

다중 증상 디지털 정신건강 플랫폼을 위한 AI 기반 개인 맞춤형 치료 경로 엔진: 알고리즘 아키텍처, 프라이버시 보존 교차 모달 동기화, 그리고 감사 추적 가능한 임상 추론

An AI-Driven Personalized Treatment Pathway Engine for Multi-Symptom Digital Mental Health Platforms: Algorithmic Architecture, Privacy-Preserving Cross-Modal Synchronization, and Audit-Traceable Clinical Reasoning

Alex Kwon¹

¹Boston Neuromind LLC, Boston, MA, USA

Former Visiting Scholar, Harvard Graduate School of Education (Mind, Brain, and Education)

Correspondence: alex@bostonneuromind.com

초록

배경. 디지털 정신건강 중재(DMHI)는 빠르게 확산되었으나, 증상 평가와 개인화된 세션 수준의 치료 계획 사이에는 지속적인 격차가 존재한다. 대부분의 플랫폼은 진단 결과를 표시하지만, 개인의 맥락에 적응하는 근거 기반 치료 경로로 연결하지 못한다. 두 번째 격차는 대화형 에이전트(봇)와 임상 평가 모듈의 통합에 관한 것이다: 봇이 수집한 데이터는 임상적으로 가치 있으나, 사용자의 명시적 통제 없이 공유될 경우 심각한 프라이버시 문제를 야기한다. **목적.** 본 연구는 다중 증상 디지털 정신건강 플랫폼 (Boston Neuromind의 NeuroCatchers + TalkCatcher 이중 모듈 아키텍처)에서 평가-치료 경로를 운영화하는 6개 알고리즘 엔진을 제시한다. 10개 증상 영역(ADHD, 우울증, 불안, 수면, 번아웃, PTSD, OCD, 양극성장애, 자폐, 최고 수행)을 다룬다. **방법.** (a) 환자 평가, 훈련 라이브러리, 추천 추론, 세션 기록을 연결하는 통합 관계형 데이터베이스 스키마; (b) 증상 일치, 바이오마커 일치, 근거 강도(Cochrane d 정규화), Fischer 역동적 기술 호환성, 실용적 적합성, 개인 이력의 6개 구성요소를 통합한 추천 점수 산정자(ALG-1); (c) Fischer 수준을 인식하여 N개 세션에 기술 진행을 분배하는 세션 계획 생성자(ALG-2); (d) 세션 수 변경, 목표 수정, 중도 이탈 신호에 의해 트리거되는 적응형 재계획자(ALG-3); (e) 옵트인 기본 OFF 의미를 가진 프라이버시 보존 봇 ↔ Catcher 브리지(ALG-4); (f) 임상가/내담자 대면 맞춤 훈련 빌더(ALG-5); (g) 임상 버전과 환자 버전을 병렬로 생성하는 한·영 이중 언어 결과 렌더러(ALG-6)를 명세한다. **결과.** 아키텍처는 종단 간 감사 추적성을 제공한다: 모든 추천은 점수 구성요소, 근거 인용(PMID/DOI), 양 언어로 된 추론, 그리고 이를 생성한 모델 버전을 영구 저장한다. ADHD는 후속 증상 영역 복제를 위한 황금 템플릿 역할을 한다. **결론.** 본 설계는 다중 증상 개인 맞춤형 디지털 정신건강을 위한 일반화 가능하고 규제 정렬된(FDA SaMD, HIPAA, GDPR 9조 및 22조) 프레임워크를 제공하며, 추천 알고리즘이 대화형 데이터와 임상 데이터 스트림 간의 프라이버시 경계를 존중하면서 투명하게 작동할 수 있는 방법을 보여준다.

키워드: 디지털 정신건강; 개인화; 임상 의사결정 지원; 적응형 치료; 프라이버시 보존 AI; 설명 가능한 AI; Fischer 역동적 기술 이론; 감사 추적성

2. 서론

2.1 디지털 정신건강 환경과 개인화 격차

지난 10년간 디지털 정신건강 중재(DMHI)는 스마트폰 앱, 웹 기반 프로그램, 대화형 에이전트, 가상현실 경험을 포함하여 현저히 확장되었다(Torous et al., 2021). 메타분석 증거는 앱 기반 중재가 우울증, 불안 및 관련 상태에 대해 소-중 효과 크기를 산출하며, 적응적 또는 개인화된 요소를 포함할 때 효과 크기가 일관되게 더 크다는 점을 보여준다(Linardon et al., 2024). 개인화 자체는 문헌에서 이질적인 구성 개념이다: Hornstein 등(2023)은 우울증에 대한 DMHI 에서만 최소 6 개의 개인화 차원을 식별했으며, Wanniarachchi 등(2025)은 청소년 우울 및 불안에 대한 DMHI 67 개 연구에서 67 개의 개별 개인화 변수를 분류했다.

그러나 구조적 격차가 지속된다. 배포된 DMHI 에서 지배적인 패턴은 증상 선별 후 정적 심리교육 콘텐츠, 일반적 기술 연습 또는 비구조화된 챗봇 대화로 이어진다. 진단 신호(예: 환자 건강 설문지-9 점수, ADHD 자가보고 선별 결과, 바이오마커 패턴)에서 환자의 증상 프로파일, 바이오마커 서명(가용 시), 발달 기술 수준, 실용적 제약, 개인 이력에 매핑되는 개인화된 순서가 있는 근거 기반 치료 활동 순서로의 번역—임상가, 환자, 규제 기관이 감사할 수 있는 명시적 추론과 함께—은 운영 시스템에서 거의 다루어지지 않았다. 이것이 본 연구가 다루는 중심 문제이다.

2.2 평가에서 행동으로: 치료 경로 문제

본 연구에서 사용하는 치료 경로(treatment pathway)는 환자의 증상 프로파일, 바이오마커 서명(가용 시), 발달 기술 수준, 실용적 제약, 개인 이력에 매핑되는 근거 기반 치료 활동의 순서화된 개인 맞춤형 시퀀스이다. Murphy(2005)가 개척하고 Almirall 등(2014) 및 Nahum-Shani 등(2018)이 정교화한 적응형 치료 전략 프레임워크는 정신건강에서 순서화 결정을 위한 개념적 발판을 제공하지만, 이러한 프레임워크는 10 개의 별개 증상 영역을 담당하는 실제 플랫폼이 평가-치료 번역을 대규모로 운영화할 수 있는 데이터 구조나 알고리즘 메커니즘을 명세하지 않는다.

신체 의학의 임상 의사결정 지원 시스템(CDSS)은 영구 저장된 추론을 갖춘 알고리즘 추천의 가치를 입증해 왔으나(예: 간세포암에 대한 Lee et al., 2024; 개관을 위한 Sutton et al., 2020), 정신건강 CDSS 는 부분적으로 (a) 진단 범주 내 증상 이질성이 크고, (b) 바이오마커 근거가 종양학이나 심장학보다 약하며 (c) 환자의 주관적 선호와 발달 단계 자체가 결정에 대한 실질적 입력이라는 점에서 뒤쳐져 왔다. 우리는 이러한 장애물이 알고리즘 추천을 배제하기는커녕 오히려 필요한 아키텍처를 정의한다고 주장한다: 주관적 선호, 발달 기술 매칭, 감사 추적 가능한 추론을 명시적으로 처리하는 다중 구성 요소 점수 산정자가 필요하다.

2.3 봇-임상 통합의 프라이버시 역설

덜 자주 명확히 표현되는 두 번째 격차는 대화형 에이전트와 임상 평가 모듈의 통합에 관한 것이다. 생성형 AI 챗봇은 엄격한 시험의 첫 번째 물결에서 임상적 효능을 입증했으며(Heinz et al., 2025; Wang et al.,

2025; 체계적 검토는 Reyes-Portillo et al., 2025 참조), 현대 플랫폼은 대화형 프론트엔드와 구조화된 임상 백엔드를 모두 점점 더 배포하고 있다. 대화 스트림은 상당한 임상적 가치를 갖는 정보—기분 궤적, 생활 사건, 순응도 보고, 증상 변동—를 포함하지만, 명시적 사용자 통제 없이 이 스트림을 임상 기록으로 라우팅하는 것은 일반 데이터 보호 규정(GDPR) 9 조(특수 범주 데이터)와 22 조(자동화된 개별 의사결정) 그리고 미국의 건강보험 이동성 및 책임법(HIPAA)에 따른 상당한 윤리적·규제적 위험을 초래한다 (Cohen & Mello, 2018; Voigt & von dem Bussche, 2017; Wachter et al., 2017).

우리는 이를 봇-임상 통합의 프라이버시 역설로 프레임링한다: 봇 파생 임상 신호는 가치 있으나, 동의되지 않은 공유는 대화형 공개가 기반을 두는 신뢰 전제를 훼손한다. 우리가 제안하는 아키텍처적 해결책은 대화형 모듈(TalkCatcher)과 임상 평가 모듈(NeuroCatchers) 사이의 명시적 오프인 기본 OFF 브리지가—아래에서 ALG-4 로 상세히 기술된다.

2.4 임상 및 규제 요구사항으로서의 감사 추적성

최근 규제 궤적은 공유된 요구사항으로 수렴한다: AI 지원 임상 소프트웨어는 추천에 대해 투명하고 추적 가능한 추론을 제공해야 한다. 미국 식품의약국의 AI/ML 의료기기 소프트웨어(SaMD) 행동 계획(FDA, 2021; Warraich et al., 2024)은 우수한 머신러닝 관행과 수명주기 감사를 강조한다. 2024 년 발효된 유럽연합의 AI 법은 의료에 사용되는 AI 시스템을 고위험으로 분류하고 투명성, 인간 감독, 시판 후 모니터링 요구사항을 부과한다. GDPR 22 조는 데이터 주체에게 법적 또는 유사하게 중대한 영향을 미치는 전적으로 자동화된 의사결정의 대상이 되지 않을 권리를 부여하며, 부수적인 "설명에 대한 권리"의 범위에 대한 해석 논쟁이 동반된다(Goodman & Flaxman, 2017; Wachter et al., 2017). 임상적으로, 감사 추적 가능한 추론은 치료 표준 원칙에 따른 전문가 책임을 지원하고, 회고적 품질 개선을 가능하게 하며, 임상가가 알고리즘 추천을 실질적 근거에 따라 도전하거나 수용할 수 있도록 한다(Doshi-Velez & Kim, 2017; Ghassemi et al., 2021; Tjoa & Guan, 2021).

따라서 우리의 아키텍처는 감사 추적성을 사후 추가가 아닌 일급 설계 제약으로 다룬다. 엔진이 생성하는 모든 추천은 점수 구성요소, 근거 인용(가능한 경우 PubMed 식별자 또는 디지털 객체 식별자 포함), 한국어와 영어 양쪽의 추론 텍스트, 그리고 이를 생성한 모델 버전을 영구 저장한다. 12 절에서 감사 스키마를 상세화한다.

2.5 기여 및 논문 구성

본 논문은 다음과 같은 기여를 한다. 첫째, 우리는 평가-치료 경로, 프라이버시 보존 봇-임상 브리지, 맞춤형 훈련 저작, 이중 언어 결과 렌더링을 집합적으로 운영화하는 6 개 알고리즘 엔진—ALG-1 부터 ALG-6 까지—을 명세한다. 둘째, 우리는 이러한 알고리즘을 뒷받침하는 관계형 데이터베이스 스키마를 감사 추적에 명시적 주의를 기울여 제공한다. 셋째, 우리는 보편적 증상 영역 아키텍처를 명확히 한다: 10 개 증상 트랙(ADHD 를 황금 템플릿으로, 우울증, 불안, 수면, 번아웃, 외상후스트레스장애, 강박장애, 양극성장애, 자폐, 최고 수행을 병렬 인스턴스로)에 걸쳐 매개변수화된 단일 스키마와 알고리즘 세트이다.

넷째, 우리는 우리의 설계를 현재 규제 프레임워크에 매핑하고 FDA SaMD 클래스 II 및 GDPR 9 조/22 조에 따른 적합성 경로를 식별한다.

논문의 나머지는 다음과 같이 구성된다. 3 절은 개인화, 적응형 치료, 의료 추천 시스템, Fischer 역동적 기술 이론, 프라이버시 보존 건강 데이터 아키텍처, 임상 AI에서의 감사 추적의 관련 작업을 검토한다. 4 절은 이중 모듈 시스템 아키텍처를 기술한다. 5 절은 데이터베이스 스키마를 제시한다. 6 절부터 11 절은 6 개 알고리즘을 차례로 명세한다. 12 절은 감사 추적성 계층을 상세화한다. 13 절은 보편적 증상 아키텍처와 ADHD 황금 템플릿 전략을 기술한다. 14 절은 강점, 한계, 임상적 함의, 규제 경로를 논의한다. 15 절은 결론을 맺는다.

3. 배경 및 관련 작업

3.1 디지털 정신건강에서의 개인화

Hornstein 등(2023)은 콘텐츠, 제시, 타이밍, 피드백 차원을 구분하면서 DMHI의 개인화 전략을 분류하기 위한 가장 영향력 있는 최근 프레임워크를 제공한다. Wanniarachchi 등(2025)은 이 프레임워크를 청소년 집단으로 확장하고 각 차원의 유병률을 문서화한다. 이러한 검토에서 일관된 발견은 적시 적응 로직을 구현하는 중재가 적으며, 개인화의 알고리즘적 기반을 사용자나 임상가에게 노출하는 경우는 훨씬 적다는 것이다. 우리의 작업은 영구 저장된 추론을 갖춘 명시적 알고리즘 프로세스로 개인화를 다룸으로써 이 격차를 해결한다.

3.2 정신건강을 위한 대화형 AI

Heinz 등(2025)은 주요우울장애, 범불안장애, 임상적으로 고위험인 섭식장애 전반에 걸쳐 N=210명 성인을 대상으로 임상 수준 정신건강 증상에 대한 미세 조정된 생성형 AI 챗봇(Therabot)의 첫 번째 무작위 대조 시험을 보고했다. 효과 크기는 컸고 참여도는 높았다(4주에 걸쳐 사용자당 평균 260개 메시지). Wang 등(2025)은 중국 대학생의 우울증에 대한 인지행동치료 챗봇으로 증거 기반을 확장했다. Reyes-Portillo 등(2025)은 메타분석적 종합을 제공한다. 이러한 시험은 대화형 AI의 임상적 실행 가능성을 확립하지만, 대화 내용이 구조화된 임상 기록으로 어떻게 안전하게 연결되어야 하는지 명세하지 않는다—이는 우리의 ALG-4가 직접 다루는 격차이다.

3.3 의료 분야의 추천 시스템

Lee 등(2024)은 간세포암에 대한 순위가 매겨진 대안과 생존 예측을 갖춘 치료 추천을 생성하는 머신러닝 CDSS의 최근 npj Digital Medicine 예시를 제공한다. Sutton 등(2020)은 더 광범위한 CDSS 환경을 검토하면서 이점(의사결정 지원, 오류 감소)과 위험(경고 피로, 자동화 편향)을 식별한다. Wiens 등(2019)은 임상 맥락에서 책임 있는 ML을 위한 "해를 끼치지 말라" 로드맵을 명확히 한다. 우리의 추천 엔진(ALG-1)은 다중 구성요소 점수 산정 전통을 계승하면서 정신건강을 위해 보정된 명시적 발달 및 선호 구성요소로 확장한다.

3.4 적응형 치료 전략

Murphy(2005)의 SMART(순차 다중 할당 무작위 시험) 프레임워크는 적응형 치료의 다단계 의사결정 문제를 공식화한다. Almirall 등(2014)은 체중 감량을 포함한 행동 중재로 SMART를 확장하고, Nahum-Shani 등(2018)은 모바일 건강을 위한 적시 적응 중재를 정교화한다. SMART 시험이 적응 전략을 발견하기 위한 경험적 방법론을 명세하지만, 운영 플랫폼에서 그러한 전략을 운영화하기 위한 데이터 아키텍처를 처방하지는 않는다. ALG-2와 ALG-3은 함께 환자 측 변경에 의해 트리거되는 재계획과 세션 수준 계획을 통합하는 그러한 운영화 중 하나를 제공한다.

3.5 Fischer의 역동적 기술 이론

Fischer(1980)의 역동적 기술 이론은 Fischer와 Bidell(2006) 및 Mascolo와 Fischer(2015)에서 정교화되어, 기술이 4개 계층(반사, 감각운동 행동, 표상, 추상화)에 걸쳐 계층적 순서로 발달하며, 계층당 4개 수준이 발달 척도를 산출한다고 주장한다. 기술은 일반 능력이 아닌 맥락화된 통제 구조로 개념화되며 기술 수행은 맥락적 지원의 함수로서 변동한다. 우리는 Fischer의 프레임워크를 ALG-1에서 `fischer_compatibility` 구성요소(훈련 진입 수준을 추정된 현재 수준에 매칭)로, ALG-2에서 세션별 진행 일정으로 운영화한다. 우리가 아는 한, 이는 배포된 정신건강 플랫폼에서 Fischer 프레임워크의 첫 번째 명시적 계산적 사용이다.

3.6 프라이버시 보존 건강 데이터 아키텍처

Voigt와 von dem Bussche(2017)는 9조(건강을 포함한 특수 범주 데이터)와 22조(자동화된 의사결정)에 특히 중점을 두면서 GDPR 준수에 대한 표준 실용 가이드를 제공한다. Cohen과 Mello(2018)는 현대 데이터 흐름에 비추어 HIPAA 현대화를 검토한다. 우리가 2.3절에서 명확히 한 프라이버시 역설은 특수 범주 데이터 처리 전 명시적 동의에 대한 GDPR 9조(2)(a) 요구사항에 매핑되며, 이는 ALG-4의 옵트인 기본 OFF 아키텍처의 법적 근거이다.

3.7 임상 AI에서의 설명 가능성과 감사

Doshi-Velez와 Kim(2017)은 해석 가능한 ML의 과학적 기초를 명확히 한다; Tjoa와 Guan(2021)은 명시적 의료 응용을 갖춘 설명 가능한 AI 방법을 조사한다; Ghassemi 등(2021)은 의료 분야에서 사후 설명에 대한 과도한 의존에 대해 경고하며, 처음부터 투명한 설계를 주장한다. Wachter 등(2017)은 GDPR에서 파생된 "설명에 대한 권리"의 논쟁의 여지가 있는 범위를 검토하며, Goodman과 Flaxman(2017)은 더 광범위한 해석을 제공한다. 우리의 감사 추적성 계층(12절)은 사후 설명 방법에 의존하기보다는 각 추천에 대한 실질적 입력을 영구 저장함으로써 설계상 투명한 입장을 채택한다.

4. 시스템 개요

Figure 1. Two-Module Architecture with Six-Algorithm Integration

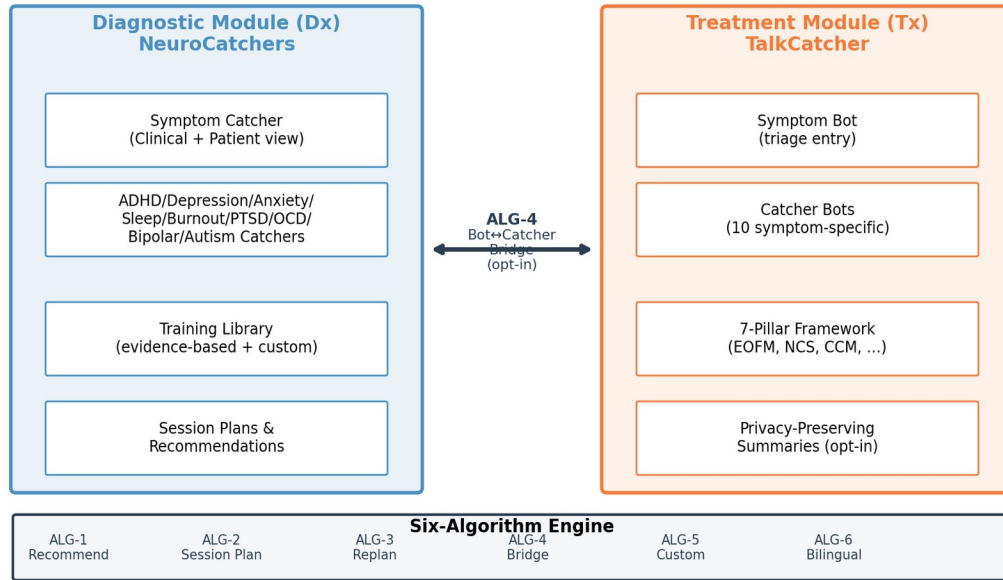


그림 1. 이중 모듈 시스템 아키텍처: NeuroCatchers(진단)와 TalkCatcher(치료)가 ALG-4 프라이버시 보존 브리지로 연결됨.

4.1 이중 모듈 아키텍처: 진단 및 치료

플랫폼은 두 개의 상보적 모듈로 조직된다(그림 1). NeuroCatchers 로 브랜딩된 진단 모듈(Dx)은 최상위 증상 캐처와 10 개 증상별 캐처(ADHD, 우울증, 불안, 수면, 번아웃, 외상후스트레스장애, 강박장애, 양극성장애, 자폐, 최고 수행)를 통해 구조화된 임상 평가를 제공한다. 각 캐처는 근거 기반 훈련 라이브러리를 포함하고 세션 계획과 추천을 노출한다. TalkCatcher 로 브랜딩된 치료 모듈(Tx)은 증상 봇(분류 진입)과 임상 증상 영역에 해당하는 10 개 캐처 봇을 통해 대화형 지원을 제공한다. 대화 모듈은 정서 전용 피드백 모델(EOFM), 서사 공동 구성 시스템(NCS), 임상 보정 모듈(CCM) 및 본 논문에서 중심적이지 않은 4 개의 추가 기둥을 포함하는 Boston Neuromind 의 7 기둥 프레임워크를 인스턴스화한다.

두 모듈은 환자가 명시적으로 옵트인한 경우에만 봇 파생 임상 신호를 진단 기록으로 동기화하는 프라이버시 보존 브리지인 ALG-4 로 연결된다. 이 분리는 2.3 절에서 명확히 한 프라이버시 역설을 반영한다.

4.2 6 개 알고리즘 엔진

6 개 알고리즘은 평가-치료 경로와 모듈 간 브리지를 집합적으로 운영화한다. ALG-1(추천 엔진)은 환자 프로파일과 후보 훈련을 입력으로 받아 추천이 포함된 순위 추천을 생성한다. ALG-2(세션 계획 생성자)는

선택된 훈련을 받아 Fischer 수준 진행이 포함된 N 개 세션 계획을 생성한다. ALG-3(적응형 재계획)은 트리거 이벤트에 대응하여 계획을 업데이트한다. ALG-4(봇 ↔ 캐처 브리지)는 프라이버시 보존 동기화를 관장한다. ALG-5(맞춤형 훈련 빌더)는 임상가나 환자가 보편적 스키마에 부합하는 새로운 훈련 기록을 저장할 수 있게 한다. ALG-6(이중 언어 결과 렌더러)은 진단 출력의 한국어 및 영어 버전을 생성한다.

4.3 보편적 증상 영역 구조

핵심 아키텍처 약속은 증상 영역 전반의 보편성이다. 단일 데이터베이스 테이블 세트, 단일 알고리즘 세트, 단일 렌더링 파이프라인이 10 개 캐처 모두를 담당한다. 증상별 콘텐츠는 코드가 아닌 데이터이다: `training_library` 테이블은 주어진 캐처에 대해 어떤 훈련이 보이는지 선택하는 `symptom_track` 열을 포함하고; `contraindications` 필드는 증상별 안전 규칙을 인코딩하며; `quick_input_schema` 는 증상에 적합한 측정 도구를 명세한다. 우리는 ADHD 를 초기 개발 및 검증을 위한 황금 템플릿으로 채택하며, 9 개의 후속 증상 영역이 코드 작성이 아닌 데이터 채우기로 인스턴스화되도록 하는 명시적 의도가 있다.

4.4 프라이버시 우선 설계 원칙

5 가지 설계 원칙이 아키텍처를 관장한다. 첫째, 옵트인 기본 OFF 의미는 모든 모듈 간 데이터 흐름에 적용된다. 둘째, 동의는 세밀하고 철회 가능하며, GDPR 7 조와 일치한다. 셋째, 감사 추적성은 모든 추천의 일급 출력이다. 넷째, 이중 언어 렌더링은 한국어 및 영어 사용자 모두를 위한 진단 출력에 필수이다. 다섯째, 아키텍처는 자율적 환자 사용과 임상가 감독 사용을 모두 수용해야 하며, 동일한 데이터 구조가 두 모드를 모두 지원해야 한다.

5. 데이터베이스 스키마 설계

Figure 2. Core Database Schema — Entity Relationship Diagram

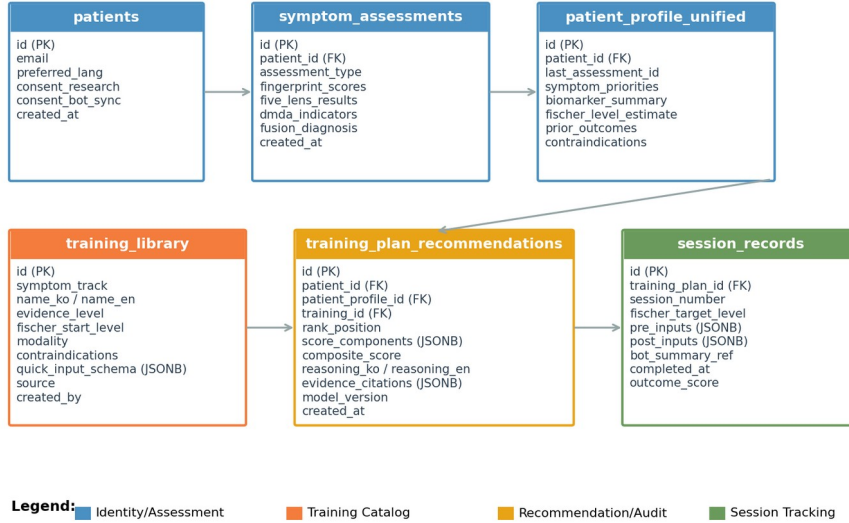


그림 2. 6 개 핵심 테이블과 그 관계를 보여주는 엔티티-관계 다이어그램.

표 1. 핵심 데이터베이스 테이블과 관계.

Table	Primary Purpose	Key Fields	Used By
patients	Identity and consent flags	id, language_pref, consent_research, consent_bot_sync	All algorithms
symptom_assessments	Raw assessment outputs	id, patient_id, fingerprint_jsonb, dmda_jsonb	ALG-6
patient_profile_unified	Synthesized profile for recommendation	symptom_priorities, biomarker_summary, fischer_level_estimate	ALG-1
training_library	Curated intervention catalog	id, symptom_track, evidence_level, fischer_start_level, contraindications, quick_input_schema	ALG-1, ALG-2, ALG-5
training_plan_recommendations	Persisted recommendation audit	score_components, composite_score, reasoning_ko, reasoning_en, evidence_citations, model_version	ALG-1, audit layer

Table	Primary Purpose	Key Fields	Used By
session_records	Per-session tracking	pre_input, post_input, bot_summary, completion_status	ALG-2, ALG-3, ALG-4

5.1 핵심 엔터티와 관계

스키마는 6 개의 핵심 테이블로 구성된다(그림 2). patients 테이블은 정체성, 언어 선호도, consent_research 및 consent_bot_sync 를 포함한 동의 플래그를 보유한다. symptom_assessments 테이블은 임상 도구의 결과를 저장하며, Fingerprint Engine 출력, Five-Lens 점수, 감별 정신장애 알고리즘(DMDA) 지표, 융합된 진단 프로파일을 위한 JSONB 열로 구성된다. patient_profile_unified 테이블은 가장 최근의 평가를 합성된 증상 우선순위 벡터, 바이오마커 요약, 추정된 Fischer 기술 수준, 금기 플래그와 통합한다; 이 테이블은 ALG-1 에 대한 주요 입력이다.

training_library 테이블은 근거 기반 중재를 카탈로그화한다(6 절). training_plan_recommendations 테이블은 모든 ALG-1 출력을 전체 점수 분해 및 추론과 함께 영구 저장한다. session_records 테이블은 세션 전후 빠른 입력 및 선택적 봇 요약 참조를 포함하여 세션별 활동을 추적한다.

5.2 추천 입력으로서 patient_profile_unified 테이블

중심적 설계 선택은 이질적 평가 출력을 추천에 대한 입력 역할을 하는 단일 통합 프로파일로 통합하는 것이다. 통합 프로파일은 다음을 포함한다: symptom_priorities(증상 영역 간 상대적 가중치를 나타내는 정규화 벡터), biomarker_summary(뇌전도 또는 기타 생리학 데이터가 가용한 경우), fischer_level_estimate(환자의 추정된 현재 발달 수준), prior_outcomes(과거 참여, 중도 이탈, 결과 데이터), contraindications(특정 중재를 차단하는 플래그 세트). 이 통합은 ALG-1 이 원시 평가 데이터를 탐색하기보다 안정적이고 잘 정의된 입력에 대해 작동할 수 있게 한다.

5.3 training_library 테이블: 근거 등급 및 스키마

training_library 의 각 행은 하나의 치료 중재를 나타낸다. 핵심 열에는 다음이 포함된다: id, symptom_track(어떤 캐처가 이 훈련을 노출하는지), name_ko 및 name_en, evidence_level(rct, meta_analysis, consensus_guideline, expert_opinion, custom 중 하나), fischer_start_level(발달 진입점), modality(인지적, 행동적, 신체적, 뉴로피드백, 노출, 마음챙김, 심리교육), contraindications(JSONB), quick_input_schema(세션별 측정 도구를 명세하는 JSONB), source(어떤 임상가 또는 조직이 항목을 큐레이션했는지). 이중 evidence_level + custom 설계는 큐레이션되고 검증된 항목이 추천 엔진의 근거 강도 점수 산정을 손상시키지 않으면서 임상가 저작 맞춤형 훈련과 공존할 수 있도록 한다.

5.4 추천 추론 영구 저장

모든 ALG-1 호출은 하나 이상의 행을 `training_plan_recommendations`에 기록한다. 핵심 필드에는 다음이 포함된다: `score_components`(하위 구성요소 및 방법과 함께 6개 구성요소 점수를 모두 포함하는 JSONB), `composite_score`(최종 가중 합계), `reasoning_ko` 및 `reasoning_en`(양 언어로 된 사람이 읽을 수 있는 설명), `evidence_citations`(이를 지원하는 특정 주장과 함께 PubMed 식별자 또는 디지털 객체 식별자의 JSONB 배열), `contraindications_considered`(어떤 환자 플래그가 평가되었고 어떻게), `model_version`(이 추천을 생성한 점수 산정 모델의 식별자). `model_version` 필드는 종단 감사에 필수적이다: 엔진이 학습되고 업데이트됨에 따라 모든 과거 추천은 이를 생성한 모델의 맥락에서 해석 가능하게 유지된다.

5.5 봇-캐처 동기화 스키마

`consent_bot_sync`가 TRUE일 때, ALG-4는 세션별 요약을 `session_records.bot_summary`에 기록한다. 이는 고정 스키마를 갖는 JSONB 필드이다: `mood_trend`(구성 가능한 창에 대한 이동 평균), `reported_symptoms`(대화에서 추출된 임상적으로 관련된 신호의 배열), `life_events`(이사나 직업 변경과 같이 사용자가 공개한 중요한 사건), `adherence_self_report`(처방된 활동에 대한 사용자의 명시된 참여). 스키마는 임상 의사결정을 지원하면서 불필요한 데이터 흐름을 최소화하기 위해 의도적으로 좁다.

6. 알고리즘 1: 추천 엔진

6.1 입력 및 하드 필터

ALG-1은 `patient_profile_unified`의 행과 `symptom_track`이 활성 캐처와 일치하는 `training_library` 항목의 부분집합을 입력으로 받는다. 점수 산정 전, 하드 필터가 적용된다: `contraindications` 배열이 환자의 금기 플래그와 교차하는 모든 훈련은 후보 집합에서 제거된다. 하드 필터는 가중되지 않으며; 특정 환자에게 임상적으로 부적절한 훈련을 범주적으로 배제한다. 예에는 약물 상호작용(환자가 금기 심장 상태를 플래그할 때 자극제 기반 프로토콜이 필터링됨), 동반이환 배제(활동성 정신증이 플래그될 때 OCD에 대한 노출 프로토콜이 필터링됨), 발달 하한선(환자의 추정 수준이 훈련의 `fischer_start_level`에서 허용 격차를 뺀 값보다 낮을 때 높은 Fischer 수준 훈련이 필터링됨)이 포함된다.

6.2 6개 구성요소 점수 산정 함수

살아남은 후보는 6개 구성요소에 걸쳐 점수가 매겨진다(그림 3, 표 2). `symptom_match` 구성요소는 환자의 `symptom_priorities` 벡터와 훈련의 `symptom_target_vector` 사이의 가중 코사인 유사도를 계산한다. `biomarker_summary`가 채워진 경우 적용 가능한 `biomarker_match` 구성요소는 환자 바이오마커 서명과 훈련의 `biomarker_target_pattern` 사이의 Z 점수 가중 정렬을 계산한다. `evidence_strength` 구성요소는 훈련의 원천 근거에서 Cochrane d(또는 유사한 효과 크기)를 정규화한다. 메타분석적 효과 크기는 단일 시험 추정치보다 높은 가중치를 받으며, 합의 지침 훈련은 보정된 기준선을 받는다.

fischer_compatibility 구성요소는 환자의 fischer_level_estimate 와 훈련의 fischer_start_level 사이의 일치를 점수화하며, 정확한 일치에서 최대 점수와 격차에 대한 점진적 감소를 보인다. practical_fit 구성요소는 양상 선호도, 시간 가용성 제약, 비용 요인을 통합한다. personal_history 구성요소는 이전 성공과 관련된 훈련에 보상을 주고 이전 중도 이탈과 관련된 훈련에 페널티를 부여한다. 각 구성요소는 문서화된 도출 방법으로 단위 구간 [0, 1]의 점수를 산출한다.

6.3 종합 점수 및 순위 매기기

composite_score 는 6 개 구성요소 점수의 가중 합으로 계산된다. 기본 가중치는 임상가 상담에서 도출되며 증상 트랙별로 조정될 수 있다. 후보는 내림차순 composite_score 로 순위가 매겨지며, 상위 3 개는 주요 추천으로 제시되고; 4 번째부터 10 번째는 대안으로 유지되어 사용자 요청 시 노출된다. 종합 가중치 자체는 감사를 지원하기 위해 model_version 메타데이터에 저장된다.

6.4 추론 생성

각 순위 후보에 대해 ALG-1 은 템플릿 + 대규모 언어 모델 다듬기 파이프라인을 통해 한국어 및 영어 추론 텍스트를 생성한다. 템플릿은 임상가가 작성하며 환자의 두드러진 특징에 대한 자리표시자를 포함한다(예: "주의력 조절의 보고된 우선순위와 sm3 의 추정 발달 수준을 고려하면..."). LLM 다듬기 단계는 템플릿의 사실적 내용을 보존하면서 어조 일관성과 자연스러운 표현을 보장한다. 생성된 추론은 근거 인용을 절대 조작하지 않도록 제약된다; 원천 훈련의 evidence_citations 필드에 존재했던 PMID 와 DOI 만이 렌더링된 추론에 나타날 수 있다.

6.5 콜드 스타트 처리

환자가 이전 이력이 없는 경우(콜드 스타트), personal_history 구성요소는 기본적으로 중립 0.5 이며 이 구성요소에 대한 종합 가중치는 나머지 5 개 구성요소에 일시적으로 재분배된다. 마찬가지로 biomarker_summary 가 비어 있는 경우 biomarker_match 가 생략되고 그 가중치가 재분배된다. 재분배는 score_components 에 기록되어 코호트 간 비교 분석이 각 추천에서 가용한 데이터를 올바르게 고려할 수 있도록 한다.

6.6 예시: ADHD 추천 출력

표 3 은 경미한 부주의 증상, 바이오마커 데이터 없음, sm3 의 추정 Fischer 수준을 가진 성인 ADHD 환자에 대한 샘플 출력을 보여준다. 1 순위 추천은 composite_score 0.89 인 집중 주의 뉴로피드백 프로토콜로, Arns 등(2009)과 Van Doren 등(2019)이 지원한다. 2 순위는 포모도로 기반 행동 프로토콜 (composite_score 0.84)이다. 3 순위는 마음챙김 기반 주의 훈련(composite_score 0.81)이다. 각각은 한국어 및 영어 추론 텍스트와 관련 근거 인용을 동반한다.

7. 알고리즘 2: 세션 계획 생성자

7.1 입력 및 Fischer 수준 진행

ALG-2 는 선택된 `training_library` 행, 환자나 임상가가 선택한 `N_sessions` 수, 환자의 현재 `fischer_level_estimate` 를 입력으로 받는다. Fischer 수준이 훈련의 시작 수준에서 환자의 발달 궤적과 훈련의 구조에 적합한 종료 수준으로 진행되는 N 개 세션 명세의 순서화된 시퀀스를 생성한다(그림 4, 표 4).

진행 알고리즘은 Fischer 의 계층 경계를 존중한다: 계층 내 전환(감각운동, 표상, 추상)은 계층 횡단 전환보다 빠르게 발생하며, 후자는 추가 통합 세션을 필요로 한다. 알고리즘은 기술 구성이 엄격히 선형적이지 않고 반복적이라는 Fischer 와 Bidell(2006)의 특성화에 의존한다: 주어진 수준에서의 연습은 진행 전에 기술을 통합하고 안정화한다.

7.2 세션별 콘텐츠 생성

각 세션에 대해 ALG-2 는 목표 Fischer 수준, 목표 기술, 활동 설명, 예상 지속 시간, 환자가 세션 전후에 완료할 `quick_input_schema` 인스턴스화를 포함하는 콘텐츠 명세를 생성한다. `quick_input_schema` 는 `training_library` 행에서 상속되지만 세션의 특정 초점에 맞게 매개변수화된다. 예를 들어, ADHD 집중 주의 훈련은 세션 전 집중도 평가에 0-10 슬라이더, 세션 동안 완료된 작업에 정수 카운트, 세션 후 집중도 평가에 0-10 슬라이더를 사용할 수 있다; 이러한 레이블은 환자 선호에 따라 한국어 또는 영어로 현지화된다.

7.3 30 초 세션 입력 설계

핵심 설계 제약은 세션별 데이터 입력이 30 초 이내에 완료 가능해야 한다는 것이다. 입력 부담이 이 임계값을 초과하면 디지털 중재 순응도가 빠르게 저하된다(Baumel et al., 2019; Fleming et al., 2018). 따라서 `quick_input_schema` 는 세션당 최대 4 개 필드로 제한되며, 각 필드는 4 개 원시 입력 유형(슬라이더, 정수, 부울, 시간) 중 하나를 사용한다. 자유 텍스트 입력은 선택적이며 계획 계속을 위해 결코 필수가 아니다.

8. 알고리즘 3: 적응형 재계획

8.1 트리거 조건

ALG-3 은 세 가지 트리거 클래스에 대응하여 발화한다(그림 5). 첫째, 세션 수를 변경하거나, 주요 목표를 수정하거나, ALG-1 이 순위 매긴 대안에서 다른 훈련을 선택하는 명시적 사용자 또는 임상가 행동. 둘째, 중도 이탈 신호: 구성 가능한 임계값을 초과하는 결석 세션(기본: 연속 2 회 결석), 또는 구성 가능한 허용 범위를 초과하는 세션 전후 결과 점수의 악화. 셋째, 환자의 업데이트된 프로파일에 대해 계획을 재평가하는 구성 가능한 주기에서의 정기 검토(기본: 6 세션마다).

8.2 계획 diff 및 이력 보존

ALG-3 이 발화할 때, 기존 계획을 제자리에서 수정하지 않는다; 대신 이전 계획과 새 계획 사이의 diff와 함께 plan_modification 행을 기록한 다음 새 계획을 작성한다. 이는 환자의 치료 궤적의 완전한 이력을 보존하여 임상적 성찰과 종단 분석 모두를 지원한다. 과거 계획은 무기한 쿼리 가능하게 유지된다.

8.3 추론 재생성

각 재계획 호출은 ALG-1 이 사용하는 동일한 템플릿 + LLM 다듬기 파이프라인을 통해 이중 언어 추론 텍스트를 재생성하며, 명시적 "무엇이 변경되었고 왜" 서사를 추가한다. 서사는 트리거(예: "환자 요청에 따라 10 에서 12 세션으로 연장됨" 또는 "이전 계획의 연속 2 회 결석으로 인해 대안 훈련 선택됨")와 관련 업데이트된 프로필 구성요소를 인용한다.

9. 알고리즘 4: 프라이버시 보존 봇 ↔ 캐치 브리지

9.1 통합 플랫폼의 프라이버시 역설

2.3 절에서 명확히 한 바와 같이, 대화형 에이전트와 임상 평가 모듈의 통합은 프라이버시 역설에 직면한다. 봇 파생 임상 신호는 가치 있으나, 동의되지 않은 공유는 대화형 공개가 기반을 두는 신뢰 전제를 훼손한다. 완전한 통합(높은 임상 효용, 낮은 프라이버시 보호)과 완전한 분리(높은 프라이버시 보호, 낮은 임상 효용) 사이의 선택은 잘못된 이분법이다. ALG-4 는 세 번째 옵션을 구현한다: 기본 OFF 기준선을 갖춘 명시적이고, 세밀하며, 철회 가능한 옵트인.

9.2 옵트인 아키텍처

patients.consent_bot_sync 플래그는 동기화를 권장한다(그림 6). 계정 생성 시 기본값은 FALSE 이다. 플래그는 프라이버시 설정 인터페이스를 통해 환자가 언제든지 토글할 수 있다; FALSE 로 토글하면 추가 동기화는 즉시 중단되지만 이전에 동기화된 요약물 소급하여 삭제하지는 않는다(이는 GDPR 17 조 및 HIPAA 접근권 조항에 따른 표준 데이터 삭제 요청 대상이다). 임상가가 환자 치료에 관여할 때, 임상가는 환자에게 동기화를 활성화하도록 요청할 수 있으나, 환자는 GDPR 9 조 (2)(a)의 특수 범주 데이터에 대한 명시적 동의 요구사항과 일치하여 플래그 값에 대한 단독 권한을 유지한다.

9.3 요약 추출 파이프라인

동기화가 활성화되면, 봇의 대화 이력은 세션 종료 시(또는 일괄 운영의 경우 매일) 큐레이션된 임상 어휘에 고정된 검색 증강 생성 파이프라인을 통해 처리된다. 파이프라인은 bot_summary 스키마(5.5 절)와 일치하는 콘텐츠를 추출하고 이 스키마 외부의 콘텐츠는 폐기한다. 결정적으로, 원시 대화 로그는 진단 기록에 결코 기록되지 않는다; 스키마 부합 추출만 전송된다. 이는 GDPR 5 조 (1)(c) 데이터 최소화와 일치하여 데이터 흐름을 최소화한다.

9.4 치료사 가시성 통제

진단 기록 내에서, bot_summary 필드는 표준 치료사 접근 통제 대상이다. 환자는 치료사 보기에서 특정 bot_summary 필드를 추가로 제한할 수 있다(예: mood_trend 를 허용하면서 life_events 숨기기); 이러한 필드 수준 제한은 session_records 행의 access_policy JSONB 열에 기록된다. bot_summary 에 대한 모든 접근은 감사를 위해 기록된다.

10. 알고리즘 5: 맞춤형 훈련 빌더

10.1 임상가 및 내담자 입력 UI

임상가는 개별 환자를 위한 맞춤형 프로토콜을 자주 개발한다(예: 특정 OCD 의식에 대한 점진적 노출 계획 실행 기능 지원을 위한 개인화된 아침 루틴). ALG-5 는 8 개 필드 UI 를 통해 그러한 훈련 저작을 지원한다(그림 7): 증상 트랙, 목표 행동, 범주, 두 개의 측정 명세(각각 제약된 드롭다운에서 선택), 권장 세션, 세션당 분, 가시성(임상가 전용 대 비공개). 자유 형식 JSON 저작은 명시적으로 금지된다; 제약된 UI 는 맞춤형 훈련이 보편적 스키마에 부합하고 ALG-1 점수 산정과 호환되도록 보장한다.

10.2 AI 지원 다듬기

선택적 "AI 로 다듬기" 작업은 다음을 위해 언어 모델을 호출한다: 임상적으로 적절한 설명 생성, 증상 트랙의 표준 금기 라이브러리에서 도출된 금기 제안, 저자의 참조를 위한 유사한 기존 훈련 노출. 저자는 AI 생성 콘텐츠에 대한 완전한 편집 통제권을 유지한다. 모든 AI 생성 텍스트는 training_library 행의 source 필드에 플래그된다.

10.3 검증 및 저장

삽입 전, 검증은 다음을 확인한다: (a) 모든 필수 필드가 채워졌는지, (b) contraindications 가 증상 트랙의 필수 포함과 충돌하지 않는지, (c) quick_input_schema 가 유효한 입력 유형을 산출하는지. 맞춤형 훈련은 evidence_level = 'custom'으로 저장되며 source 는 저작 임상가 또는 내담자를 기록한다. 그것들은 메타분석적으로 지원되는 훈련보다 낮은 보정된 evidence_strength 기준선으로 ALG-1 점수 산정 자격이 있다.

11. 알고리즘 6: 이중 언어 결과 렌더러

11.1 진단 결과를 위한 두 가지 출력 경로

증상 캐치 출력은 두 가지 병렬 버전으로 렌더링된다: DSM/ICD 코드, Z 점수, 신경학적 해석을 사용하는 임상 버전과 일상 언어, 은유, 격려를 사용하는 환자 버전(그림 8). 대조적으로, 증상 봇 출력은 단일 통합 버전으로 렌더링된다: 봇의 의사소통 맥락은 임상/환자 구분을 덜 의미 있게 만들며, 봇의 어조 보정은 이미 ALG-6 의 매개변수에 의해 제약된다.

11.2 템플릿 + LLM 다듬기 파이프라인

두 경로 모두 ALG-1의 추론 생성과 유사한 템플릿 + LLM 다듬기 파이프라인을 사용한다. 템플릿은 구조적 요소를 인코딩한다(어떤 발견이 어떤 순서로 어떤 강조로 보고되는지); LLM 다듬기 단계는 어조를 보정하고 각 언어의 자연스러운 표현을 보장한다. 어조 보정 매개변수는 출력에 따라 다르다: 임상 버전은 정확하고 전문적인 어조를 사용하며; 환자 버전은 따뜻하고 격려하는 어조를 사용하고; 붓 버전은 접근성과 진단 본질의 균형을 맞춘다.

11.3 한국어-영어 병렬 렌더링

두 언어는 사후 번역이 아닌 병렬로 렌더링된다. 이는 사후 번역의 두 가지 실패 모드를 피한다: 관용구 보존(한국어 관용구는 영어 등가물로 직접 번역되지 않으며, 그 반대도 마찬가지) 및 임상 용어 등가성(한국어 임상 용어는 때때로 직접적인 영어 유사물이 부족하고 설명적 렌더링을 필요로 함). 템플릿 엔진은 각 언어에 대한 별도 템플릿을 유지하며, 데이터 입력만 공유한다.

11.4 문화적 보정

Hwang 등(2008)에 따르면, 문화적 맥락은 증상 제시와 치료 추천의 수용 가능성 모두를 형성한다. 한국어 렌더링은 임상적 만남에 적절한 존칭, 문화적으로 표시될 때 심리적 고통의 신체적 프레이밍, 개인 프레이밍 언어가 문화적으로 부조화할 수 있는 곳에서 가족 체계 프레이밍을 사용한다. 이러한 적응은 사후에 도입되지 않고 템플릿에 인코딩된다.

12. 감사 추적성

표 6. *training_plan_recommendations*의 감사 필드 스키마.

Field	Type	Purpose
score_components	JSONB	All 6 component scores + sub-components + methods
composite_score	Numeric (0-1)	Final weighted total
reasoning_ko	Text (immutable)	Korean human-readable explanation
reasoning_en	Text (immutable)	English human-readable explanation
evidence_citations	JSONB array	PMID/DOI with claim and component link
contraindications_considered	JSONB	Which patient flags evaluated and outcome
model_version	Text	Identifier of scoring model that produced recommendation
created_at	Timestamp	Creation time

Field	Type	Purpose
patient_id	UUID	Foreign key to patients

표 7. 감사 설계의 규제 매핑.

Regulation	Requirement	How Audit Design Addresses It
FDA SaMD Good Machine Learning Practice	Transparency, lifecycle audit	score_components + model_version persistence
FDA SaMD Post-Market Monitoring	Tracking real-world performance	training_plan_recommendations + session_records linked
GDPR Article 9 (Special Category Data)	Explicit consent for health data	consent_bot_sync flag; opt-in default-off in ALG-4
GDPR Article 22 (Automated Decision-Making)	Right not to be subject to solely automated decisions	Clinician-in-the-loop design; reasoning_ko/en for meaningful review
GDPR Article 5(1)(c) Data Minimization	Limit data flow to necessary	Fixed bot_summary schema; raw conversation never transferred
EU AI Act High-Risk System	Transparency, human oversight, post-market monitoring	Audit fields + clinician collaboration model
HIPAA Privacy Rule	Patient control over PHI	Granular consent flags; withdrawal supported

12.1 일급 감사 설계

2.4 절에서 명확히 한 바와 같이, 우리는 감사 추적성을 일급 설계 제약으로 다룬다. training_plan_recommendations 테이블은 모든 추천에 대해 6 개 점수 구성요소와 하위 구성요소, 종합 점수, 한국어와 영어 양쪽의 추론 텍스트, 가능한 경우 PubMed 식별자 또는 디지털 객체 식별자가 포함된 근거 인용, 고려된 금기사항(각각에 대해 도달한 결론과 함께), 그리고 추천을 생성한 점수 산정 모델의 버전 식별자를 영구 저장한다. 이 설계는 미국 식품의약국의 인공지능/머신러닝 의료기기 소프트웨어 행동 계획(FDA, 2021)의 최소 요구사항을 초과하며 Warraich 등(2024)의 수명주기 감사 강조와 일치한다.

12.2 점수 구성요소 영구 저장

score_components JSONB 열은 각 구성요소의 숫자 값뿐만 아니라 하위 구성요소와 값 계산에 사용된 방법도 저장한다. 예를 들어, symptom_match 구성요소는 코사인 유사도 값, 적용된 가중치 벡터, 환자 프로파일에서 추출된 symptom_priorities 슬라이스, 그리고 훈련 카탈로그의 일치 슬라이스를 기록한다. 이러한 세부 수준은 회고적 복제를 가능하게 한다: 영구 저장된 기록과 model_version 이 주어진다면, 감사자는 라이브 엔진 상태에 접근하지 않고도 점수를 재계산할 수 있다.

12.3 근거 인용 스키마

evidence_citations 필드는 JSONB 배열이다. 각 항목은 pmid 또는 doi, 인용이 지원하는 주장의 간단한 설명, 그리고 그 주장과 추천의 관계를 포함한다. 예를 들어, ADHD 뉴로피드백 추천은 지속적 효과 증거에 대해 Van Doren 등(2019)을 인용할 수 있으며, "ADHD 에 대한 뉴로피드백은 6 개월 추적 시 중간 효과 크기를 보인다"는 주장이 evidence_strength 구성요소에 연결된다. 이 스키마는 치료 시점에서의 임상가 검증과 규제 검토 중 외부 감사를 모두 지원한다.

12.4 이중 언어 추론 불변성

reasoning_ko 및 reasoning_en 필드는 추천 시점에 작성되며 이후 불변이다. 동일한 환자에 대한 후속 추천은 기존 행을 업데이트하기보다 새로운 행을 생성한다. 이 설계 선택은 종단 감사와 환자의 치료 과정 동안 엔진의 추론이 어떻게 진화했는지 이해하는 데 필요한 역사적 기록을 보존한다. 이중 언어 불변성은 단일 감사 프레임워크 하에서 한국어와 영어 사용자를 모두 담당하는 데 필수적이다.

12.5 규제 매핑

표 7 은 감사 필드를 특정 규제 요구사항에 매핑한다. 미국 식품의약국의 의료기기 소프트웨어 지침, 우수 머신러닝 관행 원칙, 일반 데이터 보호 규정 22 조의 전적으로 자동화된 의사결정의 대상이 되지 않을 권리, 그리고 유럽연합 AI 법의 고위험 시스템 투명성 요구사항이 모두 영구 저장 설계에 의해 다루어진다. 우리는 이 설계가 완전한 적합성에 필요하지만 충분하지는 않다고 주장한다; 다양한 인구 전반의 결과 검증과 시판 후 모니터링은 어떤 승인 추구에도 앞서 여전히 요구된다.

13. 보편적 증상 아키텍처

표 8. 10 개 증상 영역과 주요 훈련 범주.

Symptom Domain	Primary Training Categories	Key Evidence Source
ADHD (Gold Template)	Neurofeedback, CBT, mindfulness, environmental modification, skills training	Cortese et al. (2018); Van Doren et al. (2019)
Depression	CBT, behavioral activation, interpersonal therapy, mindfulness-based cognitive therapy	Cuijpers et al. (2020)
Anxiety	Exposure therapy, CBT, mindfulness, relaxation training	Carpenter et al. (2018)
Sleep	CBT-I, sleep hygiene, stimulus control, sleep restriction	Riemann et al. (2023)
Burnout	Workload management, recovery activities, boundary-setting, values clarification	Maslach & Leiter (2016)
PTSD	Trauma-focused CBT,	Bisson et al. (2020)

Symptom Domain	Primary Training Categories	Key Evidence Source
	prolonged exposure, EMDR, psychoeducation	
OCD	Exposure and response prevention, CBT, acceptance-based approaches	Skapinakis et al. (2016)
Bipolar	Psychoeducation, mood monitoring, social rhythm therapy, family-focused therapy	Goodwin et al. (2016)
Autism	Social skills training, sensory regulation, communication support, family support	Lord et al. (2022)
Peak Performance	Goal-setting, attention training, performance routine development	Fischer & Bidell (2006); skill-based literature

13.1 황금 템플릿으로서의 ADHD

우리는 초기 개발과 임상 검증을 위한 황금 템플릿으로 주의력 결핍 과잉행동 장애(ADHD)를 선택했다. 이 선택은 세 가지 고려사항을 반영한다. 첫째, ADHD는 주관적 증상 보고, 신경심리 평가(지속 수행 과제, n-back 과제), 바이오마커 데이터(정량적 뇌전도 세타/베타 비율; Arns et al., 2009; Van Doren et al., 2019) 사이에 가장 강력한 수렴을 가지며, 이는 6개 구성요소 점수 산정 체계의 완전한 사용을 가능하게 한다. 둘째, ADHD에 대한 근거 기반 중재는 약물요법, 인지행동치료, 뉴로피드백, 마음챙김, 환경 수정, 기술 훈련에 걸쳐 있으며(Cortese et al., 2018), 추천 엔진이 실제 다양성을 탐색할 수 있게 한다. 셋째, ADHD는 소아 및 성인 인구를 모두 포착하고 초기 구체적 조작에서 고급 추상화에 이르는 Fischer 기술 범위에 걸쳐 나타나며, 발달 매칭 구성요소를 작동시킨다.

황금 템플릿 전략은 다른 증상 영역으로의 복제 전에 ADHD의 심층 임상 검증을 약속한다. 검증 목표는 전문 임상가 선택과의 추천 일치(목표: 카파 ≥ 0.70), 외부 임상 검토자가 판단한 추론 품질(리커트 척도, 목표 평균 ≥ 5 점 만점에 4점), 그리고 이중 언어 추론 텍스트의 환자 이해 가능성(한국어 및 영어 표본 전반에서 목표 평균 ≥ 5 점 만점에 4점)을 포함한다.

13.2 증상 영역 복제 전략

우울증, 불안, 수면, 번아웃, 외상후스트레스장애, 강박장애, 양극성장애, 자폐, 최고 수행으로의 복제는 코드 작성이 아닌 데이터 채우기로 진행된다. 각 새로운 증상 영역에 대해, 팀은 훈련 라이브러리 시드(전체 quick_input_schema, fischer_start_level, contraindications 필드가 포함된 근거 기반 중재 세트)를 큐레이션하고, 금기사항 규칙 세트를 조립하며, 대표적 사례 표본에 대한 전문 임상 판단에 대비하여 추천 엔진의 출력을 검증한다. 동일한 6개 알고리즘과 동일한 데이터베이스 스키마가 모든 영역을 담당한다. 영역별 근거 출처에는 우울증에 대한 Cuijpers 등(2020), 불안에 대한 Carpenter 등(2018), 수면에 대한 Riemann 등(2023), 번아웃에 대한 Maslach 과 Leiter(2016), 외상후스트레스장애에 대한 Bisson 등

(2020), 강박장애에 대한 Skapinakis 등(2016), 양극성장애에 대한 Goodwin 등(2016), 자폐에 대한 Lord 등(2022)이 포함된다.

13.3 훈련 라이브러리 시드 데이터 큐레이션

증상별 훈련 라이브러리 시드는 박사 수준 훈련, 뉴로피드백 보드 인증, Harvard 교육대학원 소속(전직 방문학자)을 결합한 임상가에 의해 큐레이션된다. 시드는 가용한 경우 rct 또는 meta_analysis evidence_level의 중재를 우선시하며, consensus_guideline 및 expert_opinion 항목이 격차를 채운다. 각 시드는 한국과 영어 임상 실무 전반의 문화적 적절성에 대해 검토된다(Hwang et al., 2008). 큐레이션된 라이브러리 자체가 알고리즘 엔진과는 구별되지만 상호 보완적인 지적재산권의 한 형태이다.

13.4 증상 간 동반이환 처리

동반이환은 실제 정신건강 인구에서 예외가 아닌 규칙이다. 통합 프로파일 스키마는 범주적 단일 증상 할당이 아닌 영역 간 상대적 가중치를 인코딩하는 symptom_priorities 벡터를 통해 동반이환을 수용한다. ALG-1은 symptom_priorities의 argmax로 주요 증상 영역을 선택하지만 전체 벡터를 사용하여 관련된 부수적 고려사항을 표시한다. 예를 들어, 주요 ADHD와 이차적 우울증을 가진 환자는 우울증 친화적 양식(행동 활성화 중독)이 선호된다는 명시적 표기와 함께 ADHD 훈련 추천을 받을 수 있다. 동시 동시 치료 전반의 전체 다중 증상 계획은 향후 작업으로 보류된다.

14. 논의

14.1 강점

아키텍처의 주요 강점은 통합, 프라이버시, 추적성이다. 단일 스키마 하에서 평가, 추천, 세션 계획, 적응형 재계획을 통합함으로써, 설계는 대부분의 현대 디지털 정신건강 중재를 특징짓는 진단과 치료 사이의 구조적 격차를 제거한다. 붓-임상 브리지를 옴트인 기본 OFF 명시적 결정으로 다룸으로써, 설계는 환자 자율성을 존중하면서 적절한 경우 통합할 옵션을 보존한다. 모든 추천과 함께 전체 점수 분해, 근거 인용, 이중 언어 추론, 모델 버전을 영구 저장함으로써, 설계는 규제 기관, 임상가, 환자에게 의미 있는 감사 및 책임의 기질을 제공한다.

추가 강점은 보편적 증상 영역 구조이다. 동일한 6개 알고리즘과 동일한 데이터베이스 스키마가 10개 캐처 모두를 담당하며; 새로운 증상 영역은 코드 작성이 아닌 데이터 채우기로 추가된다. 이는 영역당 구현 비용을 최소화하고 전체 플랫폼 전반의 감사, 프라이버시 처리, 이중 언어 렌더링에서 일관성을 보장한다.

14.2 한계

몇 가지 한계가 적용된다. 첫째, 엔진은 아직 임상적으로 검증되지 않았다; 계획된 검증에는 전문 임상가 선택에 대한 추천 일치 연구, 외부 검토자와의 추론 품질 연구, 이중 언어 이해 가능성 연구가 포함된다. 둘째, 콜드 스타트 문제가 지속된다: 사전 결과 데이터가 없는 경우, personal_history 구성요소는 중립적 사전을 기여하며, 이는 추천이 근거 강도 지배로 되돌아갈 수 있게 한다. 셋째, 엔진은 임상가가 큐레이션한

콘텐츠가 정확하다고 가정한다; training_library 시드의 편향은 추천으로 전파된다. 넷째, 실시간 Fischer 기술 추정은 사소하지 않다; 현재 구현은 자가 보고와 임상가 시행 단축 평가를 사용하며, 이는 세밀한 수준 식별에 불충분할 수 있다.

14.3 임상적 함의

실무 임상가에게 아키텍처는 임상가 대체가 아닌 임상가-AI 협업 모델을 제공한다. ALG-1은 추론과 함께 순위 추천을 생성하며; 임상가는 선택, 수정 또는 무시한다. ALG-5는 임상가가 추천 엔진의 기대에 부합하는 맞춤형 훈련을 저작할 수 있게 한다. 봇 브리지 아키텍처는 임상가가 명시적 환자 동의가 있을 때만 봇 파생 임상 신호에 접근할 수 있게 하여 치료 동맹을 보존한다. 우리는 여기서 입증하지는 않지만, 이 협업 모델이 자율적 알고리즘 추천 시스템보다 임상가 수용을 증가시킨다고 가정한다.

14.4 규제 경로

아키텍처는 미국 식품의약국 의료기기 소프트웨어 클래스 II 승인 추구를 지원하도록 설계되었다. 감사 추적성 계층은 투명성, 시판 후 모니터링, 수명주기 감사에 관한 우수 머신러닝 관행 원칙을 다룬다. 오픈인 기본 OFF 봇 브리지는 일반 데이터 보호 규정 9 조와 유럽연합 AI 법의 고위험 시스템 투명성 요구사항을 다룬다. 이중 언어 렌더링 파이프라인은 국제 배포를 지원한다. 어느 프레임워크에 대한 적합성도 다양한 인구 전반의 결과 검증을 요구하며, 이는 ADHD 황금 템플릿 검증 단계의 주요 목표이다.

14.5 향후 작업

세 가지 향후 작업 라인이 우선순위가 된다. 첫째, 결과 데이터 통합: 환자가 세션을 완료하고 결과를 보고함에 따라, 엔진은 감사 추적을 보존하면서 지도 학습을 통해 점수 산정을 업데이트해야 한다. 둘째, 다중 모달 바이오마커 확장: 웨어러블 생리학적 데이터(소비자 기기의 심박변이도, 수면 아키텍처)를 biomarker_summary에 통합하면 biomarker_match 구성요소가 강화될 것이다. 셋째, 다중 증상 동시 치료 계획: 실제 환자는 종종 여러 증상 영역에 대한 동시 주의를 정당화한다; 순차적이 아닌 동시 훈련을 계획하도록 ALG-2를 확장하는 것은 상당한 설계 문제이다.

15. 결론

우리는 다중 증상 디지털 정신건강 플랫폼을 위한 AI 기반 개인 맞춤형 치료 경로 엔진을 제시했다. 아키텍처는 통합 관계형 데이터베이스 스키마에 대해 작동하는 6개 알고리즘으로 구성된다: 다중 구성요소 추천 엔진, Fischer 인식 세션 계획 생성자, 적응형 재계획자, 프라이버시 보존 봇-임상 브리지, 맞춤형 훈련 빌더, 이중 언어 결과 렌더러. 설계는 현대 디지털 정신건강의 세 가지 지속적인 격차를 다룬다: 평가와 개인화된 치료 계획 사이의 단절, 봇-임상 통합의 프라이버시 역설, 그리고 알고리즘 추천에 대한 감사 추적 가능한 추론의 부재.

세 가지 아키텍처적 약속이 이 작업을 구별한다. 첫째, 치료 경로 로직은 비구조화된 챗봇 대화나 정적 심리교육 콘텐츠에 암묵적이기보다 명시적이고, 모듈적이며, 감사 가능하다. 둘째, 대화 및 임상 데이터

스트림 간의 브리지는 일괄 통합이 아닌 명시적 옵트인 기본 OFF 동의 의미에 의해 권장된다. 셋째, 전체 아키텍처는 10 개 증상 영역 전반에 걸쳐 보편적이며, ADHD 가 우울증, 불안, 수면, 번아웃, 외상후스트레스장애, 강박장애, 양극성장애, 자폐, 최고 수행으로의 복제 전 임상 검증을 위한 황금 템플릿 역할을 한다.

시스템은 미국 식품의약국 의료기기 소프트웨어 지침, 유럽연합 AI 법, 일반 데이터 보호 규정, 건강보험 이동성 및 책임법과 규제 정렬되도록 설계되었다. 아키텍처 자체는 규제 승인에 필요하지만 충분하지 않다; 다양한 인구 전반의 결과 검증과 시판 후 모니터링은 어떤 승인 추구에도 앞서 여전히 요구된다. 다음 단계 작업은 ADHD 황금 템플릿의 임상 검증이며, 이후 나머지 9 개 증상 영역으로의 복제는 주로 소프트웨어 엔지니어링이 아닌 데이터 큐레이션 활동으로 진행될 수 있다.

직접적인 플랫폼 맥락을 넘어, 아키텍처는 프라이버시나 투명성을 희생하지 않고 대화형 및 임상 구성요소를 통합하는 다중 증상 디지털 정신건강 플랫폼 구축을 위한 일반화 가능한 템플릿을 제공한다. 특히 Fischer 역동적 기술 호환성 구성요소는 발달적으로 보정된 치료 매칭이 필요한 경우 본 플랫폼을 넘어 적용될 수 있다. 우리는 이 작업이 개인화, 프라이버시, 책임이 긴장 관계가 아니라 처음부터 공동으로 설계되는 디지털 정신건강 생태계에 기여하기를 희망한다.

References

- Almirall, D., Nahum-Shani, I., Sherwood, N. E., & Murphy, S. A. (2014). Introduction to SMART designs for the development of adaptive interventions: With application to weight loss research. *Translational Behavioral Medicine*, 4(3), 260–274. <https://doi.org/10.1007/s13142-014-0265-0>
- Arns, M., de Ridder, S., Strehl, U., Breteler, M., & Coenen, A. (2009). Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity: A meta-analysis. *Clinical EEG and Neuroscience*, 40(3), 180–189. <https://doi.org/10.1177/155005940904000311>
- Baumel, A., Muench, F., Edan, S., & Kane, J. M. (2019). Objective user engagement with mental health apps: Systematic search and panel-based usage analysis. *Journal of Medical Internet Research*, 21(9), e14567. <https://doi.org/10.2196/14567>
- Bisson, J. I., van Gelderen, M., Roberts, N. P., & Lewis, C. (2020). Non-pharmacological and non-psychological approaches to the treatment of PTSD: Results of a systematic review and meta-analyses. *European Journal of Psychotraumatology*, 11(1), 1795361. <https://doi.org/10.1080/20008198.2020.1795361>
- Carpenter, J. K., Andrews, L. A., Witcraft, S. M., Powers, M. B., Smits, J. A. J., & Hofmann, S. G. (2018). Cognitive behavioral therapy for anxiety and related disorders: A meta-analysis of randomized placebo-controlled trials. *Depression and Anxiety*, 35(6), 502–514. <https://doi.org/10.1002/da.22728>
- Cohen, I. G., & Mello, M. M. (2018). HIPAA and protecting health information in the 21st century. *JAMA*, 320(3), 231–232. <https://doi.org/10.1001/jama.2018.5630>
- Cortese, S., Adamo, N., Del Giovane, C., Mohr-Jensen, C., Hayes, A. J., Carucci, S., Atkinson, L. Z., Tessari, L., Banaschewski, T., Coghill, D., Hollis, C., Simonoff, E., Zuddas, A., Barbui, C., Purgato, M., Steinhausen, H.-C., Shokraneh, F., Xia, J., & Cipriani, A. (2018). Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: A systematic review and network meta-analysis. *The Lancet Psychiatry*, 5(9), 727–738. [https://doi.org/10.1016/S2215-0366\(18\)30269-4](https://doi.org/10.1016/S2215-0366(18)30269-4)
- Cuijpers, P., Karyotaki, E., Eckshtain, D., Ng, M. Y., Corteselli, K. A., Noma, H., Quero, S., & Weisz, J. R. (2020). Psychotherapy for depression across different age groups: A systematic review and meta-analysis. *JAMA Psychiatry*, 77(7), 694–702. <https://doi.org/10.1001/jamapsychiatry.2020.0164>
- Darcy, A., Daniels, J., Salinger, D., Wicks, P., & Robinson, A. (2021). Evidence of human-level bonds established with a digital conversational agent: Cross-sectional, retrospective observational study. *JMIR Formative Research*, 5(5), e27868. <https://doi.org/10.2196/27868>
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608. <https://arxiv.org/abs/1702.08608>

- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6), 477–531. <https://doi.org/10.1037/0033-295X.87.6.477>
- Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action and thought. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (6th ed., Vol. 1, pp. 313–399). Wiley.
- Fleming, T., Bavin, L., Lucassen, M., Stasiak, K., Hopkins, S., & Merry, S. (2018). Beyond the trial: Systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *Journal of Medical Internet Research*, 20(6), e199. <https://doi.org/10.2196/jmir.9275>
- Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, 3(11), e745–e750. [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9)
- Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a 'right to explanation'. *AI Magazine*, 38(3), 50–57. <https://doi.org/10.1609/aimag.v38i3.2741>
- Goodwin, G. M., Haddad, P. M., Ferrier, I. N., Aronson, J. K., Barnes, T. R. H., Cipriani, A., Coghill, D. R., Fazel, S., Geddes, J. R., Grunze, H., Holmes, E. A., Howes, O., Hudson, S., Hunt, N., Jones, I., Macmillan, I. C., McAllister-Williams, H., Miklowitz, D. R., Morriss, R., ... Young, A. H. (2016). Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *Journal of Psychopharmacology*, 30(6), 495–553. <https://doi.org/10.1177/0269881116636545>
- Heinz, M. V., Mackin, D. M., Trudeau, B. M., Bhattacharya, S., Wang, Y., Banta, H. A., Jewett, A. D., Salzhauer, A. J., Griffin, T. Z., & Jacobson, N. C. (2025). Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI*, 2(4). <https://doi.org/10.1056/Aloa2400802>
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.) (2024). *Cochrane Handbook for Systematic Reviews of Interventions* (Version 6.5). Cochrane. <https://training.cochrane.org/handbook>
- Hornstein, S., Zantvoort, K., Lueken, U., Funk, B., & Hilbert, K. (2023). Personalisation strategies in digital mental health interventions: A systematic review and conceptual framework for depressive symptoms. *Frontiers in Digital Health*, 5, 1170002. <https://doi.org/10.3389/fdgth.2023.1170002>
- Hwang, W.-C., Myers, H. F., Abe-Kim, J., & Ting, J. Y. (2008). A conceptual paradigm for understanding culture's impact on mental health: The cultural influences on mental health (CIMH) model. *Clinical Psychology Review*, 28(2), 211–227. <https://doi.org/10.1016/j.cpr.2007.05.001>
- Lee, K. H., Choi, G. H., Yun, J., Choi, J., Goh, M. J., Sinn, D. H., Jin, Y. J., Kim, M. A., Yu, S. J., Jang, S., Lee, S. K., Jang, J. W., Lee, J. S., Kim, D. Y., Cho, Y. Y., Kim, H. J., Kim, S., Kim, J. H., Kim, N., & Kim, K. M. (2024). Machine learning-based clinical decision support system for treatment

- recommendation and overall survival prediction of hepatocellular carcinoma: A multi-center study. *npj Digital Medicine*, 7, 2. <https://doi.org/10.1038/s41746-023-00976-8>
- Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2024). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry*, 23(1), 139–149. <https://doi.org/10.1002/wps.21183>
- Lord, C., Charman, T., Havdahl, A., Carbone, P., Anagnostou, E., Boyd, B., Carr, T., de Vries, P. J., Dissanayake, C., Divan, G., Freitag, C. M., Gotelli, M. M., Kasari, C., Knapp, M., Mundy, P., Plank, A., Scahill, L., Servili, C., Shattuck, P., ... McCauley, J. B. (2022). The Lancet Commission on the future of care and clinical research in autism. *The Lancet*, 399(10321), 271–334. [https://doi.org/10.1016/S0140-6736\(21\)01541-5](https://doi.org/10.1016/S0140-6736(21)01541-5)
- Mascolo, M. F., & Fischer, K. W. (2015). Dynamic development of thinking, feeling, and acting. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology and developmental science: Theory and method* (7th ed., Vol. 1, pp. 113–161). Wiley. <https://doi.org/10.1002/9781118963418.childpsy104>
- Maslach, C., & Leiter, M. P. (2016). Understanding the burnout experience: Recent research and its implications for psychiatry. *World Psychiatry*, 15(2), 103–111. <https://doi.org/10.1002/wps.20311>
- Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, 13, 23–47. <https://doi.org/10.1146/annurev-clinpsy-032816-044949>
- Murphy, S. A. (2005). An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine*, 24(10), 1455–1481. <https://doi.org/10.1002/sim.2022>
- Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, 52(6), 446–462. <https://doi.org/10.1007/s12160-016-9830-8>
- National Board for Certified Counselors (2023). National Clinical Mental Health Counseling Examination (NCMHCE). NBCC. <https://www.nbcc.org/exams/ncmhce>
- Reyes-Portillo, J. A., Mufson, L., Greenhill, L. L., Gould, M. S., Fisher, P. W., Tarlow, N., & Rynn, M. A. (2025). Generative AI mental health chatbots as therapeutic tools: Systematic review and meta-analysis. *JMIR Mental Health*, 12, e69220. <https://doi.org/10.2196/69220>
- Riemann, D., Espie, C. A., Altena, E., Arnardottir, E. S., Baglioni, C., Bassetti, C. L. A., Bastien, C., Berzina, N., Bjorvatn, B., Dikeos, D., Dolenc Groselj, L., Ellis, J. G., Garcia-Borreguero, D., Geoffroy, P. A., Gjerstad, M., Gonçalves, M., Hertenstein, E., Hoedlmoser, K., Hion, T., ... Spiegelhalder, K. (2023). The European Insomnia Guideline: An update on the diagnosis and treatment of insomnia 2023. *Journal of Sleep Research*, 32(6), e14035. <https://doi.org/10.1111/jsr.14035>

- Skapinakis, P., Caldwell, D. M., Hollingworth, W., Bryden, P., Fineberg, N. A., Salkovskis, P., Welton, N. J., Baxter, H., Kessler, D., Churchill, R., & Lewis, G. (2016). Pharmacological and psychotherapeutic interventions for management of obsessive-compulsive disorder in adults: A systematic review and network meta-analysis. *The Lancet Psychiatry*, 3(8), 730–739. [https://doi.org/10.1016/S2215-0366\(16\)30069-4](https://doi.org/10.1016/S2215-0366(16)30069-4)
- Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *npj Digital Medicine*, 3, 17. <https://doi.org/10.1038/s41746-020-0221-y>
- Tjoa, E., & Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Toward medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4793–4813. <https://doi.org/10.1109/TNNLS.2020.3027314>
- Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Whelan, P., Carvalho, A. F., Keshavan, M., Linardon, J., & Firth, J. (2021). The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*, 20(3), 318–335. <https://doi.org/10.1002/wps.20883>
- U.S. Food and Drug Administration (2021). Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan. U.S. Department of Health and Human Services. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>
- Van Doren, J., Arns, M., Heinrich, H., Vollebregt, M. A., Strehl, U., & Loo, S. K. (2019). Sustained effects of neurofeedback in ADHD: A systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, 28(3), 293–305. <https://doi.org/10.1007/s00787-018-1121-4>
- Voigt, P., & von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR): A practical guide*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-57959-7>
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76–99. <https://doi.org/10.1093/idpl/ix005>
- Wang, Y., Li, X., Zhang, Q., Yeung, D., & Wu, Y. (2025). Effect of a cognitive behavioral therapy-based AI chatbot on depression and loneliness in Chinese university students: Randomized controlled trial with financial stress moderation. *JMIR mHealth and uHealth*, 13, e72365. <https://doi.org/10.2196/72365>
- Wanniarachchi, V., Greenhalgh, M., Choi, A., & Warren, A. (2025). Personalization variables in digital mental health interventions for depression and anxiety in adolescents and youth: A scoping review. *Frontiers in Digital Health*, 7, 1500220. <https://doi.org/10.3389/fgdth.2025.1500220>

Warraich, H. J., Tazbaz, T., & Califf, R. M. (2024). FDA perspective on the regulation of artificial intelligence in health care and biomedicine. *JAMA*, 333(3), 241–247.
<https://doi.org/10.1001/jama.2024.21451>

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, 25(9), 1337–1340.
<https://doi.org/10.1038/s41591-019-0548-6>