memoNetworkLink + 5 memoTagLink
 - save/close 크래시 해결 확인 (ATLAS.ti 25 버그 자체 해결)
- ATLAS.ti 빌드 패턴 지식베이스 업데이트 (코드 관계, 네트워크 뷰, HITL)

2026-03-23

- 1차 코딩 완료 검증

- 1,018 코드 / 5,707 인용문 / 25,204 코드-인용문 링크
- 200/200 문서 코딩 완료
- 이론적 포화 달성

- ATLAS.ti Build 1 (여성리더십_1_1차코딩완료.atlasti) 검증 완료

2026-03-20

- 프로젝트 단계 이관: 개시 단계 → 검증 중

2024-06-28

- 팀 업무 분장 메모 (조연주 → 채충일: 문헌 검토 완성, 분석 작성, Findings)
- 목표: 7월 말 초고 완성, 8월 19일 투고

2024-06-23

- 조연주 교수 수정본 제출 (draft_Jun23_2024_from_Cho.docx)

2024-05

- 토픽 레이블 검토 세션 (이수미, 유지은 검토)

2024-03-29

- 팀 미팅
 - 채충일: 연구 흠 소개, STM 5개 모델 작업 중
 - 레이블링: 유지은·채충일 → 조연주·이수미 검토 순서 제안
 - 조연주: 구글드라이브 draft 업데이트 중
- 다음 미팅: 4월 13일 오후 11시 (중국 시간)

2023-09-29

- 키워드 네트워크 분석 완료 (A Brief Report.docx)
- 단어 빈도 분석 (wordfreq.csv)

2023

- STM 분석 반복 실행: Analysis0 → Analysis8
- Draft 1 작성

2015–2022

- 8개 질적 연구 수행
- 200개 심층 인터뷰 수집 (다양한 섹터)

2026-03-24 — (+ Claude Code)

- 기존문서 전수 검토 완료
- 연구방향 A, B, C 확정

주요 결정사항: - 논문 A: 이중 STM 분석 (전문 + 코드·인용문) - 논문 B: 문헌 기반 개념 모델 → 질적 분석 (ai-work-change 접근 적용) - 논문 C-1: AI-human-in-loop GT 방법론 페이퍼 (방법론 전문 저널) - 논문 C-2: 한국 여성 리더십 실질 발견 (주제 저널)

확인 사항: - STM 분석 결과물(Analysis0~8) 존재. 방향 확정 후 업데이트 예정 - 조연주·유지은·이수미 교수팀 협업 활성화 - 목표: 2026-04-30 초고(A 논문) 완성

작성 문서: RD-002, RD-003, RD-004 (docs/ /)

2024-03-29 —

- 채충일: 연구 홈 소개 (https://chadchae.github.io/rp_2024_woman_leadership/)
 - 현재 5개 STM 모델 작업 중
 - 제안: 유지은·채충일이 레이블링 → 조연주·이수미 검토
- 조연주: 구글드라이브 draft 업데이트 중
- 다음 미팅: 4월 13일 오후 11시 (중국 시간)

GT — 1

Version 2.0 (2026-03-23)

- 방법: 근거 이론 초기 코딩 (AI-human-in-loop)
- 도구: Claude Code + ATLAS.ti 25
- 결과:
 - 1,087 코드 생성 → 병합 후 1,018 활성 코드
 - 5,707 인용문 (5개 스킵: offset 매핑 불가)
 - 25,204 코드-인용문 링크
 - 200/200 문서 코딩 완료
 - 이론적 포화 달성
- 산출물: _1_1 .atlasti (5.3 MB, 검증 완료)
- 데이터 위치: Analysis/Qual/ /1 _v2/

Version 1.0 (deprecated)

- 이전 코딩 시도. v2 재코딩으로 대체.
-

STM —

Analysis8 ()

- 상태: 완료. 연구방향 확정 후 업데이트 예정.
- 데이터: 전체 인터뷰 전문 (200개)
- 방법: STM (Structural Topic Modeling), R stm 패키지
- 위치: /Quan/Analysis8/

Analysis0~7 ()

- 전처리, 불용어, 토픽 수 최적화 반복 실험
 - 위치: /Quan/Analysis0~7/
-

GT — 2 (ATLAS.ti)

Phase 5N — (2026-03-24)

- 방법: tagTagLink + networkView + 메모 + HITL 시스템
- 도구: Python (native XML) + ATLAS.ti 25
- 결과:
 - 타입 폴더 3개: in-vivo / gerund / meaning
 - 네트워크 뷰 5개 (초점코드 간 관계 시각화)
 - tagTagLink 21개 (여성리더십 / AI-Work 검증 패턴 확립)
 - 메모-네트워크 연결 (memoNetworkLink 5개, memoTagLink 5개)
 - Human-in-the-Loop 시스템 구현 완료
- 산출물: _0_1_Phase5_ .atlasti (8.7 MB, 검증 완료)

Phase 5 — (2026-03-24)

- 결과:
 - 타입 폴더 3개: in-vivo / gerund / meaning (전체 코드 타입별 분류)
 - 메모 3개 (3층 구조 설명, 최종구조 요약, 다음단계 로드맵)
 - 누적 메모 109개
- 산출물: _0_1_Phase5_ .atlasti

Phase 4 — (2026-03-24)

- 결과:
 - 초점 코드 TagGroup: gerund 동명사형 상위 40개 선별
 - 선별 알고리즘: $score = freq \times 2 + doc_freq$
 - 메모 42개 (초점코드 40개 개별 + 선별기준 + 절차)
- 산출물: _0_1_Phase4_ .atlasti

Phase 3 — (2026-03-24)

- 결과:
 - freq 2-9 코드 438개 Phase 2 범주에 통합
 - 매핑 코드 276개 → 57개 범주 배정
 - 미분류 그룹 (162개 코드 보류)
 - 메모 3개 (통합절차, 미분류그룹 설명, 코드폴더vs코드그룹)
- 산출물: _0_1_Phase3_.atlasti

Phase 2 — (2026-03-24)

- 결과:
 - TagGroup 57개 (CAT-A01~005) 생성
 - 메모 58개 (범주별 설명 메모)
 - GT 범주화: 1,018 활성 코드 → 57개 의미적 군집
 - 산출물: _0_1_Phase2_.atlasti
-

STM (— A)

- 분석 A-1: 전문 (200개) → STM 업데이트
- 분석 A-2: 코드+인용문 (5,707개) → STM 신규 분석
- 목적: 이중 방법 비교 (논문 A의 핵심 기여)

GT 2 (— C-2)

- SK-02: 속성-차원 분석
- SK-03: 패러다임 모델 적용
- SK-07: 코드 동시출현 분석
- SK-10: 핵심 범주 선정 (초점코드 40개 → 5-10개 핵심범주)
- SK-11: 스토리라인 구성

A — STM

Ver 0.4 (2024-06-23) —

- 초록 업데이트
- 서론 정제
- 문헌 검토: 질적 연구의 계산적 분석 섹션 확장
- Findings: 여전히 비어 있음
- 파일: /draft_Jun23_2024_from_Cho.docx

Ver 0.3 —

- 방법 섹션 상세화
- 분석 절차 기술
- 참고문헌 완성
- 파일: /draft 2_Chad.docx

Ver 0.2 ()

- Draft 2 --- 구조 확립
- 파일: /google drive-20240622/draft/draft 2.docx

Ver 0.1 ()

- Draft 1 --- 초기 구조
 - 파일: /google drive-20240622/draft/draft 1.docx
-

B — +

- 상태: 방향 확정 (2026-03-24). 개념 모델 선택 미결정.

C-1 — AI-CGT

- 상태: 방향 확정 (2026-03-24). ai-work-change 방법론 노트 기반 확장.

C-2 — GT

- 상태: 방향 확정 (2026-03-24). 2차 코딩 완료 후 착수.

Research Note

2026-03-24

A — STM

논문 A의 핵심 방법론적 통찰: 동일 데이터셋에서 두 가지 STM 분석을 실행하면 내재적 삼각검증이 가능하다.

- 분석 1 (전문): 기계 주도, 연구자 선택 편향 없음, 노이즈 포함
- 분석 2 (코드+인용문): 연구자 정제 신호, 노이즈 제거, 선택 편향 존재

두 분석에서 수렴하는 토픽 = 강건한 발견 (robust findings) 발산하는 토픽 = 방법론 논의 소재 (각 접근이 무엇을 더 포착하는가)

이 프레임은 논문 A를 단순히 ``STM을 적용했다''가 아니라 ``동일 데이터에서 두 계산적 접근을 비교했다''는 더 강한 방법론적 기여로 포지셔닝한다.

B —

ai-work-change 방법론을 여성 리더십에 적용:

1. 문헌 리뷰 → 개념 모델 선택
2. 선택한 모델을 1,018개 GT 코드 + 5,707개 인용문에 적용
3. 모델 검증, 확장, 또는 도전

RD-001 모델 선택 힌트 (코드 빈도 기반): - 젠더와 리더십 (3,758회) --- 역할 일치 이론 (Eagly & Karau, 2002) - 리더 정체성 (2,815회) --- 이바라의 리더 정체성 전환 (2010) - 일-가정 (2,507회) --- Greenhaus & Powell (2006)

가능한 통합 모델: 젠더 장벽 → 정체성 작업 → 리더십 이머전스

C-1 — AI-CGT

GT에서 AI가 기여하는 지점: - 초기 line-by-line 코딩의 속도·일관성 - 200개 문서 전체에 걸친 대규모 지속적 비교 - 대규모 코퍼스에서 동시출현 패턴 파악

인간 판단이 필수인 지점: - 이론적 민감성 --- 어떤 코드가 분석적으로 유의한가 - 문화적 해석 --- 한국 특유의 표현과 사회적 맥락 - 포화 판단 --- 범주가 포화되었는지 최종 판단 - 이론 구축 --- 범주를 일관된 이론으로 연결

AI-CGT의 독특한 특성: - 연구자 역할의 전환: ``코더''에서 ``큐레이터·검증자''로
- AI-인간 결정의 투명한 감사 추적 - 새로운 성찰성(reflexivity) 실천 가능성 (AI는 위치성이 없음 --- 구성주의 GT에서 어떻게 처리할 것인가?)

C-2 — : “ ”

RD-001 분석 기반 (81개 문서, 감정):

[.]
↓ (421)
[+]
↓
[(grant)]
↓
[+]
↓
[. .]
↓
[- (283)]

핵심 명제: 한국 여성 리더십은 근본적으로 ``장벽의 자원화(adversity-as-resource)'' 과정이다.

``어머니 리더십'' --- 새로운 구성개념 (77회): 돌봄 + 훈육 + 희생 + 장기적 관점.
서구 리더십 이론에 없는 한국 문화 고유 구성개념. 서번트 리더십과 유사하지만 문화적으로 다름.

Q&A

2026-03-24

Q: vs. + ?

전문(Full Transcripts) → STM

장점: - 표준적 접근, 기존 STM 문헌과 직접 비교 가능 - 연구자 선택 편향 없음 ---
기계가 전체 텍스트에서 패턴 발견 - 코퍼스가 클수록 LDA/STM 결과 안정적 - 전처리만
동일하면 재현 가능

단점: - 노이즈 다량 포함: 면담자 질문, 추임새, 주제 이탈 발화 - 인터뷰 구조 편향:
토픽이 질문 카테고리를 반영할 수 있음 - 한국어 형태소 처리 부담 (자모분리 문제
실제 존재) - 1,018개 GT 코딩 자산 미활용

코드+인용문 → STM

장점: - 고밀도 신호: 연구자가 선별한 5,707개 인용문, 노이즈 제거 - STM 공변량 완전
활용 (섹터, 연도, 직위, 혼인 상태) - 방법론적 독창성: Computational Grounded Theory
(Nelson, 2020) 실제 구현 - 삼각검증: 동일 데이터에서 GT 코드 vs. STM 토픽 비교 -
한국어 처리 유리: 의미 있는 문장 단위, 전처리 부담 감소

단점: - 선택 편향: 코딩되지 않은 발화는 분석에서 제외 - 짧은 텍스트 세그먼트:
LDA가 매우 짧은 텍스트에서 불안정할 수 있음 (5,707개 물량으로 보완) - 잠재적
순환논리: 연구자 선별 텍스트 → 기계가 연구자 코딩 확인

결정 (2026-03-24): 둘 다 수행. 결과 비교 자체가 논문 A의 방법론적 기여.

Q: STM (Analysis0~8) ?

답변: 결과물 존재. 연구방향 확정 후 업데이트 예정. 분석 파이프라인(R 스크립트,
전처리)은 /Quan/Analysis0~8/에 있음.

Q: ?

답변: 그렇다. 조연주·유지은·이수미 교수 협업 중.

Q: 2026-04-30 ?

답변: 초고(논문 A) 완성 목표.

Q: GT AI ?

GT는 연구자를 주된 해석 주체로 전제한다. AI-human-in-loop GT에서:

GT 단계	AI 역할	인간 역할
초기 코딩 (line-by-line)	높은 일관성으로 수행 가능	검토·검증
이론적 민감성	부분 가능 (문헌 기반 패턴 인식)	필수 --- 맥락적 판단
지속적 비교	200개 문서 전체에서 대규모 수행 가능	방향 설정·해석
포화 판단	빈도 패턴 보조 가능	최종 해석적 판단
메모잉	분석 메모 초안 작성 가능	심화·수정·검증
이론 구축	독립적으로 불가	주도 필수

이 경계 지도가 논문 C-1의 핵심 기여다.

[Phase 1]

8 (2015-2022) → 200

[Phase 2]

DOC/DOCX → TXT → →

[Phase 3] GT 1 (2026-03-23)

AI-human-in-loop

1,018 / 5,707 / 25,204

[Phase 4] (2026-03-24)

A, B, C-1, C-2

[Phase 5-A] STM → A ()

1: → STM

2: + → STM

→ Findings

[Phase 5-B] → B ()

→ → 1

[Phase 5-C] GT 2 → C-2 /

SK-01 (57 Phase2) → SK-09 (Phase5N)

SK-02 → SK-03 → SK-07

SK-10 → SK-11 →

[Phase 6] → C-1 ()

ai-work-change +

AI-CGT

[Phase 7]

GT

1 ()

1. 배치 단위 (5개 문서씩) 코딩
2. Claude Code가 line-by-line 초기 코딩 수행
3. 연구자 검토 및 수정
4. 코드 병합 및 정제 (1,087 → 1,018)
5. ATLAS.ti native XML 빌드 (atext3 역공학 방식)
6. 검증: ATLAS.ti 25에서 수동 확인

2 — ATLAS.ti ()

HITL (Human-in-the-Loop) 체크포인트 구조:

```
Phase 2: 57 TagGroup →  
    ↓ (researcher_edits_phase2.json )  
Phase 3: freq2-9 →  
    ↓ (researcher_edits_phase3.json )  
Phase 4: 40 →  
    ↓ (researcher_edits_phase4.json )  
Phase 5: + →  
    ↓  
Phase 5N: +
```

각 체크포인트에서 연구자가 ATLAS.ti에서 직접 편집 후 저장하면, 다음 Phase 빌드에 자동 반영됨.

C-2: GT 2

1. 연구자 HITL 체크포인트 실행 (Phase 2~5 검토 및 편집)
2. 초점코드 40개 간 관계 심화 분석 (NetworkView 활용)
3. 핵심범주 도출: 초점코드 군집 → 5-10개 핵심범주 후보
4. 이론적 포화 검증: 부정사례 탐색, 인터뷰 재독
5. 선택코딩: 핵심범주 통합 스토리라인 구성

A: STM

1. Analysis8 R 스크립트 검토
2. 전처리 업데이트 (200개 전체 기준 재실행)
3. 분석 1 실행: 전문(200개) → STM
4. 분석 2 실행: 코드+인용문(5,707개) → STM
5. 토픽 레이블링 (팀 검토)
6. Findings 섹션 작성
7. Discussion, Implications 작성
8. 초고 완성 (2026-04-30 목표)

Theoretical Framework

1 (RD-001,)

리더십 이론	코드 빈도	핵심 문헌
젠더와 리더십	3,758회	Eagly & Carli (2007) 미로 이론; 역할 일치 이론
리더 정체성	2,815회	Ibarra et al. (2010) 리더 정체성 전환
일-가정·경력	2,507회	Greenhaus & Powell (2006) 일-가정 풍요 이론
회복탄력성	1,522회	Luthans (2002) 심리적 자본
리더십 개발	1,498회	Day (2000); McCauley et al. (2010) 70:20:10
서번트 리더십	1,293회	Greenleaf (1977)
리더십 이머전스	922회	DeRue & Ashford (2010) claiming-granting
진정성 리더십	852회	Avolio & Gardner (2005)
관계적 리더십	689회	Uhl-Bien (2006)
변혁적 리더십	659회	Bass & Avolio (1994)

(Role Congruity Theory)

Eagly & Karau (2002). 여성에 대한 사회적 역할 기대와 리더십 역할 간의 불일치가 여성 리더에 대한 편견을 만든다. 여성이 리더십을 주장(claim)할 때 역할 불일치 비용이 발생한다.

데이터 연결: 한국 여성 리더가 선언적 claim보다 타인 인정(grant) 경로로 이머전스하는 현상 설명.

(Leader Identity Work)

Ibarra et al. (2010). 리더십 역할을 맡는 것은 자아 개념의 전환을 수반한다. 정체성 작업(identity work)이 리더십 개발의 핵심이다.

데이터 연결: ``스스로 모드 체인지를 하려고 많이 노력을 해요'' (150회) --- 자기주도적 정체성 재구성 코드.

Claiming-Granting

DeRue & Ashford (2010). 리더십 이머전스는 개인이 리더 정체성을 주장(claim)하고 타인이 이를 인정(grant)하는 상호작용 과정이다.

데이터 연결: 한국 여성 리더는 grant-based 이머전스가 지배적 --- 실적 기반(895회) + 관계 기반(254회).

(Psychological Capital)

Luthans (2002). 자기효능감, 희망, 낙관주의, 회복탄력성의 통합 구성개념.

데이터 연결: ``버틸 수 있는 힘이 길러진 것 같아요'' (136회) --- 역경 후 회복탄력성.

:

서구 리더십 이론에 없는 한국 문화 고유 구성개념 (77회 등장).

구성 요소: 돌봄(care) + 훈육(discipline) + 희생(sacrifice) + 장기적 관점 팀원을 ``자식처럼'' 육성하는 리더십 철학

서번트 리더십과 유사하나 유교 문화적 의미 층위가 다름: - Servant Leadership: 섬기는 자아(serving self) - 어머니 리더십: 양육하는 자아(nurturing self) + 훈육 권위

Paper C-2에서 독자적 구성개념으로 이론화 예정.

Theoretical Relationship

이론적 프레임워크는 논문별로 별도 구성. 아래는 각 논문의 이론적 포지셔닝.

Paper A —

Computational Analysis in Qualitative Research

질적 연구에서 주관성 문제를 보완하기 위한 계산적 분석의 역할: - CAQDAS(NVivo, ATLAS.ti) → 연구자 코딩 지원 - Computational GT(Nelson, 2020) → 기계 주도 패턴 탐지 - Topic Modeling → 대규모 텍스트에서 잠재 토픽 추출

연구 질문 (기존): > ``기존 8개 질적 연구의 테마와 STM 토픽 모델링에서 도출된 핵심 토픽을 어떻게 비교할 것인가?``

확장된 연구 질문 (2026-03-24 업데이트): > ``동일 데이터에서 전문(full transcript) 기반 STM과 코드+인용문 기반 STM은 어떻게 다른 토픽을 생성하며, 그 차이의 방법론적 함의는 무엇인가?``

Paper B —

개념 모델 선택 기준 (선정 예정)

후보 모델 (RD-001 코드 빈도 기반): 1. 역할 일치 이론(Eagly & Karau, 2002) + 리더 정체성 전환(Ibarra, 2010) 통합 2. 심리적 자본 모델(Luthans, 2002) 3. 리더십 이머전스 claiming-granting(DeRue & Ashford, 2010)

연구 질문 (예정): > ``선택한 개념 모델은 한국 여성 리더십 데이터에서 어떻게 검증되고 확장되는가?``

Paper C-2 —

``장벽의 자원화 (Adversity-as-Resource)'' 모형 (감정, 2차 코딩 후 검증 필요)

```
[ . ]
  ↓ (421)
[ + ]
  ↓
[ (grant ) ]
  ↓
[ + ]
  ↓
[ . . ]
  ↓
[ - (283) ]
```

핵심 명제: 한국 여성 리더십은 ``장벽의 자원화(adversity-as-resource)'' 과정 --- 젠더 차별이 리더십 개발의 외부 압력으로 작용하여, 역경·자기주도 학습을 강제하고, 실적 기반 이머전스와 모성·공적 스타일로 귀결된다.

서구 이론과의 차별점: - Day(2000) 70:20:10 모형 수정 필요: 역경 학습이 독립 구성요소로 포함되어야 함 - Claiming 보다 Granting 이머전스가 지배적 --- 젠더 역할 규범이 원인 - 어머니 리더십은 서구 이론에 없는 한국 고유 구성개념

논문	이론적 기여 유형	핵심 주장
A	방법론 혁신	두 STM 접근의 비교가 질적 재분석의 새로운 표준 제시
B	이론 검증·확장	기존 이론이 한국 문화 맥락에서 어떻게 수정되어야 하는가

논문	이론적 기여 유형	핵심 주장
C-1	방법론 개발	AI-CGT 절차 프로토콜 --- GT에서 AI의 역할 경계 설정
C-2	이론 생성	장벽의 자원화 + 어머니 리더십 --- 한국 여성 리더십 고유 이론

Methodology and Method

본 연구는 세 가지 방법론적 접근을 통합한다.

방법론	적용 논문	핵심 특징
Structural Topic Modeling (STM)	논문 A	계산적 텍스트 분석, 귀납적
개념 모델 기반 질적 분석	논문 B	연역-귀납 혼합
AI-Computational Grounded Theory (AI-CGT)	논문 C-1, C-2	AI-human-in-loop GT

Grounded Theory — Strauss & Corbin (1998)

핵심 원칙: - 데이터에서 이론을 귀납적으로 생성 - 지속적 비교법(constant comparison)
- 이론적 표본 추출(theoretical sampling) - 이론적 포화(theoretical saturation)

주요 단계:

1. 개방 코딩(open coding): 개념 생성
2. 축 코딩(axial coding): 패러다임 모델 적용
3. 선택 코딩(selective coding): 핵심 범주 통합

패러다임 모델 구성요소: 인과적 조건 → 현상 → 맥락 → 중재 조건 → 전략/행위 → 결과

Grounded Theory — Charmaz (2014, GT)

핵심 원칙: - 연구자와 참여자가 공동으로 의미를 구성 - 연구자의 위치성(positionality)과 성찰성(reflexivity) 강조 - 동명사 코딩(ground coding): 과정과 행위 중심

주요 단계:

1. 초기 코딩(initial coding): 동명사 형태, 행위 포착
2. 초점 코딩(focused coding): 가장 유의미한 초기 코드 선별
3. 이론적 코딩(theoretical coding): 코드 간 관계 개념화

AI-Computational Grounded Theory (AI-CGT)

정의: AI-human-in-loop 방식으로 GT를 수행하는 절차적 방법론. Nelson(2020)의 Computational Grounded Theory를 AI와의 협업으로 확장한 접근.

GT 단계별 AI 역할과 인간 역할:

GT 단계	AI 역할	인간 역할
초기 코딩	높은 일관성으로 수행 가능	검토·검증
이론적 민감성	부분 가능 (문헌 기반)	필수 --- 맥락적 판단
지속적 비교	대규모 수행 가능	방향 설정·해석
포화 판단	빈도 패턴 보조	최종 해석적 판단
메모잉	초안 작성 가능	심화·수정·검증
이론 구축	독립적으로 불가	주도 필수

이 연구에서의 구현: - Claude Code를 AI 파트너로 활용 - 200개 인터뷰 배치 코딩 (5개 문서씩) - 연구자 검토 후 ATLAS.ti 25에 통합 - 논문 C-1에서 절차와 방법론적 특이점 전면 보고

Structural Topic Modeling (STM)

방법: Roberts et al. (2019)의 STM --- LDA 기반에 문서 메타데이터를 공변량으로 통합

특징: - 문서별 토픽 비율이 메타데이터(섹터, 연도, 직위 등)에 따라 달라질 수 있음 - R `stm` 패키지 사용

본 연구 적용 방식 (논문 A): - 분석 1: 200개 전문 텍스트 입력 - 분석 2: 5,707개 코드+인용문 입력 - 공변량: 섹터, 인터뷰 연도, 위계 레벨, 혼인 상태

참고 문헌:

- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). `stm`: An R package for structural topic models. *Journal of Statistical Software*, 91(2), 1--40.
- Nelson, L. K. (2020). Computational grounded theory. *Sociological Methods & Research*, 49(1), 3--42.
- Baumer, E. P. et al. (2017). Comparing grounded theory and topic modeling. *JASIST*, 68(6), 1397--1410.

Method

Data Collection

항목	내용
총 인터뷰	200개
수집 기간	2015--2022
수집 방법	반구조화 심층 인터뷰
원본 형식	DOC (158개) + DOCX (42개)
변환 형식	TXT (ATLAS.ti 분석용)

8

연구	연도	참여자 수	주제
Cho et al.	2015	18명	일-가정 균형
Cho et al.	2016	50명	기업 여성 리더 일-가정 불균형
Cho et al.	2017	19명	경력 성공의 성별 차이
Han et al.	2018	21명	---
Cho et al.	2019	15명	여성 CEO 경력 도전 (MNC)
Cho et al.	2020	3명	여성 창업가
Cho et al.	2021	23명	---
Cho et al.	2022	50명	---
합계		약 200명	

항목	범주
직종	기업임원, 팀리더, 의료(의사/교수), 정치, 언론, 학교, 비영리, 대학박사, 정부
위계레벨	7단계 (최고경영진 ~ 팀원)
위신점수	20--100 연속 척도
혼인 상태	기혼, 미혼, 이혼, 사별
자녀 유무	있음 / 없음

상세 코드북: Analysis/Data_Original/ _ _ .md 참조

GT 1 (2026-03-23)

항목	수치
총 코드 (병합 전)	1,087개
총 코드 (병합 후, 활성)	1,018개
코드 유형	In-vivo / Gerund / Meaning
총 인용문	5,707개 (5개 스킵)
Code-Quotation 링크	25,204개
코딩 완료 문서	200 / 200
이론적 포화	달성

(, RD-001)

테마	코드 빈도
성차별·유리천장	3,742회
일-가정 양립	2,799회
정체성·내적동력	2,360회

테마	코드 빈도
경력·성취	1,881회
사회적 역할인식	1,527회
리더십 스타일	1,315회
역경·회복력	1,218회
네트워크·관계	737회

- IRB: 원 8개 연구 --- Indiana University IRB 승인
- 본 재분석 --- UT Tyler IRB Exempt (Category 4: 2차 데이터 분석)
- 익명화: 참여자 식별정보 제거, 그룹코드(A-D) + 가명 사용
- 데이터 보관: 연구팀 내부 보안 스토리지

Literature Review

Searching and Inclusion & Exclusion

Round 1

Search keywords and category

keywords combination

Round 2

Search keywords and category

keywords combination

Round 3

Search keywords and category

keywords combination

PRISMA

Reference List

Category, Classification and Decision Note for Selected Literature in Rounds

Round 1

Round 2

Round 3

Additional (during and after writing)

Reserach Problems

Key References

Quotes and Paraphrases

Fri Jul 21 02:27:35 KST 2022

()

- What is : (definition)
 - human resource analytics
 - people analytics
 - workplace analytics
 - learning analytics
 - computational social science
 - marketing analytics
 - audit analysis
 - business Analytics
- What is the characteristics of analytics
- What is the similarity of the analytics
- What is difference between analytics
- What is main frame(core/idea) of the analytics

HRD master class

- What do we mean by Data Analytics
 - how it is defined and how does our understanding of the term vary around the world?
- What's the relationship between Data Analytics and the concepts of work and the workplace?
 - how has that relationship changed over the years?
- What's the connection between Data Analytics and HRD?
 - How does research, theory and practice in one influence research, theory and practice in the other?

- If we look at the challenges facing HRD in organizations over the next few years, how does Data Analytics help HRD in understanding those challenges and working on them?
- What do we mean by Data Analytics
 - how is it defined and how does our understanding of the term vary around the world?
- What's the relationship between Data Analytics and the concepts of work and the workplace?
 - how has that relationship changed over the years?
- What's the connection between Data Analytics and HRD?
 - how does research, theory and practice in one influence research, theory and practice in the other?
- If we look at the challenges facing HRD in organizations over the next few years, how does Data Analytics help HRD in understanding those challenges and working on them?
- I'm wondering what happens if HRD takes a more data analytics approach to addressing organizational challenges
 - to pick on one for an example, if we consider how best to help employees to learn from each other especially with employees working remotely or in hybrid forms,
 - how can data analytics help HRD to address that challenge?
- Within organizations, what HRD questions could you see data analytics answering in the coming years that, in most organizations, currently go unanswered?
- Earlier in the episode, you mentioned leveraging data that already exists within organizations -- for a typical organization, what existing data could HRD use and where does that data reside?
- What advice do you have for HRD professionals who want to learn how to do data analytics
 - how to learn the basics of data analysis and what tools to use?
- What advice do you have on how HRD professionals could tap into data experts working within organizations
 - what job titles might those people have, and are there any best practices around how to work with them?

TWS

- 먼저 HRA/PA란 무엇인지 더 명확하게 알고 싶습니다
 - HRA/PA이란 무엇을 의미합니까?
 - HRD 분야 연구/ 실무 측면에서 HRA/PA 에 대한 이해 및 활용이 왜 중요할까요?
 - HRD 입장에서 조직의 어떤 문제들을 해결해 줄까요?
- 현재까지 학계와 현장에서는 얼마나 활용되고 발전해 왔는지 궁금합니다.
 - 먼저 연구 측면에서는 현재까지 어떤 발전/변화가 있었나요?
 - HRD 분야에서는 연구가 이루어 지고 있나요?
- 또 실무에서 HRD가 현장의 문제를 해결하기 위해 더 많은 데이터 분석 접근 방식을 취하면 어떻게 될지 궁금합니다.
 - HRD 현장/실무에서는 어느 정도 도입 및 활용이 되고 있나요?
 - 현장에서의 HRA/PA 적용 사례(Best Practice)가 있을까요?
- 그렇다면 당장 HRD 실무자들이 HRA/PA 도입 및 활용을 위해 무엇을 할 수 있을지 궁금합니다.
 - HRD에서 당장 사용할 수 있는 기존 데이터는 무엇이 있고 어떻게 활용 가능한지?
 - 어떤 결과를 기대 할 수 있는지?
 - 조직 내 타부서 또는 외부 데이터 전문가를 활용할 수 있는 방법도 있을까요?
- 마지막으로 앞으로의 방향성 그리고 HRD 실무자에게 주는 조언에 대해 여쭙고 싶습니다.
 - HRA/PA에 대한 전망과 HRD에 주는 시사점은 무엇인지?
 - 앞으로 HRA/PA를 도입 하고자 하는 HRD 전문가에게 해주고 싶은 조언

##

Products

References