

데이터 수집 및 저장 프로젝트 기획서

□ 개요

- 산출물 단계 : 데이터 수집 및 저장
- 평가 산출물 : 프로젝트 기획서
- 제출 일자 : 2025.02.10
- 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN05-final-3Team>
- 작성 팀원 : 최영민

1. 프로젝트 개요

FinPilot 프로젝트는 사용자가 구글 Docs를 활용한 비즈니스 문서 작업을 더욱 효율적으로 수행할 수 있도록 돕기 위한 목적으로 기획되었습니다. 이 프로젝트는 구글 Docs의 기본 기능을 확장하여 사용자 경험을 향상시키고, 생산성을 높이는 것을 목표로 하고 있습니다. 구글 Docs는 많은 사람들이 문서 작업을 수행하는 데에 사용하고 있는 인기 있는 플랫폼으로, 공동 작업 등 다양한 기능을 제공하지만, 여전히 문서 작업의 전 과정을 사용자가 직접 수행해야 한다는 비효율성을 가지고 있습니다. 이러한 한계를 극복하기 위해, 본 프로젝트는 사용자들이 비즈니스 문서 작업 간 필요로 하는 다양한 도구와 기능(초안 작성, 단락 생성, 데이터 시각화 등)을 추가하여 보다 직관적이고 유용한 작업 환경을 제공하고자 합니다. 이를 통해 사용자들은 문서 작성, 편집, 협업 등의 과정을 더욱 원활하게 진행할 수 있으며, 결과적으로 시간과 노력을 절약할 수 있게 될 것입니다. 이와 같은 노력은 구글 Docs의 활용도를 높이고, 사용자들에게 보다 풍부한 경험을 제공하는 데 기여할 것입니다.

1.1 프로젝트 배경

FinPilot 팀은 다음 세 가지 측면에 주목하였습니다.

먼저 **비효율적인 문서 작업 과정**입니다. 기존의 문서 작업 방식은 ‘주제 설정’, ‘자료 수집’, ‘데이터 분석’, ‘초안 작성’, ‘편집’, ‘배포’의 전 과정을 사용자가 직접 수행해야 했습니다. 즉, 문서를 작성하기 위한 자료를 수집하고, 데이터를 분석하여 초안을 작성한 후 편집하는 과정까지 사용자가 일일이 수행해야 했습니다. 이 과정에서 정보와 기능, 도구가 다양한

곳에 흩어져 있어 다양한 인터페이스를 넘나들며 자료를 수집하고 데이터를 분석하며 문서 작업을 수행해야 한다는 불편함을 가지고 있었습니다. 이 점에 주목하여 ‘자료 수집’, ‘데이터 분석’, ‘작성’, ‘편집’을 하나의 인터페이스에 수행할 수 있도록 문서 작업을 혁신하고자 하였습니다.

다음으로 **도메인 특화 솔루션** 입니다. 비즈니스 문서, 특히 금융/증권 도메인의 문서는 복잡한 용어와 정교한 데이터 분석을 요구하며, 작성과 검토에 많은 시간과 노력을 필요로 합니다. 그러나 기존의 솔루션은 템플릿 제공에 국한되어 문서 작성의 자유도와 실시간 피드백 기능이 부족했습니다. 또한, 데이터 분석과 인사이트 제공 기능이 미흡하고 금융 관련 **API**를 활용하여 데이터를 분석하고 문서화 하는 국내 서비스 사례는 부족했습니다. 이에 따라 금융/증권 분야의 용어를 이해하고 데이터 분석과 문서 작업을 도와주는 비즈니스 문서 작업 어시스턴트 의 필요성을 인식하였습니다.

마지막으로 **사용성** 입니다. **Google Workspace**는 사용자가 전 세계적으로 30억명 이상의 사용자를 보유하고 있으며, 기업과 교육 기관에서 널리 활용하고 있습니다. **Google Workspace** 생태계의 일원인 **Google Docs**는 전 세계적으로 널리 사용되는 클라우드 기반의 문서 작성 도구로, 협업 기능과 접근성 측면에서 많은 장점을 가지고 있습니다. 다양한 도구와 통합 사용이 가능하고, 기업용 플랜도 저렴하여 중소기업, 스타트업에게 경제적 입니다. 따라서 강력한 협업 도구인 **Google Docs** 에서 제공되는 **AI** 어시스턴트 솔루션을 확장 프로그램으로 개발하여 사용자 맞춤형 기능을 제공하고자 합니다.

1.2 프로젝트 목표

1. **사용자 경험 향상:** 확장 프로그램 형태로 구현해 구글 **Docs**의 기본 기능에 더하여 사용자들이 보다 직관적이고 편리하게 문서를 작성하고 편집할 수 있도록 지원합니다.
2. **생산성 증대:** 반복적인 작업을 자동화하고, 문서 작성 및 편집 시간을 단축시켜 사용자들의 생산성을 높입니다.
3. **정보/기능 집약적 설계:** 자료 수집, 분석, 작성, 편집 을 하나의 인터페이스에서 수행할 수 있도록 서비스하여 문서 작업을 혁신합니다.

1.3 주요 기능

- **초안 작성:** 사용자가 요청한 문서의 초안을 작성합니다. **FinPilot**은 사람의 문서 작성 과정을 모방하여, 먼저 문서의 섹션을 구성하고 각 섹션에 맞는 내용을 작성합니다. 각 섹션의 내용을 작성하는 과정에서 필요한 자료를 스스로 판단하여 수집합니다.

주식, 재무제표, 뉴스 등의 관련 자료를 수집한 후 이를 분석하여 문서를 작성합니다.

- **단락 생성:** 사용자가 원하는 특정 주제에 대한 단락을 생성합니다. 초안 작성 후 해당 문서의 내용을 보완하는 과정에서 사용할 수 있습니다. PDF를 업로드하면 해당 문서의 내용을 기반으로 단락을 생성하며 PDF가 제공되지 않아도, 웹 검색을 통해 정보를 수집하여 단락을 생성합니다. 기본적으로 **Self-RAG** 과 **Corrective RAG** 아키텍처를 활용하여 단락을 생성합니다.
- **문서 요약 / 확장:** 문서의 특정 섹션을 요약하거나 확장합니다. 문서의 내용을 해치지 않으면서 문서의 분량을 쉽게 조절할 수 있습니다.
- **데이터 시각화:** 사용자가 원하는 데이터를 분석하고 시각화 합니다. 웹 상의 데이터를 수집하여 시각화가 가능하며, 사용자가 시각화를 원하는 **csv** 파일을 업로드하면 이를 기반으로 데이터 분석과 시각화를 수행할 수 있습니다.

1.4 기대 효과

이 프로젝트를 통해 업무 간 **Google Docs** 활용하는 기업과 사용자들은 보다 효율적이고 생산적인 문서 작성 환경을 경험할 수 있을 것입니다. 또한, 기업 및 교육 기관 등 다양한 분야에서의 활용도를 높여, 업무 프로세스의 효율성을 한층 더 강화할 수 있을 것으로 기대됩니다.

2. 시장 분석

구글 **Docs** 확장 프로그램 시장은 최근 몇 년간 급격한 성장을 보이고 있으며, 이는 클라우드 기반 협업 도구의 수요 증가와 맞물려 있습니다. 구글 **Workspace**의 일환으로 제공되는 구글 **Docs**는 다양한 기기에서 접근 가능하며, 무료로 제공된다는 점에서 많은 사용자들에게 인기를 끌고 있습니다. 이러한 배경에서 구글 **Docs** 확장 프로그램 시장은 더욱 확장될 것으로 예상됩니다. 또한, **AI**를 활용한 기업 업무 프로세스 효율성 증대는 이미 많은 기업에서 시도하고 있습니다. 따라서 관련 시장은 계속 커질 것으로 예상됩니다.

2.1 국내 기업 AI 활용 사례

1. 미래에셋증권 - 네이버클라우드의 'HCX-DASH'

- 미래에셋증권은 네이버클라우드의 금융 특화 **sLLM**을 도입하여, 자체 데이터를 학습시켜 금융업무에 최적화 된 모델을 사내 온프레미스 환경에서 운영하고 있습니다.
- 금융 보고서 작성을 자동화 및 효율화하고 있으며, 내부 데이터를 분석하고

문서 요약 작업을 지원합니다. 금융 규제를 준수하도록 지원하면서 AI를 통한 비즈니스 프로세스를 혁신한 사례입니다.

2. KB금융 - 삼성SDS와 협력한 'Gen AI 기반 sLLM'

- KB금융은 삼성SDS와 패브릭스(FabriX) 플랫폼 기반 생성형 AI 솔루션 사용하고 있습니다. 이를 기반으로 그룹 내 통합된 AI 거버넌스 체계 마련하였으며, 계열사 간 AI 솔루션 도입 및 운영 효율화.
- 금융 소비자에게 안전하고 신뢰성 있는 AI 서비스를 제공함과 동시에, 대용량 데이터 전송 톨, 프록시 및 모니터링 기능, 문서 복호화를 위한 DRM 솔루션 포함을 포함하고 있습니다.

2.2 해외 기업 AI 활용 사례

1. 데이터 요약 및 분석 (BloombergGPT)

- 금융문서 요약 및 핵심 정보 추출
- 시장 감성 분석
- 재무 보고서 및 시장 동향 데이터 기반 예측 인사이트 제공

2. 고객 서비스 개선 (Bank of America Erica)

- 고객 맞춤형 재무 조언, 다국어 지원
- 24/7 서비스 제공

3. 리스크 관리 및 사기 탐지 (Wells Fargo)

- LLM 및 딥러닝 활용 대규모 거래 패턴 분석 및 사기 탐지

4. 투자 포트폴리오 최적화 (Wealthfront)

- 시장 동향 분석, 세금 최적화, 포트폴리오 조정

5. 규제 준수 자동화 (HSBC와 Quantexa)

- 고객확인제도(KYC) 자동화
- 금융 네트워크에 숨겨진 패턴 및 관계 식별, 의심스러운 활동 보고서 생성

6. 문서 처리 효율화 (Allganize Finance 13B)

- 미국,일본,한국 금융 법규 학습, 재무 문서 내 주요 정부 추출 및 요약
- 법률적 맥락에서 규정 준수 여부 검토, 다국어 지원.

2.3 국내 기업 Google WorkSpace 활용 사례

1. 세아 그룹

- 온프레미스 기반의 그룹웨어와 파일 관리 시스템을 Google Workspace로 전환하여 파일공유와 협업을 개선한 사례입니다.

2. 대한항공

- Google Workspace 도입을 통해 협업과 커뮤니케이션 중심의 업무 환경을

조성하였습니다.

- 운항 및 객실 승무원들은 Google Docs를 통해 최신의 기내 교본과 정보를 손쉽게 접근하는 등 업무 효율성 증대를 도모한 사례입니다.

2.4 Google Docs 시장 규모 및 성장 전망

Google Workspace는 2023년 약 25억 8,223만 달러에서 2031년 약 49억 2,717만 달러로 성장할 것으로 예상되는 만큼 사용성이 우수한 플랫폼입니다. 이는 구글 Docs와 같은 응용 프로그램 내에서 자동화할 수 있는 앱, 워크플로, 프로세스를 만들 수 있는 기능의 수요 증가에 기인합니다.

주요 트렌드

1. **생산성 및 접근성 강화:** 최근 구글은 크롬 확장 프로그램 중 생산성과 접근성에 중점을 둔 프로그램들을 선정하여 발표했습니다. 이는 사용자들이 보다 효율적으로 작업할 수 있도록 돕는 도구들로, 오디오 회의를 텍스트로 녹음하는 '트랜스크립터'와 같은 프로그램이 포함되어 있습니다.
2. **협업 기능의 중요성:** 구글 Docs는 다양한 기기 간의 공유와 협업 작업을 중요시하며, 이는 사용자들이 언제 어디서나 문서 작업을 할 수 있도록 지원합니다. 이러한 기능은 특히 원격 근무 환경에서 더욱 중요해지고 있습니다.
3. **스타트업과의 협력:** 구글플레이는 창구 프로그램을 통해 스타트업들이 해외 시장에 진출하고 AI 기술을 도입할 수 있도록 지원하고 있으며, 이는 구글 Docs 확장 프로그램의 생태계 확장에 기여하고 있습니다.

이러한 시장 분석을 바탕으로, FinPilot 프로젝트는 사용자 요구에 부합하는 기능을 제공함으로써 시장에서의 경쟁력을 확보할 수 있을 것입니다.

2.5 경쟁사 분석

구글 Docs 확장 프로그램 시장에서 주요 경쟁사로는 마이크로소프트의 Office 365가 있습니다. Office 365는 구글 Docs와 유사한 클라우드 기반의 문서 작성 및 협업 도구를 제공하며, 특히 기업 환경에서 강력한 입지를 가지고 있습니다.

마이크로소프트 Office 365의 강점

1. **통합된 생태계:** Office 365는 Word, Excel, PowerPoint 등 다양한 생산성 도구를 통합하여 제공하며, 이는 사용자들이 하나의 플랫폼 내에서 모든 작업을 수행할 수

있도록 지원합니다.

2. **보안 및 관리 기능:** 기업 환경에서 중요한 보안 및 관리 기능을 강화하여, IT 관리자가 사용자 및 데이터에 대한 통제력을 유지할 수 있도록 합니다.
3. **AI 및 자동화 기능:** 마이크로소프트는 AI 기술을 활용하여 문서 작성 및 데이터 분석을 자동화하는 기능을 제공, 사용자들이 보다 효율적으로 작업할 수 있도록 돕고 있습니다.

FinPilot의 차별화 전략

Google Docs는 무료로 제공되며, 다양한 기기에서 접근 가능하다는 점에서 개인 사용자 및 소규모 팀에게 매력적입니다. 또한, 실시간 협업 기능이 강력하여, 팀원 간의 원활한 커뮤니케이션을 지원합니다. 이러한 점에서 구글 Docs는 Office 365와 차별화된 사용자 경험을 제공하고 있습니다.

2.6 타겟 사용자 정의

FinPilot의 타겟 사용자는 주로 업무 상 문서 작업의 효율화를 원하는 그룹으로 구성됩니다. 이들은 주로 구글 Docs의 기본 기능을 확장하여 보다 효율적이고 맞춤형 작업 환경을 원하는 사용자들입니다.

주요 타겟 사용자 그룹

1. **기업 및 비즈니스 사용자:** 기업 내 팀원들이 문서 작성 및 협업을 위해 구글 Docs를 사용하는 경우 FinPilot을 사용하여 업무 효율화를 이룰 수 있습니다. 이들은 문서 관리, 버전 관리, 보안 기능 강화와 같은 비즈니스 요구를 충족시키는 확장 프로그램을 필요로 합니다.
2. **프리랜서 및 개인 사용자:** 기업 뿐 만 아니라 비즈니스 문서 작업을 구글 Docs를 통해 클라이언트와의 협업 및 프로젝트 관리를 수행하는 개인의 경우도 FinPilot은 훌륭한 동반자가 될 것입니다. 이들은 작업 효율성을 높이는 자동화 도구와 클라이언트와의 원활한 커뮤니케이션을 지원하는 기능을 선호합니다.

3. 주요 기능

구글 Docs 확장 프로그램의 주요 기능은 사용자 경험을 향상시키고, 문서 작성 및 협업을 보다 효율적으로 수행할 수 있도록 설계되었습니다. 이 섹션에서는 확장 프로그램이 제공할

주요 기능들을 상세히 설명합니다.

문서 작성/편집

- **초안 작성:** 사용자가 요청한 문서의 초안을 작성합니다. FinPilot은 사람의 문서 작성 과정을 모방하여, 먼저 문서의 섹션을 구성하고 각 섹션에 맞는 내용을 작성합니다. 각 섹션의 내용을 작성하는 과정에서 필요한 자료를 스스로 판단하여 수집합니다. 주식, 재무제표, 뉴스 등의 관련 자료를 수집한 후 이를 분석하여 문서를 작성합니다.
- **단락 생성:** 사용자가 원하는 특정 주제에 대한 단락을 생성합니다. 초안 작성 후 해당 문서의 내용을 보완하는 과정에서 사용할 수 있습니다. PDF를 업로드하면 해당 문서의 내용을 기반으로 단락을 생성하며 PDF가 제공되지 않아도, 웹 검색을 통해 정보를 수집하여 단락을 생성합니다. 기본적으로 **Self-RAG** 과 **Corrective RAG** 아키텍처를 활용하여 단락을 생성합니다.
- **문서 요약 / 확장:** 문서의 특정 섹션을 요약하거나 확장합니다. 문서의 내용을 해치지 않으면서 문서의 분량을 쉽게 조절할 수 있습니다.
- **문서 반영:** 생성된 텍스트를 버튼으로 손쉽게 문서에 반영합니다. 생성되는 텍스트의 양식을 복사하여 반영할 수 있으며, 스타일을 제외한 텍스트를 반영하고 싶다면 곧바로 문서에 반영하는 것도 가능합니다.

데이터 분석/시각화

- **데이터 분석/시각화:** 사용자가 원하는 데이터를 분석하고 시각화 합니다. 웹 상의 데이터를 수집하여 시각화가 가능하며, 사용자가 시각화를 원하는 **csv** 파일을 업로드하면 이를 기반으로 데이터 분석과 시각화를 수행할 수 있습니다. 텍스트 반영과 마찬가지로 버튼으로 손쉽게 문서에 반영이 가능합니다.

파일 업로드/삭제

- **PDF 업로드:** 문서 내용 작성 간 반영을 원하는 PDF 파일을 업로드합니다. 업로드 된 PDF 파일은 단락 생성 과정에서 참조합니다. 여러 파일을 업로드 후 특정 파일만 삭제하는 것도 가능하며, 1인 당 최대 4개의 PDF 파일 업로드를 지원합니다.
- **CSV 업로드:** 데이터 시각화 과정에서 분석을 원하는 CSV 파일을 업로드합니다. 분석/시각화 과정에서 해당 데이터를 기반으로 수행합니다.

출처 확인

- **자료 출처 확인:** 비즈니스 문서에서 정확한 출처를 확인하고 명시하는 것은 매우

중요합니다. FinPilot은 생성된 콘텐츠의 출처를 확인할 수 있도록 생성에 활용된 모든 출처를 확인할 수 있도록 지원합니다.

이러한 주요 기능들을 통해, FinPilot은 사용자 요구를 충족시키고, 문서 작업 간 효율성을 극대화 할 수 있을 것입니다.

4. 개발 일정 및 단계

FinPilot 프로젝트의 성공적인 진행을 위해 프로젝트의 주요 개발 단계와 각 단계의 목표를 설정합니다.

4.1. 개발 일정

- 1. 프로젝트 계획 및 요구사항 분석 (1주):**
 - 프로젝트의 목표와 범위를 정의하고, 요구 사항을 정의합니다.
 - 시장 조사 및 경쟁사 분석을 통해 제공할 기능을 정의합니다.
- 2. 프로토타입 개발 (3주):**
 - 정의한 요구 사항을 기반으로 프로토 타입을 개발하고 주요 기능을 테스트 합니다.
 - 중간 평가를 통해 피드백을 수렴하여 개선 점을 도출합니다.
- 3. 기능 개발 및 통합 (3주):**
 - 도출된 피드백을 바탕으로 전체 시스템을 개선합니다.
 - 시스템 전체 기능의 구현을 완료하고 테스트 합니다.
- 4. 테스트 (0.5주):**
 - 기능 테스트 및 사용자 테스트를 통해 버그를 식별하고 수정합니다.
 - 성능 최적화 및 보안 검토를 수행합니다.
- 5. 배포(0.5주):**
 - 확장 프로그램을 웹 스토어에 배포하고, 사용자 교육 자료를 제공합니다.

4.2. 개발 단계

- 초기 계획 단계:** 프로젝트의 전반적인 계획을 수립하고, 팀 구성 및 역할을 정의합니다.

이름	직책	주요 업무
최영민	PM AI 엔지니어	<ul style="list-style-type: none"> 프로젝트 일정/리소스 관리, 커뮤니케이션 총괄 sLLM 어플리케이션 개발 및 최적화
허상호	Back-end 개발자 AI 엔지니어	<ul style="list-style-type: none"> 서버, 데이터베이스 설계 sLLM 모델 개발 및 최적화
서장호	Front-end 개발자 AI 엔지니어	<ul style="list-style-type: none"> Chrome Extension UI 설계 및 개발 sLLM 모델 개발 및 최적화
조주영	데이터 수집 및 검증	<ul style="list-style-type: none"> 시장조사를 통한 데이터 수집 및 데이터 검증

- **개발 단계:** sLLM/LLM 어플리케이션을 구현하고 Chrom Extension과 서버를 구축하여 통합합니다.
 - sLLM/LLM 어플리케이션 모델링 방안
 - LangGraph를 활용한 sLLM/LLM 어플리케이션 구현
 - sLLM/LLM Agent 정의
 - 초안 작성 Agent
 - 사용자가 요청한 문서에 대해 목차(outline)를 생성하고 각 목차에 대한 내용을 작성합니다.
 - 목차별 내용 작성 시 필요한 데이터를 수집하여 반영하며, 관련 출처를 반환합니다.
 - 단락 생성 Agent
 - 초안 생성과 달리 목차를 구성하지 않고 구체적인 요청에 대한 단락을 생성합니다.
 - self RAG와 Corrective RAG의 아키텍처를 사용하여 웹 기반 자료와 사용자가 업로드한 자료를 기반으로 단락을 작성합니다.

- 요약 / 확장 Agent
 - 사용자가 요청한 문장을 요약/확장 하여 문서 분량 조절을 지원합니다.
- 데이터 분석 / 시각화 Agent
 - 사용자가 업로드한 **CSV** 파일의 내용이나 웹 에서 수집한 데이터를 분석하여 시각화 자료를 제공합니다.
- 사용 데이터
 - 재무 제표 데이터 : 대차대조표, 손익계산서, 현금흐름표 등 다양한 기업의 재무 제표 데이터는 문서 작성 과정에서 반드시 필요합니다. 다양한 산업 및 규모의 기업을 포괄하며, **API** 활용해 데이터를 수집합니다.
 - 주식 데이터 : 주식 데이터는 시장을 분석하여 문서를 작성하는 과정에서 반드시 필요합니다. **API**를 활용해 데이터를 수집합니다.
 - 웹/뉴스 데이터 : 뉴스 기사, 소셜 미디어 게시물, 블로그 글 등 텍스트 데이터는 시장 심리 분석 및 문서 작성 과정에서 필요합니다. 금융 시장 관련 다양한 주제의 텍스트 데이터를 수집하여 문서 작성, 데이터 시각화 등에 활용합니다.
- 테스트 단계: 사용자 유즈 케이스를 연구하고 이를 바탕으로 테스트 계획서를 작성합니다. 작성된 테스트 계획서를 바탕으로 체계적인 테스트를 통해 프로그램의 안정성과 성능을 검증합니다.
- 배포 단계: 최종 제품을 **Chrome** 웹 스토어에 배포합니다.

* 비교 : 이 문서는 FinPilot의 도움을 받아 작성되었음.