[첨부 1]

AEGIS: 목적 중심 AI 협업 아키텍처 구조

1. 설계 개요

AEGIS는 다양한 작업에서 AI의 판단과 실행을 목적 중심으로 통합 관리하기 위해 설계한 AI 협업 아키텍처입니다.

일반적인 AI 활용이 개별 요청에 대한 답변이나 결과물 생성에 초점을 맞춘다면, AEGIS는 작업의 목적과 성공 기준을 먼저 정의하고, 이에 필요한 실행 절차와 기능을 구성하여 결과를 검증하는 전체 과정을 하나의 구조로 연결합니다.

이를 통해 사용자가 모든 작업 절차를 직접 설계해야 하는 부담을 줄이고, 다양한 작업에서도 공통된 기준을 유지하면서 목적에 적합한 결과를 도출하고자 합니다.

AEGIS는 별도의 AI 모델을 개발한 것이 아니라, 기존 생성형 AI에 공통 원칙과 작업별 실행 기준을 적용하여 활용하는 협업 구조입니다. 작업의 목적과 환경에 따라 필요한 기능과 절차를 구성하고, 실행 결과를 검토하며 개선하는 방식으로 운영합니다.

AEGIS는 다음 세 가지 핵심 요소를 중심으로 구성됩니다.

2. 핵심 아키텍처

2.1 AEGIS PRIME — 공통 실행 기준(Core)

AEGIS PRIME은 AEGIS 전체에 적용되는 핵심 기준이자 최상위 관리 요소입니다.

작업의 목적과 성공 기준을 중심으로 AI의 판단과 실행을 정렬하고, 모든 실행 과정에서 공통 원칙과 품질 기준이 유지되도록 합니다.

주요 역할은 다음과 같습니다.

- 사용자의 요청을 분석하여 작업의 목적과 성공 기준을 식별하도록 합니다.
- AI가 목적과 성공 기준에 맞춰 필요한 관점, 작업 모듈 및 실행 방법을 구성하도록 공통 기준을 제공합니다.
- 기능별 작업과 전체 실행 과정에 공통 원칙을 적용합니다.
- 최종 결과가 목적과 제약조건을 충족하는지 검증하도록 합니다.

AEGIS PRIME은 개별 작업에 따라 임의로 변경되는 것이 아니라, 전체 구조의 일관성을 유지하는 공통 기준으로 기능합니다.

2.2 Work Blocks — 기능별 실행 모듈

Work Blocks는 특정 목적의 작업을 수행하기 위해 구성한 기능별 실행 모듈입니다.

각 모듈은 고유한 목적, 입력, 실행 규칙, 검증 방법 및 출력으로 구성됩니다. AEGIS는 작업의 특성에 따라 필요한 Work Blocks를 선택하고 조합하여 실행 경로를 구성합니다.

주요 특징은 다음과 같습니다.

- **기능 분리:** 복잡한 작업을 목적별 기능으로 나누어 수행합니다.
- **선택적 조합:** 작업에 필요한 모듈만 선택하여 활용합니다.
- **공통 기준 적용:** 모든 모듈은 AEGIS PRIME의 공통 원칙을 따릅니다.
- **재사용 및 확장:** 기존 모듈을 다른 작업에서도 활용하고 필요한 기능을 확장할 수 있습니다.

이를 통해 모든 작업을 하나의 고정된 절차로 처리하지 않으면서도 공통 실행 기준을 유지할 수 있도록 설계했습니다.

2.3 Workflow — 통합 실행 절차

Workflow는 작업의 시작부터 최종 결과 출력까지 실행 과정을 연결하고 관리하는 절차입니다.

AEGIS PRIME이 공통 기준을 제시하고 Work Blocks가 개별 기능을 수행한다면, Workflow는 각 단계의 결과를 연결하고 전체 작업이 목적에 맞게 진행되도록 합니다.

Workflow는 모든 작업에 동일한 순서를 강제하지 않습니다. 아래 절차는 기본적인 실행 흐름을 나타내며, 실제 작업에서는 목적과 복잡도, 요구사항 및 제약조건에 따라 단계를 생략하거나 조정하고 필요한 부분을 반복할 수 있습니다.

기본 실행 절차는 다음과 같습니다.

요구사항 분석 → 실행 계획 수립 → 기능별 실행 → 결과 통합 → 검증 → 수정 및 개선 → 최종 출력

단계	주요 역할
요구사항 분석	요청, 목적, 성공 기준 및 제약조건을 식별합니다.
실행 계획 수립	필요한 Work Blocks와 실행 경로를 결정합니다.
기능별 실행	선택된 모듈을 활용하여 개별 작업을 수행합니다.
결과 통합	개별 결과를 연결하고 논리적 일관성을 검토합니다.
검증	목적, 성공 기준 및 제약조건의 충족 여부를 확인합니다.
수정 및 개선	문제가 발견되면 필요한 부분을 수정하고 다시 검증합니다.
최종 출력	검증 결과를 목적과 활용 환경에 맞는 형태로 제공합니다.

작업의 목적이나 핵심 조건이 불분명한 경우에는 사용자에게 필요한 정보를 요청하며, 문제가 발견되면 전체 과정을 무조건 반복하기보다 필요한 부분을 중심으로 수정하도록 설계했습니다.

3. 전체 구조의 작동 원리

AEGIS의 세 가지 핵심 요소는 하나의 실행 체계로 연결됩니다. AEGIS PRIME은 작업 전반에 적용할 공통 원칙과 판단 기준을 제공하고, AI는 사용자가 제시한 목적과 성공 기준에 따라 필요한 Work Blocks와 Workflow를 구성합니다. 이후 AI는 구성된 작업 모듈과 실행 절차에 따라 개별 작업을 수행하고 결과를 통합·검증하며, 필요한 부분을 수정합니다. 이 과정에서 중요한 판단과 최종 결과 확인은 사람이 담당합니다.

이 구조의 핵심은 작업마다 실행 방식은 달라질 수 있지만, 목적을 중심으로 판단하고 실행하며 결과를 검증한다는 공통 기준은 유지된다는 점입니다.

4. 핵심 설계 원칙

AEGIS는 다음 원칙을 바탕으로 전체 구조의 일관성과 실행 효율성을 확보하도록 설계했습니다.

- **목적 중심성:** 모든 판단과 실행을 현재 작업의 목적과 성공 기준에 맞춥니다.
- **사용자 협업:** 사람과 AI의 역할을 구분하고 작업 과정에서 상호 검토와 수정을 반복합니다. 사람은 목적과 성공 기준, 주요 의사결정 및 최종 승인을 담당하며, AI는 작업 구성과 실행, 결과 검증을 지원합니다.
- **적응형 실행:** 작업의 복잡도와 특성에 따라 필요한 기능과 실행 방식을 조정합니다.
- **근거 및 사실성:** 확인 가능한 근거를 우선하고, 사실과 추론 및 불확실성을 구분합니다.
- **맥락과 완전성 유지:** 이전 결정사항을 보존하고 결과의 누락과 모순을 점검합니다.
- **자원 효율성:** 기존 실행 자산을 재사용하고 불필요한 중복 작업을 줄입니다.
- **핵심 기준 보호:** AEGIS의 핵심 실행 자산은 임의로 변경하지 않으며, 변경이 필요한 경우 사용자의 승인을 받도록 합니다.

이 원칙들은 실행 방식을 지나치게 고정하기 위한 것이 아니라, 작업별 유연성을 유지하면서도 공통된 판단 기준을 확보하기 위한 것입니다.

5. 이중 검증과 자기 수정

AEGIS는 개별 모듈의 수행 결과와 전체 작업의 최종 결과를 구분하여 검증합니다.

개별 모듈 검증은 각 Work Block의 결과가 해당 작업의 목적과 실행 규칙에 부합하는지 확인하는 과정입니다. **전체 결과 검증**은 통합된 결과가 전체 목적, 성공 기준, 제약조건 및 공통 원칙을 충족하는지 확인하는 과정입니다.

검증 과정에서 문제가 발견되면 원인을 분석하고 필요한 부분을 수정한 뒤 다시 확인합니다. 해결할 수 없는 문제나 남아 있는 한계가 있다면 이를 명시하도록 합니다.

이러한 자기 수정(Self-Correction) 방식은 결과물을 생성하는 데서 작업을 끝내지 않고, 검증과 개선을 거쳐 목적에 더 적합한 결과를 도출하기 위한 것입니다. 다만 검증 절차 자체가 모든 오류의 발견이나 제거를 보장하는 것은 아닙니다.

6. 요약

AEGIS는 **공통 실행 기준(AEGIS PRIME), 기능별 실행 모듈(Work Blocks), 통합 실행 절차(Workflow)**를 연결한 목적 중심 AI 협업 아키텍처입니다.

AEGIS PRIME이 판단과 실행의 기준을 유지하고, Work Blocks가 필요한 기능을 수행하며, Workflow가 실행·통합·검증·수정 과정을 관리합니다.

이 구조를 통해 작업별 유연성을 유지하면서도 목적 중심의 일관된 실행과 결과 검증을 추구합니다.
AEGIS는 이러한 원리를 다양한 실제 작업에 적용하고, 활용 경험을 바탕으로 지속적으로 개선하기 위해 설계되었습니다.