



ФАКУЛТЕТ "ПРИЛОЖНА ИНФОРМАТИКА И СТАТИСТИКА"
КАТЕДРА "ИНФОРМАЦИОННИ ТЕХНОЛОГИИ И КОМУНИКАЦИИ"

„Методи за автоматично
генериране на предложения за
оптимизиране на бизнес процеси“

*Автореферат на дисертационен труд за
придобиване на образователна и научна
степен "доктор"*

Докторант:

Албена Минчева

Научен ръководител:

Гл. ас. д-р Гено Стефанов

София, 2025 г.

СЪДЪРЖАНИЕ

ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД	3
1. АКТУАЛНОСТ НА ПРОБЛЕМА.....	3
2. ОБЕКТ И ПРЕДМЕТ НА ИЗСЛЕДВАНЕТО.....	5
3. ЦЕЛ И ЗАДАЧИ НА ДИСЕРТАЦИОННИЯ ТРУД	6
4. РАБОТНИ ХИПОТЕЗИ НА ДИСЕРТАЦИОННИЯ ТРУД	9
5. НАУЧНИ И НАУЧНО-ПРИЛОЖНИ ПРИНОСИ.....	10
6. ОБЕМ И СТРУКТУРА НА ДИСЕРТАЦИОННИЯ ТРУД.....	14
КРАТКО ИЗЛОЖЕНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД.....	16
1. УПРАВЛЕНИЕ НА БИЗНЕС ПРОЦЕСИ.....	16
2. КОНЦЕПТУАЛНА И ТЕХНОЛОГИЧНА АРХИТЕКТУРА.....	22
3. НАБОРИ ОТ ДАННИ	36
4. ИЗВЛИЧАНЕ НА ИНФОРМАЦИЯ И ГЕНЕРИРАНЕ НА БИЗНЕС ПРАВИЛА	43
5. ГЕНЕРИРАНЕ И ОЦЕНКА НА ПРЕДЛОЖЕНИЯ ЗА ОПТИМИЗАЦИЯ	63
6. МЕТОДОЛОГИЯ ЗА АВТОМАТИЗИРАНО ГЕНЕРИРАНЕ НА ПРЕДЛОЖЕНИЯ ЗА ОПТИМИЗАЦИЯ НА БИЗНЕС ПРОЦЕСИ.....	74
7. ЗАКЛЮЧЕНИЕ	79
ЛИТЕРАТУРА	81
СПИСЪК ФИГУРИ	86
СПИСЪК ТАБЛИЦИ	87

ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

1. АКТУАЛНОСТ НА ПРОБЛЕМА

В динамичния и конкурентен пейзаж на съвременните организации, ефективността и адаптивността са от първостепенно значение. В основата на всяка организация стоят сложни и взаимосвързани бизнес процеси, които са ключов фактор за осигуряване на стойност за клиентите, устойчивост на пазара и оперативна ефективност. Управлението на тези процеси е основополагащо за постигането на стратегическите цели, тъй като позволява на организациите да дефинират, анализират, оптимизират и адаптират своите операции спрямо динамичните пазарни условия.

От друга страна в съвременния бизнес контекст организациите са изправени и пред необходимостта от непрекъснато адаптиране на своите бизнес процеси в отговор на динамично променящите се регулаторни изисквания и технологични иновации. Управлението на бизнес процеси (Business Process Management – BPM) се е утвърдило като ключова дисциплина, която позволява на организациите не само да повишават ефективността си, но и да осигуряват устойчиво съответствие с нормативната рамка. В този контекст адаптацията на процесите към законовите изисквания следва да се разглежда като форма на оптимизация, тъй като води до по-висока степен на съответствие, минимизира правните и оперативните рискове, автоматизира ръчни дейности и допринася за стратегическата устойчивост на организацията.

Нормативната среда е сред най-динамичните фактори, влияещи върху управлението на бизнес процеси. Законодателните промени и новите регулаторни изисквания могат значително да повлияят на организационните операции, като наложат модификации в съществуващите процеси или дори създаването на изцяло нови бизнес практики. Въпреки това, анализът и интеграцията на тези нормативни изменения в бизнес процесите често остават ръчен и трудоемък процес, който крие риск от пропуски, несъответствия и забавяния в изпълнението. Липсата на автоматизирани механизми за интерпретация на законовите текстове и тяхното съотнасяне към съществуващите бизнес процеси значително затруднява организациите в постигането на регулаторно съответствие и оптимална ефективност.

Съвременните технологии в областта на изкуствения интелект, и по-специално генеративният изкуствен интелект (Generative AI) и големите езикови модели (LLMs), предоставят нови възможности за автоматизация в управлението на бизнес процеси. Тези модели имат потенциала да анализират големи обеми от текстови данни, да идентифицират скрити зависимости и да генерират структурирани бизнес правила въз основа на нормативните документи. Включването на Generative AI в този контекст не само ускорява процеса на адаптация към нормативните изисквания, но и представлява форма на интелигентна оптимизация, която намалява разходите, увеличава ефективността и ограничава риска от човешки грешки.

Настоящото изследване се основава на тези тенденции и разглежда методите за автоматично генериране на предложения за оптимизация на бизнес процеси чрез комбинираното използване на големи езикови модели и нормативни документи върху съществуващи модели на бизнес процеси. Предложеният иновативен подход се фокусира върху анализа на регулаторни текстове за извличане на бизнес правила и генериране на предложения за адаптация на съществуващите модели на бизнес процеси. Тази адаптация, разбрана в разширен смисъл, надхвърля обикновеното съответствие със закона – тя представлява оптимизация на процесите в съответствие с външната нормативна среда, с цел повишаване на устойчивостта, ефективността и оперативната съвместимост.

Същата тази методология може да бъде успешно приложена и към други източници на текстови данни. Такива източници могат да включват обратна връзка от клиенти, мнения и препоръки на служители, вътрешни фирмени политики, индустриални стандарти и дори неструктурирани данни от социални медии и форуми. Тази гъвкавост разширява обхвата на автоматичното генериране на предложения за оптимизация, като позволява на организациите да адаптират своите процеси не само спрямо регулаторните изисквания, но и в отговор на динамично променящите се потребителски очаквания и вътрешнофирмени нужди.

Актуалността на проблема се обуславя от нарастващата сложност на регулаторната среда и необходимостта от бърза адаптация на бизнес процесите към нея. В този контекст изследването има значителна научна и практическа стойност, тъй като предлага методологична рамка за прилагане на Generative AI в BPM и демонстрира

потенциала на автоматизираното оптимизиране на бизнес процеси. Това е от съществено значение за ефективното и устойчиво управление на организациите в условията на динамична и висококонкурентна бизнес среда.

2. ОБЕКТ И ПРЕДМЕТ НА ИЗСЛЕДВАНЕТО

Обект на изследването са концепциите, методите и технологиите, които стоят в основата на автоматизираното генериране на препоръки за оптимизация на бизнес процеси. Изследването се съсредоточава върху интеграцията на съвременни подходи за обработка на естествен език, генеративен изкуствен интелект, както и архитектури и системи за управление на бизнес процеси, които имат потенциала да повишат адаптивността и ефективността на организационните структури спрямо регулаторната и пазарната динамика.

Предмет на изследването е разработването на методология за интеграция на големи езикови модели от ново поколение, векторни бази данни и автоматизирани механизми за извличане и интерпретация на регулаторна информация с цел генериране на структурирани предложения за промени в бизнес процеси. Методологията цели изграждането на технологично решение, способно да идентифицира възможности за подобрене, които осигуряват както съответствие с нормативната рамка, така и оптимизация на ефективността чрез автоматизирана адаптация.

Изследователският проблем се изразява в необходимостта от разработване на ефективни методи за интеграция между съвременни езикови модели и традиционни системи за управление на бизнес процеси. Основното предизвикателство е създаването на технологично решение, което може автоматично да анализира сложни нормативни текстове, да извлича от тях приложими бизнес правила и да ги трансформира в предложения за оптимизация. Това включва преодоляване на несъответствия между формалните изисквания на регулациите и неструктурираната им текстова форма, както и прилагането на изведените правила в контекста на реални организационни процеси. В дисертационния труд се приема разширено тълкуване на термина „оптимизация“, което включва и адаптацията към регулаторни изисквания, когато тя допринася за

ефективността, намалява риска от несъответствия и подпомага стратегическото управление на процесите.

3. ЦЕЛ И ЗАДАЧИ НА ДИСЕРТАЦИОННИЯ ТРУД

Целта на дисертационния труд е да се дефинират и разработят методи, както и да се предложи интегрирана методология за проектиране и изграждане на иновативно технологично решение за автоматизирано генериране на предложения за оптимизация на бизнес процеси. Предлаганото решение трябва да бъде базирано на интеграция на съвременни технологични компоненти, като големи езикови модели (LLMs), векторни и релационни бази данни, като предоставя структурирани, приложими и ефективни препоръки, съобразени, както със съществуващите регулаторни изисквания, така и с вътрешноорганизационните специфики.

Задачите, които се изпълняват в процеса на реализация на поставената цел на дисертационния труд са:

1. **Формулиране на концептуална рамка за интеграция на нормативни документи в BPM системи**

Дефиниране на концептуална архитектура за обработка на нормативни документи, обхващаща:

- Фрагментиране на текстово съдържание в отделни смислови единици;
- Извличане на релевантни фрагменти на база семантична значимост;
- Генериране на бизнес правила от идентифицираните текстови сегменти;
- Генериране на предложения за оптимизация на бизнес процеси, на основата на дефинирани бизнес правила.

2. **Анализ и систематичен преглед на съществуващи публично достъпни колекции от модели на бизнес процеси**

- Проучване на научни публикации и практически източници, за да се открият наличните публично достъпни колекции от процесни модели;
- Задълбочен критичен анализ и систематична оценка на процесните модели;

- Дефиниране и прилагане на критерии за филтрация, съобразени с изискванията на текущото изследване;
- Сформиране на представителна извадка за по-нататъшен експериментален анализ.

3. Изследване и дефиниране на методи за фрагментиране на текстови документи

- Анализ на съществуващите методи за фрагментиране на текстови документи на логически и семантично обособени сегменти;
- Дефиниране на четири различни метода за семантично фрагментиране и векторизация;
- Дефиниране на метод за оценка на методите за фрагментиране, на база на получените резултати.

4. Експериментално имплементиране и тестване на методите за фрагментиране

- Разработване и прилагане на четирите дефинирани метода за фрагментиране и векторизация в контекста на нормативни документи на български език.

5. Оценка на методите за фрагментиране въз основа на генерирани бизнес правила

- Оценка на четирите метода за фрагментиране и векторизация на базата на автоматично извлечените бизнес правила, чрез използването на AI-базирани оценители с предварително дефинирани критерии;
- Анализ на оценките и извеждане на заключения относно най-подходящия метод за обработка на нормативни документи.

6. Дефиниране на метод за генериране на предложения за оптимизация на бизнес процеси и експериментално тестване

- Дефиниране метод, които включва голям езиков модел за генериране на предложения за оптимизация на бизнес процеси;
- Селекция на шест големи езикови модела (както платени, така и с отворен код) въз основа на релевантни показатели за тяхната ефективност;
- Провеждане на експериментални тестове за анализ на тяхната производителност и релевантност при генерирането на предложения за оптимизация.

7. Разработване на метод за оценка на генерираните предложения чрез AI-базирани оценители

- Използване на три водещи големи езикови модела като AI оценители за анализ на качеството на генерираните предложения за оптимизация;
- Дефиниране на оценъчни критерии като пълнота, последователност, съгласуваност и коректност.

8. Анализ на резултатите и изграждане на интегриран модел

- Оценка на ефективността на избрания метод за фрагментиране въз основа на стандартната девиация и консистентността на AI оценителите;
- Оценка на стабилността и предвидимостта на различните LLM модели за определяне на най-подходящия модел за генериране на предложения за оптимизация;
- Изграждане на интегриран модел (фреймуърк), който включва най-подходящите методи и технологии за изграждане на приложение за генериране на предложения за оптимизация на бизнес процеси, на базата на проведените оценявания.

9. Формулиране на методологична рамка за автоматизирано извличане на бизнес правила и предложения за оптимизация

- Разработване на цялостна, възпроизводима и мащабируема методологична рамка, съставена от ясно дефинирани етапи, които обединяват процеси по обработка на нормативни документи, използване на езикови модели, оценка на резултати и генериране на оптимизационни предложения.

10. Формулиране на препоръки и насоки за бъдещи разработки

- Обобщаване на резултатите и доказване на валидността на хипотезите, заложи в дисертационния труд.
- Идентифициране на възможности за надграждане и интеграция на предложената методология в BPM системи.

За постигане на поставената цел и изпълнение на горепосочените задачи, дисертационният труд следва иновативен подход, който комбинира технологии за обработка на естествен език, автоматизация и управление на бизнес процеси. Очаква се

резултатите от дисертационния труд да допринесат за разширяване на възможностите за автоматизирана обработка на нормативни документи и тяхното прилагане при оптимизацията на бизнес процеси, като по този начин се улеснява адаптацията на организациите към динамично изменящата се нормативна среда.

4. РАБОТНИ ХИПОТЕЗИ НА ДИСЕРТАЦИОННИЯ ТРУД

Основните работни хипотези за настоящия дисертационен труд са:

Хипотеза I: Взаимодействие на технологиите за оптимизация на бизнес процеси

Формулираната хипотеза предполага, че комбинираното използване на големи езикови модели (LLMs), векторни бази данни и релационни бази данни може да създаде ефективна методологична рамка за автоматично генериране на предложения за оптимизация на бизнес процеси. Очаква се, че оптималната интеграция на предложените методи да доведе до по-добра семантична разбираемост на нормативни документи, което от своя страна да доведе до контекстуална точност и релевантност на генерираните предложения за оптимизации. Настоящото изследване ще оцени до каква степен тази интеграция допринася за генерирането на релевантни предложения за оптимизация.

Хипотеза II: Влияние на метода за фрагментиране върху качеството на генерираните предложения

Втората хипотеза изследва влиянието на различните методи за семантично фрагментиране на нормативни документи върху качеството на извлечената информация и генерираните предложения за оптимизация. Очаква се, че прецизното структуриране на текстовата информация чрез логически обособени текстови сегменти ще намали загубата на контекст и ще подобри разбирането на нормативната рамка от големите езикови модели. Следователно, правилният избор на метод за фрагментиране е критичен фактор за точността, последователността и пълнотата на автоматично генерираните бизнес правила.

Хипотеза III: Влияние на нормативната рамка върху бизнес процесите и тяхната оптимизация

Третата хипотеза се фокусира върху начина, по който нормативните актове и техните изменения влияят върху бизнес процесите и как това въздействие може да бъде обективно оценено чрез AI-базирани методи. Очаква се, че предложената методология ще покаже измеримото въздействие на законодателните промени върху структурата и логиката на бизнес процесите, като улесни адаптацията към нормативните изисквания. Чрез използването на AI оценители, изследването ще анализира до каква степен автоматизираната обработка на нормативни документи може не само да извлече релевантна информация, но и да предложи съответни адаптации на бизнес процесите.

5. НАУЧНИ И НАУЧНО-ПРИЛОЖНИ ПРИНОСИ

В резултат на проведеното изследване са формулирани съществени приноси в научната област на управлението на бизнес процеси (BPM) с изкуствения интелект, както и в тяхното интегрирано приложение в реални бизнес среди. Приносите са класифицирани в две основни категории – научни и научно-приложни – като всяка от тях е подкрепена с емпирични резултати и теоретична обосновка.

НАУЧНИ ПРИНОСИ

1. Разработена е методологична рамка за интегрирано използване на LLMs в BPM контекст

Разработена е цялостна методология, базирана на шест последователни етапа, която съчетава техники за семантично извличане, генерация на съдържание и автоматизирана оценка. Методологията формализира прилагането на Retrieval-Augmented Generation (RAG) архитектура за задачи, свързани с управление и оптимизация на бизнес процеси, и дефинира ясно ролята на всеки компонент – от събирането на данни до избора на най-ефективния езиков модел, подходящ за генериране на предложения за оптимизация.

2. Разработени са четири метода за семантично фрагментиране и векторизация на нормативни документи

Предложени са четири метода, всеки с различен подход:

- чрез семантична близост на изречения,
- чрез семантична клъстеризация,

- чрез голям езиков модел,
- чрез комбинация на голям езиков модел и BM25 индексирание.

Методите са експериментално валидирани чрез сравнение по обективни метрики, свързани с релевантност и ефективност на извличане, като е демонстрирано, че подходът чрез голям езиков модел постига най-висока ефективност при извличане на смислово релевантни фрагменти от правни текстове.

3. Разработен е метод за автоматизирано извличане и формулиране на бизнес правила въз основа на нормативни източници

Предложен е иновативен метод, който комбинира семантично извлечени текстови фрагменти с потребителско запитване, за да генерира бизнес правила чрез голям езиков модел. Методът е адаптиран специално за български език и съдържание от нормативни документи, като включва и стъпка за превод на английски език с цел интеграция в мултиезикови среди.

4. Разработен е иновативен метод за автоматизирана оценка на ефективността на фрагментиране чрез концепцията LLM-as-Judge

Приложена е концепцията за използване на LLM модели в ролята на обективни оценители, което позволява стандартизирано и възпроизводимо измерване на качеството на извлечените фрагменти. Методът намалява субективността, характерна за експертната оценка, и гарантира висока консистентност на резултатите.

5. Разработен е метод за автоматизирано генериране на предложения за оптимизация на бизнес процеси

Предложен е метод, който използва като вход бизнес правила (извлечени от нормативни документи) и BPMN модели, за да формулира предложения за оптимизация. Изследването демонстрира, че LLMs могат ефективно да интерпретират и комбинират тези два типа източници в логически консистентни оптимизационни предложения.

6. Разработен е метод за оценка на генерираните предложения чрез многомоделна AI оценка

Предложен е метод за автоматизирана оценка на предложения за оптимизация, базиран на концепцията LLM-as-Judge, при който няколко големи езикови модела с разсъждаващи способности изпълняват ролята на независими оценители. Методът използва ясно дефинирани количествени (точност, съответствие) и качествени критерии (логическа обоснованост, яснота на формулировката), по които всеки AI оценител генерира числови резултати по скала от 0 до 100. Това позволява обективна и възпроизводима оценка на качеството на предложенията, както и систематично сравнение между различни LLM модели. Методът създава надеждна основа за информирано вземане на решения при избор на най-подходящ модел за генериране и внедряване на оптимизации в BPM системи.

7. Разработена е количествена методика за сравнителен анализ на резултатите от AI-базирани оценки

Предложена е унифицирана количествена методика за обективно сравнение на резултатите, получени, както от метода за оценка на фрагментиране, така и от метода за оценка на предложенията за оптимизация. Методиката включва изчисляване на средни стойности и стандартни отклонения на оценките, предоставени от няколко AI оценителя (LLMs), по ясно дефинирани критерии като точност, приложимост, съответствие с инструкции и др. Този подход позволява статистически обоснован избор на най-ефективни модели и методи в реални условия и формира надеждна аналитична основа за вземане на решения при внедряване в BPM среди.

НАУЧНО-ПРИЛОЖНИ ПРИНОСИ

1. Формиран е интегриран модел за внедряване на методологията в реална BPM среда

Изграден е интегриран модел (framework), който обединява най-ефективните методи, идентифицирани в хода на изследването – за фрагментиране, извличане, генерация и оценка. Моделът предлага цялостна структура за внедряване на разработената методология в реални BPM системи, използвайки както структурирани (BPMN), така и неструктурирани (нормативни) източници.

2. Демонстрирана е приложимостта на методологията в различни индустриални домейни чрез емпирична валидация

Методологията е приложена в контекста на четири бизнес сектора – логистика, финанси, информационни технологии и продажби. Резултатите от приложението потвърждават нейната гъвкавост, адаптивност и ефективност при работа с различни видове бизнес процеси, регулаторни изисквания и оперативни контексти.

3. Изведени са практически насоки за избор на LLM модели спрямо организационни ограничения

Извършен е сравнителен анализ между платени и безплатни LLM модели (напр. Gemini и DeepSeek), оценени по ключови критерии като точност, стабилност и достъпност. Резултатите формират практическа основа за информиран избор на подходящ езиков модел, съобразен с техническите и бюджетни ограничения на конкретна организация.

4. Доказана е приложимостта на нормативни документи като източник на структурирани оптимизационни знания

Изследването демонстрира, че нормативните документи могат да бъдат трансформирани от статични източници на изисквания в активен ресурс за автоматизирано генериране на бизнес правила и предложения за оптимизация. Това трансформира традиционния подход към регулаторно съответствие в стратегически инструмент за подобрене на процесите.

5. Показана е съвместимостта на архитектурата със съществуващи BPM системи без необходимост от значителни промени

Предложената архитектура е проектирана така, че да бъде интегрирана в съществуващи BPM системи без нужда от значителни архитектурни или процесни преустройства. Това улеснява нейното внедряване, включително в малки и средни предприятия, и минимизира необходимите инвестиции в техническа трансформация.

6. Внедрена и експериментално валидирана е цялостна архитектурна конфигурация за автоматизирано генериране на предложения за оптимизация на бизнес процеси

Разработените методи за семантично фрагментиране, извличане на бизнес правила, генерация и оценка на предложения са обединени в единна архитектурна конфигурация, базирана на принципите на Retrieval-Augmented Generation (RAG). Архитектурата е реализирана с използване на реални инструменти с отворен код (векторни бази, LangChain, LLMs) и е тествана в контролирана среда. Визуалната и функционална организация на компонентите осигурява възпроизводимост, гъвкавост и технологична приложимост на предложената методология. Това интегрирано решение демонстрира как различни методи могат да работят съвместно в последователен технологичен поток за реална бизнес употреба.

Формулираните научни и научно-приложни приноси демонстрират не само оригинален научен подход, но и висока степен на приложимост на резултатите в реална организационна среда. Те представляват основа за бъдещо развитие на интелигентни BPM системи, базирани на генеративен изкуствен интелект.

6. ОБЕМ И СТРУКТУРА НА ДИСЕРТАЦИОННИЯ ТРУД

Дисертационният труд е в обем от 290 страници, от които 233 съдържат същинското изследване, без приложенията. Структурата на дисертационния труд включва:

- Уводна част, в която се аргументира актуалността на темата, формулират се обектът и предметът на изследването, целта и задачите, както и изследователските хипотези;
- Шест глави, в които последователно се анализира съществуващото състояние в областта на управление на бизнес процеси, представя се концептуалната и технологичната архитектура на разработеното решение, описват се използваните набори от данни, разработените методи за извличане на информация, генериране на бизнес правила и предложения за оптимизация, както и се дефинира и обосновава цялостна методология за автоматизирано генериране на предложения за оптимизация на бизнес процеси с помощта на големи езикови модели (LLMs);
- Заключение, в което се обобщават резултатите от изследването, формулирани са научните и научно-приложните приноси, представени са насоки за бъдеща

работа, както и са включени съпътстващи списъци – на използваната литература, публикациите по темата, фигурите, таблиците, термините и съкращенията.

Първата глава представя теоретичните основи на управлението на бизнес процеси, като се разглеждат основни понятия, характеристики, етапи от жизнения цикъл, моделни нотации (BPMN), системи за управление (BPMS), както и стратегии за оптимизация и моделиране.

Втората глава описва концептуалната и технологичната архитектура, върху която се изгражда предложеното решение. Представени са както логическите взаимовръзки между компонентите, така и използваните технологии: релационни и векторни бази данни, хранилища за данни, уеб приложение, големи езикови модели, включително модели с разсъждаващи способности.

Третата глава се фокусира върху използваните набори от данни – колекция от BPMN модели от хранилището SAP Signavio Academic Models (SAP-SAM) и селектирани нормативни документи. Описани са процесите по филтрация, класификация и избор на представителна извадка от модели, използвани в експерименталната част.

В **четвъртата глава** са представени методите за извличане на информация и генериране на бизнес правила въз основа на анализ на нормативни източници. Разработени са и сравнени различни подходи за фрагментиране и векторизация на текст, включително чрез големи езикови модели и методи за индексно търсене (BM25). В тази част е въведен и метод за оценка на качеството на фрагментиране и векторно представяне чрез използване на големи езикови модели като оценители (LLM-as-Judge), с което се постига обективно измерване на семантичната адекватност на извлечените фрагменти.

Петата глава се съсредоточава върху автоматичното генериране на предложения за оптимизация на бизнес процеси въз основа на извлечените фрагменти и генерираните бизнес правила. Представен е подход за оценка на предложенията чрез LLM-as-Judge, както и сравнителен анализ на шест различни големи езикови модела – платени и с отворен достъп – по отношение на тяхната ефективност и качество на резултатите. Направеното сравнение допринася за оценка на приложимостта на моделите в контекста на бизнес анализи и подпомага избора на подходящ модел според нуждите и ресурсите на конкретна организация.

Шестата глава формулира и систематизира изведените резултати под формата на цялостна методология за автоматизирано генериране на предложения за оптимизация на бизнес процеси. Представена е поетапна структура на методологията, както и интегрирана архитектура за нейната реална имплементация.

Заключението обобщава приноса на изследването, формулира основните научни и научно-приложни резултати и очертава възможностите за бъдещо развитие и разширяване на разработения подход.

КРАТКО ИЗЛОЖЕНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

1. УПРАВЛЕНИЕ НА БИЗНЕС ПРОЦЕСИ

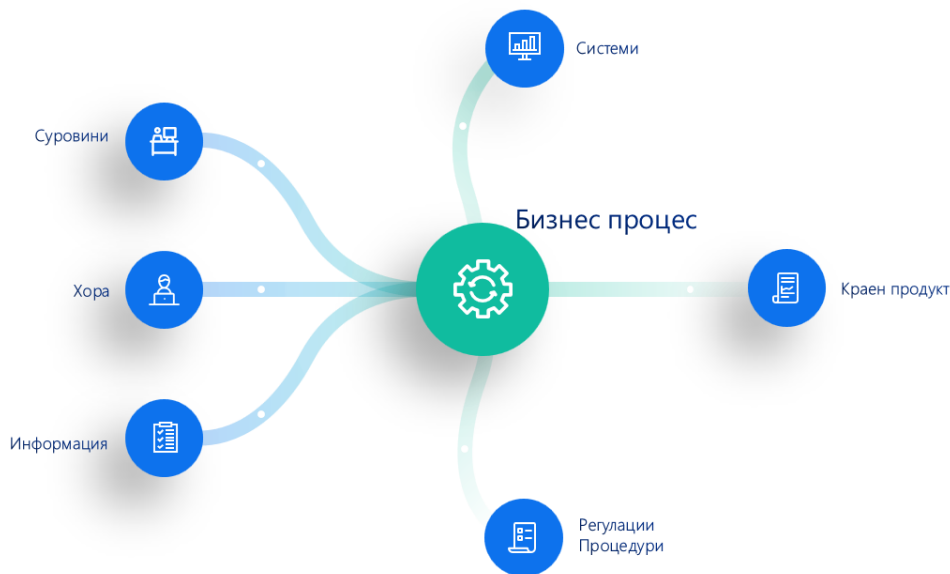
Първа глава на дисертационния труд се фокусира върху концепцията за управление на бизнес процеси (BPM), като представя ключовите теоретични основи, методи, инструменти и стратегии за тяхната оптимизация. Чрез систематично изследване на еволюцията на BPM и съвременните технологични тенденции, главата изгражда основата за предложената методология в последващите части на дисертацията.

В основата на управлението на бизнес процеси стои самото разбиране за това какво представлява бизнес процесът – неговата структура, компоненти и роля за организационната ефективност. Именно затова първоначално се разглежда същността на бизнес процесите, тяхната вътрешна логика и значение за постигане на стойностни резултати в рамките на съвременните предприятия.

Бизнес процесите представляват логически свързани дейности, насочени към трансформация на входни ресурси – като суровини, информация или капацитет – в резултати с добавена стойност под формата на продукти, услуги или решения. Те са в основата на оперативната дейност на всяка организация и допринасят за постигането на нейните стратегически цели.

Основните елементи на бизнес процеса включват: входове (ресурси), процес на трансформация (операции и действия), изходи (резултати) и контролни механизми, които гарантират ефективност, съответствие и устойчивост. Процесите не са просто последователности от задачи, а интегрирани системи за координация между ресурси, участници и регулаторни изисквания. Именно тази структурираност и възможност за

управление ги превръща в ключов фактор за организационната ефективност и конкурентно предимство.



Фигура 1. Съставни части на бизнес процес

Управлението на бизнес процеси (BPM) представлява систематичен и цялостен подход за идентифициране, анализ, моделиране, внедряване, наблюдение и непрекъснато оптимизиране на организационни процеси. Целта му е да повиши ефективността, ефикасността и адаптивността на организациите в условията на динамична пазарна среда.

BPM обхваща, както автоматизирани, така и неавтоматизирани дейности и изисква устойчив ангажимент от страна на ръководството и служителите. Чрез съгласуване на процесите с организационните цели и прилагане на подходящи технологии и ресурси, се осигурява по-добро използване на капацитет, намаляване на грешки и по-висока удовлетвореност на клиентите.

В научната литература BPM се разглежда не само като средство за автоматизация, а като стратегическа парадигма, интегрираща хора, процеси и технологии с оглед постигане на устойчиво организационно развитие. Това го превръща в ключов инструмент за повишаване на конкурентоспособността и стратегическата гъвкавост на съвременните организации.

Управлението на бизнес процеси (BPM) има за цел постигането на организационна ефективност, адаптивност и съответствие с вътрешни и външни изисквания. Основните цели включват:

- Подобряване на ефективността чрез елиминиране на излишни дейности, намаляване на разходи и оптимизиране на използваните ресурси;
- Повишаване на гъвкавостта, позволяващо бърза адаптация към пазарни и технологични промени;
- Гарантиране на регулаторно съответствие чрез стандартизирани и проследими процеси;
- Насърчаване на иновациите чрез интеграция на нови технологии и процесен реинженеринг;
- Повишаване на удовлетвореността на клиентите чрез по-добро качество и скорост на предоставяните услуги.

Приложението на BPM води до редица стратегически ползи за организацията:

- Съществено намаляване на разходите чрез оптимизация на операциите;
- Повишаване на производителността посредством автоматизация на рутинни задачи;
- Последователно и предсказуемо качество чрез стандартизирани процеси;
- По-добро вземане на решения, базирано на данни и мониторинг в реално време;
- Улеснено вътрешно сътрудничество чрез ясно дефинирани роли и отговорности.

Управлението на бизнес процеси (BPM) преминава през няколко основни етапа, всеки от които е тясно свързан с развитието на технологиите и промените в управленските парадигми.

- **Дигитализация на бизнес процеси**

Първоначалната фаза се характеризира с автоматизация на рутинни дейности чрез BPM системи, ERP платформи и формализиране на бизнес правила. Основните цели са елиминиране на оперативни неефективности, стандартизация и мащабируемост. Въпреки постигнатите ползи, този подход се ограничава от липсата на динамична адаптивност към данни в реално време.

- **Process mining подход**

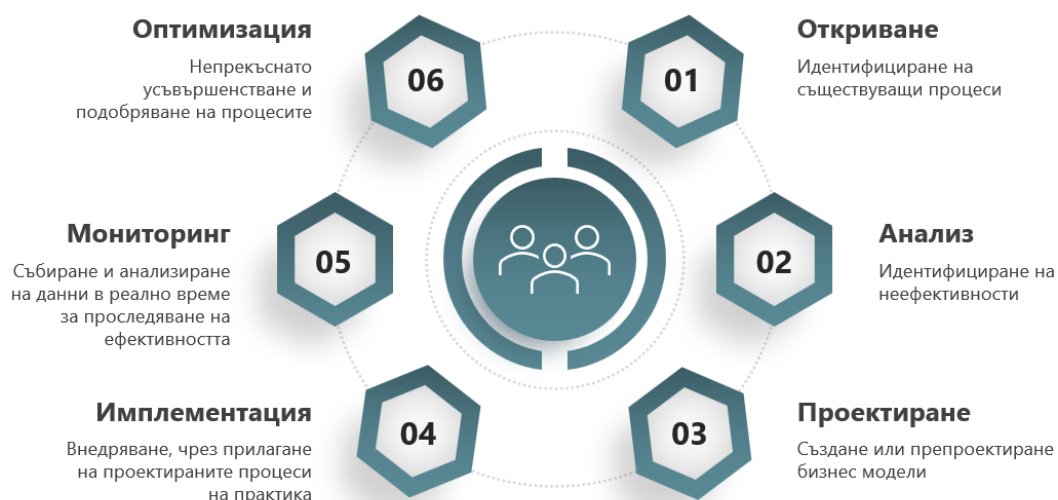
Следващата еволюционна стъпка въвежда анализ, базиран на реални събития, регистрирани в информационните системи. Process mining позволява откриване на реални процесни последователности, проверка на съответствие с моделите и прогнозен анализ. Този подход превръща BPM в динамична, управлявана от данни практика, фокусирана върху непрекъснатото подобрене.

- **Изкуствен интелект и Generative AI**

Съвременното развитие на BPM се свързва с използването на големи езикови модели (LLMs) и Generative AI. Те предоставят възможности за интелигентно препроектиране на процеси и адаптивна автоматизация. BPM вече не е статична дисциплина, а се трансформира в система, способна на съвместно проектиране с участието на AI, минимизираща човешката намеса и осигуряваща висока адаптивност и ефективност.

Тази еволюция очертава прехода на BPM от традиционен инструмент за моделиране към гъвкава, ориентирана към знанието платформа, съчетаваща човешка експертиза и изкуствен интелект в реализацията на устойчиви организационни трансформации.

Жизненият цикъл на BPM представлява структурирана и итеративна рамка, която ръководи системното подобрене на организационните процеси. Той включва шест основни фази, всяка от които има ключова роля в постигането на стратегическите цели на организацията.



Фигура 2. Жизнен цикъл при управлението на бизнес процеси

Началната стъпка включва разпознаване и картографиране на критичните процеси, определяне на техните граници, цели и заинтересовани страни. Използват се нотации като BPMN за визуализиране на текущото състояние.

- **Анализ на процеси**

Извършва се оценка на ефективността чрез събиране на количествени и качествени данни. Идентифицират се неефективните точки от процеса, дублиращи се дейности и възможности за оптимизация чрез методи като „5-те защо“ и „рибена кост“.

- **Проектиране и моделиране**

На база анализа се разработват нови или се препроектират съществуващи процеси. Чрез симулации и използване на BPM инструменти се валидират предлаганите подобрения.

- **Имплементация**

Реализира се внедряване на оптимизираните процеси чрез BPMS, включително автоматизация, системна интеграция и обучение на персонал. Провеждат се пилотни тестове за валидиране на ефективността.

- **Мониторинг и контрол**

Прилагат се показатели за изпълнение (KPIs) и системи за мониторинг в реално време за проследяване на резултатите и своевременно откриване на отклонения. Данните се използват за вземане на информирани управленски решения.

- **Оптимизация**

Въз основа на резултатите от мониторинга се извършват итеративни подобрения. Включват се съвременни технологии като автоматизация на роботизирани процеси (RPA), машинно обучение и изкуствен интелект за повишаване на ефективността и адаптивността. Насърчава се културата на непрекъснато усъвършенстване.

Жизненият цикъл на BPM не представлява еднократен процес, а динамична практика, която осигурява устойчивост и стратегическа гъвкавост в постоянно променяща се бизнес среда. За да се постигне ефективно управление на тези процеси, е необходимо използването на формални подходи за тяхното моделиране и технологични средства за тяхната реализация. В този контекст се откроява **Нотацията за управление на бизнес**

процеси (BPMN), която представлява стандартизиран език за графично описание на бизнес процеси и служи като универсален посредник между бизнес експерти и ИТ специалисти. Разработена с цел да улесни разбирането и изпълнението на процесите, BPMN 2.0 предлага не само визуализация, но и изпълнима структура, подходяща за интеграция в съвременни системи за управление.

Наред с това, **системите за управление на бизнес процеси (BPMS)** представляват технологична рамка, която поддържа пълния жизнен цикъл на BPM. Те осигуряват автоматизация, мониторинг и адаптация на процесите в реално време, като обединяват съществуващи системи в единен, процесно ориентиран модел. Чрез своите функционалности BPMS улесняват не само управлението на вътрешни и междуорганизационни процеси, но и създават условия за стратегическо развитие, дигитализация и отговорност спрямо нормативни изисквания. По този начин BPMN и BPMS се утвърждават като ключови елементи в съвременното управление на бизнес процеси и поставят основата за по-нататъшно надграждане чрез иновативни технологии като изкуствен интелект и генеративни езикови модели.

В този контекст **моделирането на бизнес процеси** се утвърждава като основен инструмент за визуализиране, анализ и последваща оптимизация. Чрез използване на формализирани нотации като BPMN, процесите могат да бъдат представени по структуриран начин, улесняващ както комуникацията между участниците, така и тяхното автоматизирано изпълнение. Моделите не само дават възможност за ясно дефиниране на дейностите и логиката зад процеса, но и създават основа за прилагане на алгоритмични подходи при оптимизация и контрол. Тъй като настоящият дисертационен труд използва именно BPMN 2.0 модели като вход за автоматично генериране на предложения за подобрене, значението на тази стъпка в жизнения цикъл е от съществено значение за изграждането на цялостната методологична рамка.

Веднъж моделирани, бизнес процесите могат да бъдат оптимизирани чрез прилагането на разнообразни **стратегии, насочени към повишаване на ефективността, производителността и адаптивността**. Сред основните подходи се открояват премахването на дейности, които не носят стойност, автоматизацията на рутинни операции чрез ИТ решения, както и преосмисляне и реструктуриране на последователността и обема на задачите. Възможностите за обединяване или

разделяне на дейности, както и тяхното паралелно изпълнение, предоставят допълнителни механизми за съкращаване на времето за изпълнение и по-рационално използване на ресурсите. Тези методи не само подобряват процесната динамика, но и повишават качеството на крайните резултати, като създават устойчиво конкурентно предимство. Те ще бъдат доразвити и използвани в предложената методология за автоматизирано генериране на оптимизационни предложения, базирана на големи езикови модели, в следващите глави на дисертационния труд.

2. КОНЦЕПТУАЛНА И ТЕХНОЛОГИЧНА АРХИТЕКТУРА

С навлизането на генеративния изкуствен интелект (Generative AI) се откриват нови хоризонти в автоматизацията и оптимизацията на бизнес процеси. Класическите подходи в управлението на бизнес процеси, макар и утвърдени, остават зависими от ръчни анализи, експертна намеса и ограничена гъвкавост, което затруднява тяхното адаптиране към динамични бизнес и регулаторни условия. Тези ограничения се проявяват най-вече при необходимостта от бързо адаптиране на процесите към променящи се нормативни изисквания и пазарна среда, където липсата на автоматизирани механизми води до забавяния и увеличени разходи.

В отговор на тези предизвикателства, големите езикови модели (LLMs), обучени чрез дълбоко машинно самообучение върху огромни корпуси от текстове, се очертават като обещаващ инструмент за интелигентна автоматизация. Чрез възможността си да интерпретират сложни текстови източници – като нормативни актове, вътрешни регламенти и индустриални стандарти – LLM моделите позволяват автоматизирано извличане на бизнес правила и формулиране на контекстуално обосновани предложения за оптимизация. Това отваря възможност за надграждане на BPM системите с ново поколение функционалност, в което регулаторната съвместимост и организационната адаптивност се реализират чрез AI-задвижван процес.

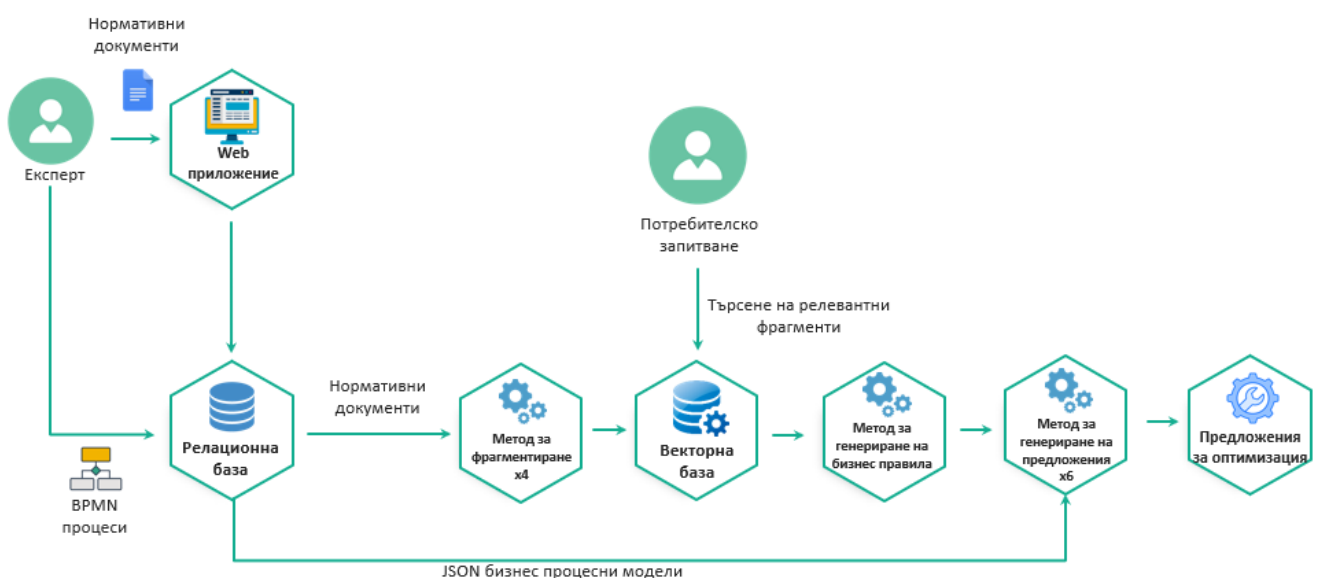
Настоящото изследване стъпва върху тази предпоставка и предлага нов архитектурен подход за внедряване на LLM модели в контекста на BPM, с цел автоматизирано генериране на предложения за оптимизация на процеси. Разработена е концептуална и технологична архитектура, която демонстрира как AI-базиран анализ на нормативни

документи може да се трансформира в практически приложими препоръки за подобрения. За тази цел е създаден експериментален прототип, който интегрира множество технологични компоненти и илюстрира приложимостта на подхода в реална среда.

КОНЦЕПТУАЛНА АРХИТЕКТУРА

В основата на предложеното решение стои концептуална рамка, която структурира основните фази и функционални компоненти, необходими за автоматизираното генериране на предложения за оптимизация на бизнес процеси на основата на нормативни документи. Рамката обхваща целия логически поток – от събирането и подготовката на данни, през обработката им с езикови модели, до извеждането на структурирани предложения, които съответстват на специфични регулаторни изисквания и организационни условия.

Предложената концепция е реализирана чрез експериментална платформа, базирана на интеграция между реляционни и векторни бази данни, големи езикови модели (LLMs) и BPMN модели, което позволява автоматизирано търсене, интерпретация и трансформация на нормативна информация в приложими бизнес правила и оптимизационни препоръки.



Фигура 3. Концептуална рамка

Концептуалната рамка включва следните ключови етапи:

1. Събиране и структуриране на данните

Първият етап включва идентифициране, събиране и систематизиране на два основни типа източници:

- Нормативни документи (закони, наредби и др.)
- Модели на бизнес процеси, представени чрез BPMN нотация

Целта е създаване на емпирична база, която да служи като основа за последващото извличане на бизнес правила и оптимизационни предложения. Данните се съхраняват в релационна база, позволяваща последователна обработка, селекция и филтрация.

2. Фрагментиране и векторизиране на нормативни документи

За да се постигне висока степен на прецизност при извличането на релевантна информация от нормативните текстове, те се обработват в два подетапа:

- Фрагментиране на текст:

Нормативните документи се разделят на смислово и логически обособени фрагменти чрез прилагане на четири различни метода за фрагментиране. Всеки фрагмент представлява самостоятелна информационна единица, пригодна за последващо векторизиране и анализ.

- Векторизиране на фрагменти:

Получените фрагменти се кодират във векторни представяния (embeddings) чрез подходящ модел за векторизация. Резултатите (семантичните вектори) се съхраняват във векторна база, което позволява ефективно търсене и филтриране на релевантни откъси в последващите стъпки.

3. Генериране на бизнес правила

В този етап, чрез използване на голям езиков модел (LLM) и векторно търсене, се извличат най-релевантните текстови фрагменти в отговор на зададено потребителско запитване (напр. въпрос относно регулаторно изискване). На база на тези фрагменти,

езиковият модел генерира структурирани бизнес правила, отговарящи на контекста и нуждите на конкретна организация.

Този процес включва:

- Дефиниране на запитване от страна на потребителя
- Векторно извличане на най-релевантни откъси от нормативната база
- Формулиране на инструкции към езиковия модел за генериране на бизнес правила

В резултат се получава списък от конкретни бизнес правила, пригодени към нуждите на организацията, които отразяват нормативната рамка и спецификите на зададения контекст.

4. Генериране на предложения за оптимизация на бизнес процеси

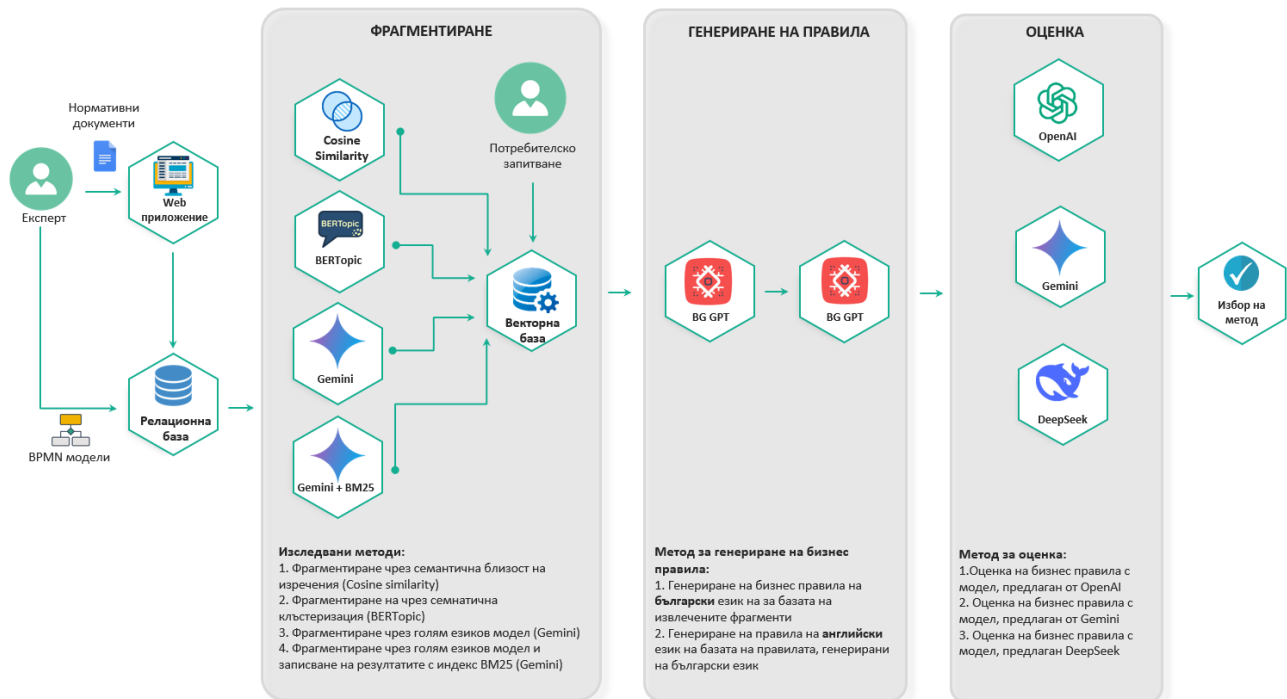
След формиране на списък от бизнес правила, концепцията предвижда тяхното прилагане към реални BPMN модели. Използвайки голям езиков модел и набор от инструкции, системата генерира конкретни предложения за промени в логиката на процеса, които са в съответствие с идентифицираните регулаторни изисквания и особеностите на дадения процесен контекст.

5. Взаимодействие между компонентите

Фигура 4 илюстрира логическата взаимовръзка между отделните етапи и компоненти: от събирането и структуриране на данни, през векторизацията и извличането на правила, до финалната фаза – генериране на предложения за оптимизация. Архитектурната логика осигурява висока степен на последователност, възпроизводимост и приложимост в реална BPM среда.

Концептуалната рамка включва два последователни и взаимосвързани компонента, чрез които се реализира пълният процес на автоматизирано генериране на предложения за оптимизация. Първият компонент обхваща събирането и структурирането на нормативни документи и BPMN модели, тяхното фрагментиране и векторизация, както и извличането и оценката на бизнес правила. Вторият компонент използва тези правила за формулиране на конкретни предложения за оптимизация на процесите и извършва

оценка на тяхната приложимост. Разделянето на архитектурата на два модула осигурява функционална яснота и проследимост на трансформациите във всяка фаза от процеса.



Фигура 4. Концептуална рамка – първи компонент

1. Въвеждане и структуриране на данни

Нормативните документи се въвеждат чрез уеб-базиран интерфейс, който съхранява метаданни в релационна база данни, докато самите файлове се записват в отделно хранилище. BPMN моделите на бизнес процеси се импортират от CSV файлове и също се структурират в релационната база. На този етап се осигурява унифицирана и достъпна среда за съхранение и последваща обработка на данните.

2. Фрагментиране на нормативни текстове

След първоначалната подготовка, нормативните документи преминават през процес на фрагментиране, чиято цел е разделянето им на логически обособени текстови единици (chunks). В рамките на разработената архитектура са приложени и сравнително анализирани четири различни подхода за фрагментиране, съобразени със спецификите на българския език и правната терминология. Всеки от тях предоставя алтернативна гледна точка върху ефективността на разделяне при обработка на регулаторни текстове.

3. Векторизация и съхранение на фрагменти

След фрагментирането, текстовите единици се преобразуват в семантични векторни представяния чрез модел за векторизация. Резултатите се съхраняват във векторна база данни, която позволява семантично търсене и сравнение между потребителски заявки и съдържанието на нормативните документи.

4. Извличане и формулиране на бизнес правила

На основата на векторно индексирани фрагменти се прилага метод за генериране на бизнес правила, който включва следните подетапи:

- Формулиране на потребителско запитване - заявка, насочена към конкретен регулаторен аспект
- Извличане на релевантна информация - потребителското запитване се векторизира и сравнява с фрагментите от нормативните документи чрез метод за векторно сходство (cosine similarity), за да бъдат извлечени 10-те най-релевантни сегмента.
- Генериране на бизнес правила - избраният голям езиков модел (Bg GPT) използва получените релевантни фрагменти, за да формира бизнес правила на български език.
- Превод и допълнителна обработка - генерираните правила на български език биват преведени на английски език, отново с помощта на големия езиков модел Bg GPT, с цел да се адаптират към други англоезични модели, с които се борави в по-нататъшни фази.

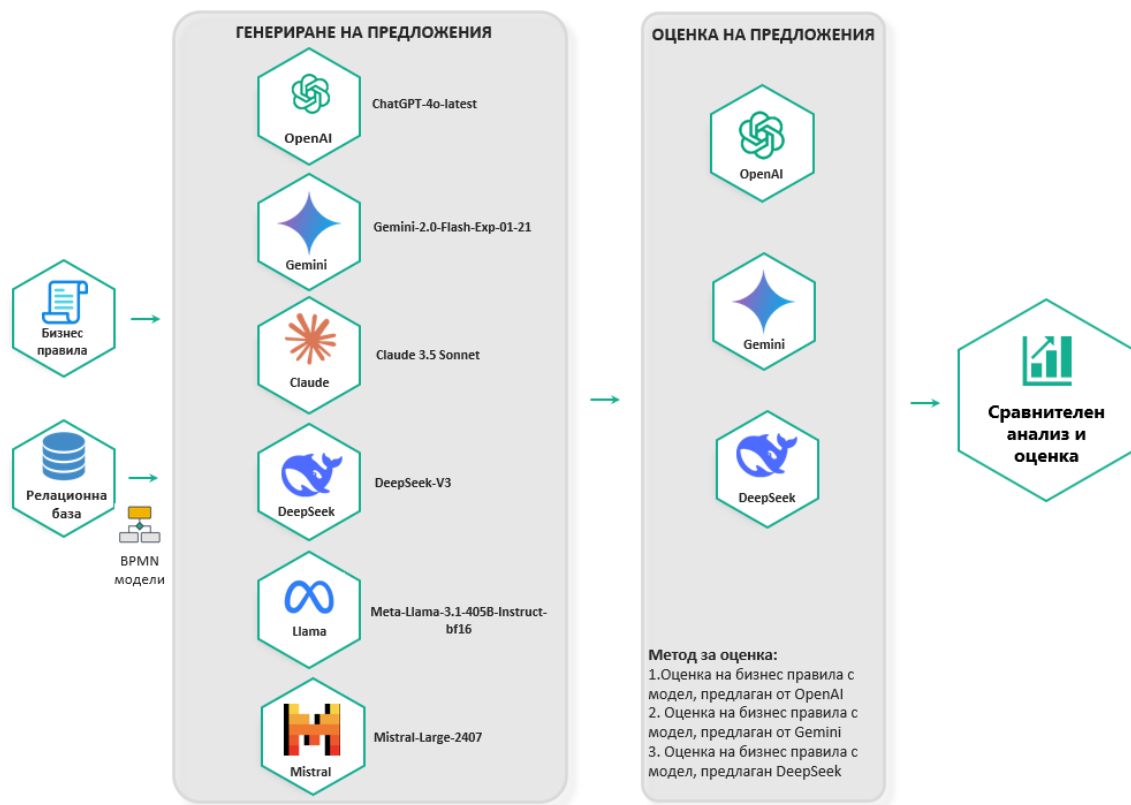
5. Оценка на генерираните бизнес правила

Оценката на качеството е критичен етап за валидиране на ефективността на генерираните правила. В тази фаза се използват три големи езикови модела с разсъждаващи способности (OpenAI, Gemini, DeepSeek), които функционират като AI-базирани оценители. Те прилагат предварително дефинирани критерии за оценка, като:

- Пълнота на формулировките
- Коректност спрямо регулаторния контекст
- Изчерпателност на съдържанието
- Степен на съответствие с инструкциите

Анализът на резултатите от тази оценка служи като основа за избор на най-подходящ метод за фрагментиране, който да бъде използван във втория компонент на архитектурата.

Вторият компонент на концептуалната архитектура е насочен към автоматизираното формулиране и оценка на предложения за оптимизация на бизнес процеси чрез използване на големи езикови модели (LLMs). Този компонент надгражда извлечените в предходния етап бизнес правила и ги използва като входни данни в процес по генериране на препоръки за подобрения на реални BPMN процеси. Архитектурата включва три основни фази: автоматизирано генериране на предложения, тяхната обективна оценка чрез AI-базирани оценители и сравнителен анализ за избор на най-ефективна конфигурация.



Фигура 5. Концептуална рамка – втори компонент

1. Автоматизирано генериране на предложения за оптимизация

На този етап се дефинира процес, при който предварително генерираните бизнес правила (на английски език), BPMN моделите и набор от инструкции се подават като вход към няколко големи езикови модела. Целта е да се създадат предложения за

подобрения на логиката на бизнес процеса, които са в съответствие с нормативните изисквания и организационния контекст.

Процесът се реализира чрез следните четири входни компонента:

- Шест големи езикови модела (LLMs) - използвани за паралелно генериране на предложения
- Бизнес правила - формулирани на база на предходно избрания метод за фрагментиране
- BPMN модел - представен в JSON формат, структуриращ текущото състояние на процеса
- Инструкции към езиковите модели - ясно дефиниращи задачата и очакваната форма на резултатите

Тази конфигурация позволява изследване на потенциала на различни езикови модели при еднакви входни условия.

2. Оценка на генерираните предложения чрез AI-базирани оценители

След генерирането на оптимизационни предложения, архитектурата предвижда прилагане на автоматизиран механизъм за оценка, осъществен чрез AI-базирани оценители – същите големи езикови модели с разсъждаващи способности, използвани и при оценката на бизнес правилата. Тази последователност гарантира съгласуваност на критериите и сравнимост на резултатите между етапите.

Всеки AI оценител прилага предварително дефинирани и теоретично обосновани критерии, сред които:

- Точност на предложението спрямо процесната логика
- Съответствие с регулаторните изисквания
- Приложимост и практическа стойност
- Яснота и пълнота на формулировката

3. Сравнителен анализ и избор на оптимален модел

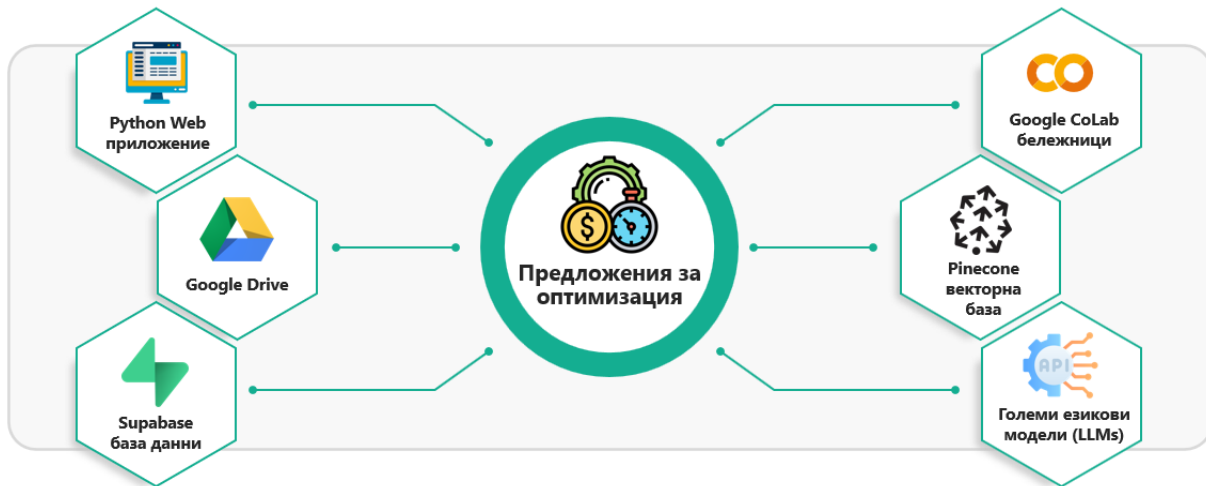
На финалния етап се извършва обобщен и количествен анализ на резултатите от AI оценителите с цел определяне на най-ефективния езиков модел. Аналитичният подход включва:

- Агрегиране на резултати от различните оценители за всяко предложение
- Изчисляване на стандартно отклонение за оценките, с цел анализ на консистентността между оценителите
- Извеждане на финален избор, базиран на средна оценка на резултатите, който определя най-подходящия LLM за внедряване в практическа среда.

Вторият компонент от концептуалната архитектура демонстрира как чрез интеграция на LLM модели и стандартизирани механизми за оценка могат да се генерират надеждни, мащабируеми и контекстуално валидни предложения за оптимизация на бизнес процеси. Този подход осигурява възможност за вземане на информирани решения относно избора на езиков модел и съответните архитектурни настройки при бъдещо внедряване в реална организационна среда.

ТЕХНОЛОГИЧНА РАМКА

Технологичната рамка на архитектурата обединява взаимосвързани софтуерни и инфраструктурни компоненти, които подпомагат основните функционалности на системата – от съхранението и обработката на текстови данни до използването на големи езикови модели и визуализация на резултатите. Подборът на технологиите е ръководен от критерии като съвместимост с архитектурата, мащабируемост, свободна достъпност и възможност за експериментално валидиране. Изградена е като експериментална конфигурация, базирана на отворен код, безплатни версии на комерсиални решения и съвременни AI модели, достъпни чрез API или локално. В следващите подраздели са описани основните технологични компоненти и тяхната роля в цялостната архитектура.



Фигура 6. Технологична рамка

- **Хранилище за данни**

В рамките на изследването, Google Drive е използван като контейнер за данни и централизирана файлова структура, осигуряваща организирано съхранение и лесен достъп до необходимите набори от данни, използвани в експерименталното проучване. Включените набори от данни са CSV файлове с BPMN бизнес модели и колекция от нормативни документи.

- **Релационна база данни**

В процеса на разработка на експерименталното приложение, предназначено за тестване и валидиране на предложената методология, се налага използването на релационна база данни за съхранение и управление на структурирана информация. Релационната база данни служи като инструмент за съхранение на извлечените от CSV файловете BPMN модели в JSON формат заедно с техните метаданни, необходими за анализ и обработка от езиковите модели. Също така съхранява пътища към нормативните документи и съответните им метаданни. За целите на разработката е избрана базата данни Supabase, чийто предимствата включват безплатен достъп, голям капацитет, автоматично генериран REST API и интеграция с векторни бази данни.

- **Уеб приложение**

Друга част от технологичната рамка на изследването е специализирано уеб приложение, разработено на Python, което осигурява интерактивен интерфейс за управление на нормативните документи, използвани по време на експерименталния анализ.

Приложението предоставя възможност за добавяне, редактиране и преглед на списъка с документи, като улеснява тяхното администриране и интеграция в процеса на обработка

- **Векторни модели**

За изграждането на цялостен поглед на използваните технологии, в научния труд са описани основните функционалности и характеристики на векторните модели. Векторните модели или понякога наричани „embedding“ модели са вид алгоритми, които преобразуват текст или други видове данни в числови вектори – т.нар. embeddings. Основната цел на тези векторни представяния е да уловят семантични и синтактични зависимости в данните, така че подобни по смисъл обекти да бъдат разположени близо един до друг в многомерното пространство, а различните – на по-голямо разстояние. По този начин сравняването на текстове, думи или изречения става възможно чрез изчисляване на сходство или разстояние между техните векторни представяния. Векторните модели могат да се разглеждат като специализирана разновидност на моделите за обработка на естествен език (NLP), с тази разлика, че те са по-лека или целенасочена версия на подобни архитектури и са оптимизирани главно за задачи по векторизиране, а не за директно генериране на текст.

Изборът на подходящ модел за векторни представяния (embeddings) е ключов етап в изграждането на Retrieval-Augmented Generation (RAG) архитектурата, тъй като той определя начина, по който текстовите фрагменти се представят в многомерно векторно пространство. Това е критично за ефективността на механизма за извличане на информация, тъй като точността на съответствията между потребителските заявки и релевантните документи зависи от способността на embedding модела да улавя семантичните връзки в текста. Неподходящият избор на модел може да доведе до неточни или нерелевантни резултати, което намалява качеството на генерираните предложения.

Работата с документи на български език поставя предизвикателство пред избора на подходящ модел, поради ограничената наличност на модели за векторно представяне на български език. За да се преодолее този проблем, е необходимо да бъде избран мултиезиков модел, който поддържа български език. В изследването е извършен анализ на мултиезикови embedding модели, които могат да работят с български език, като се

оценяват по различни архитектурни параметри. Анализът включва пет мултиезикови модела, които са оценени въз основа на критерии като размер на модела, размер на вектора, максимална дължина на входа и тип лиценз. След детайлен сравнителен анализ, Alibaba-NLP/gte-multilingual-base е избран като оптимален модел за изследването. Той предлага висока точност, баланс между ефективност и детайлност (768 измерения) и изключително голяма входна дължина (8192 токена), което позволява цялостна обработка на дълги нормативни документи.

- **Векторна база данни**

Векторните бази данни са специализирани хранилища, създадени за съхранение, индексирание и бързо търсене на векторни представяния на данни (embeddings), използвани в областта на машинното обучение и обработката на естествен език (NLP). Традиционните бази данни не са ефективни при работа с високодимензионни вектори, тъй като класическите индексирани структури не могат да осигурят бързо намиране на най-близките съседи, което е ключова задача при работа с векторни представяния. Векторните бази данни използват алгоритми за търсене на най-близки съседи (Approximate Nearest Neighbor Search – ANN) за бързо извличане на подобни вектори от големи набори от данни. За подсигурирането на този технически инструмент, в изследването, е използвана векторната база данни Pinecone, като основна технология за съхранение на векторни представяния. Изборът е базиран на нейната висока производителност, поддръжка на BM25 индекс, както и на възможността за работа с големи набори от нормативни документи.

- **Бележници (Google Colab Notebooks)**

В рамките на настоящото изследване Google Colab е избрана като основна изчислителна среда за разработка и провеждане на експериментален анализ. Google Colab представлява облачна платформа, която предоставя напълно конфигурирана среда за разработка, чрез която потребителите могат да изпълняват Python код директно в уеб браузър, без необходимост от локална инсталация на софтуер или конфигурация на хардуерни ресурси. Основните причини за избора на този инструмент са безплатният

достъп, наличието на облачна инфраструктура и възможността за използване на графични процесори без допълнително заплащане.

- **Големи езикови модели (LLMs)**

Големите езикови модели (Large Language Models - LLMs) представляват основополагащ компонент на всички дефинирани методи в настоящия научен труд. Тяхната способност да обработват, анализират и генерират текст въз основа на мащабни езикови представяния играе ключова роля в реализирането на предложената методология. В този контекст разглеждането на основните функционалности и принципи на работа на тези модели е необходимо, за да се осигури цялостно разбиране на методологичната рамка.

Големите езикови модели представляват фундаментален пробив в обработката на естествен език (NLP), като демонстрират изключителни способности за генериране, разбиране и трансформиране на текст. Тези модели са създадени чрез масивни невронни мрежи, които се обучават върху огромни текстови корпуси, като целта им е да предсказват следващия елемент в текстовата последователност. LLM моделите се характеризират с няколко ключови аспекта, които ги отличават от традиционните NLP системи:

- ✓ Голям брой параметри – повечето LLM модели съдържат стотици милиарди параметри, което им позволява да разпознават сложни езикови структури и да обработват контекст с висока точност.
- ✓ Трансформър архитектура – използват трансформър базирана невронна мрежа, която позволява ефективна обработка на дълги текстове.
- ✓ Обучение върху огромни текстови корпуси – моделите са предварително обучени върху разнообразни текстови източници, включително книги, научни статии, интернет съдържание, нормативни документи и бизнес текстове.
- ✓ Фина настройка и адаптация – след първоначалното обучение, моделите могат да бъдат дообучавани върху специфични домейни, което ги прави подходящи за експертни приложения.

Големите езикови модели представляват революционна технология, която автоматизира сложни процеси и значително улеснява обработката на текстова

информация. В рамките на настоящата дисертация тези модели играят централна роля като инструмент за анализ и оптимизация на бизнес процеси, позволявайки автоматизирано генериране на предложения за тяхното усъвършенстване.

Въпреки че LLM моделите предлагат възможност за фина настройка и дообучение в специфични домейни, настоящото изследване не разглежда тази опция. Основната причина за това е необходимостта от значителни изчислителни ресурси за провеждането на такъв процес, което би могло да ограничи неговата приложимост в реална бизнес среда, особено за организации с ограничен достъп до високопроизводителна изчислителна инфраструктура. Друг фактор обуславящ този избор е непрекъснатата актуализация на корпуса от нормативни документи. Тези динамични промени в правната рамка предполагат, че еднократното дообучаване на езиков модел върху конкретен корпус от текстове би било неефективно, тъй като бързо би загубило своята актуалност. Вместо това, представената в дисертацията методология се основава на използване на външни бази от знания и динамично извличане на релевантна информация т.нар. Retrieval-augmented generation (RAG) архитектура, която представлява по-гъвкав и икономически ефективен подход. Внедряването на външни източници на знания осигурява висока степен на актуалност на генерираните предложения за оптимизация, като същевременно намалява разходите по поддръжка и развитие на системата.

- **Големи езикови модели с разсъждаващи способности (“Thinking”)**

Развитието на големите езикови модели (LLMs) доведе до значителен напредък в обработката на естествен език, но съществуващите архитектури все още срещат предизвикателства при справянето със сложни когнитивни задачи, които изискват поэтапни разсъждения, причинно-следствен анализ и адаптивност към нова информация. В отговор на тези ограничения се появи нов клас LLM модели, наречени „мислещи (Thinking)“ модели, които демонстрират усъвършенствани способности за логически разсъждения, дълбочинно планиране и по-добро интерпретиране на контекстуални зависимости. Тези модели използват многослойна обработка на входните данни, комбинирайки техники като обосновано заключение (Chain-of-Thought reasoning), саморазсъждение (Self-reflection) и автономно търсене на допълнителна информация (Self-querying mechanisms), което ги отличава от стандартните генеративни модели.

Моделите от този тип са изградени отново върху трансформър архитектурата, но включват допълнителни механизми за подобро вземане на решения. Основните отличителни характеристики на този тип модели могат да бъдат обобщени в следното:

- ✓ Поетапно разсъждение и самооценка - вместо да генерират директен отговор, тези модели разглеждат възможни хипотези, проверяват логиката на всяка и избират най-обоснованата.
- ✓ Подобрена точност в сложни задачи - проучванията показват, че „Thinking“ моделите постигат до 20% по-висока точност в сравнение със стандартните LLMs при задачи, свързани с математически и логически разсъждения [73].
- ✓ Разбираемост и прозрачност - като документират процеса на мислене, тези модели са по-разбираеми за потребителите, особено в домейни, където се изисква обоснован анализ [74].
- ✓ Адаптивност към нови задачи - благодарение на саморазсъждението и самокорекцията, тези модели могат по-ефективно да прилагат наученото в нови контексти, без да изискват допълнително обучение.

Поради тези предимства, „Thinking“ моделите са избрани като AI оценители в настоящата методология. Основна причина за това е способността им да анализират сложни зависимости между различни части на текстовете и да преценяват логическата им последователност, пълнота и коректност. Тяхната способност за самооценка и корекция гарантира по-надеждно и обосновано оценяване, като намалява риска от пристрастни или непоследователни резултати.

3. НАБОРИ ОТ ДАННИ

Трета глава от дисертационния труд, представя използваните набори от данни, които формират емпиричната основа за разработената методология. Анализът и валидирането на предложеното решение изискват две основни групи данни: колекция от модели на бизнес процеси, които осигуряват стандартизирани процесни структури за автоматизация и оптимизация, и колекция от нормативни документи, използвани за извличане на бизнес правила и осигуряване на регулаторно съответствие. Чрез

подробно разглеждане на тези два компонента, главата обосновава избора на използваните данни и описва методологията за тяхната обработка и класификация.

- **Колекция от BPMN модели**

Моделите на бизнес процеси са основен инструмент за структуриране и управление на организационните дейности. В контекста на изследването тяхното значение се изразява в осигуряването на структурирани процесни данни, които служат за валидиране на предложената методология за автоматизирана оптимизация.

В тази част от главата се подчертава значимостта на процесните хранилища като фундаментален ресурс за научни изследвания и практическо приложение в областта на управлението на бизнес процеси (BPM). Големите колекции от модели на бизнес процеси имат ключова роля в разработването на алгоритми за анализ, машинно обучение и автоматизация. Те предоставят емпирична основа, върху която могат да бъдат валидирани иновативни методологии за оптимизация на бизнес процеси.

Въпреки големия научен интерес към BPM, изследователите често се сблъскват с ограничен достъп до реални бизнес процеси поради конфиденциалността на информацията, строгите регулаторни изисквания и стратегическата значимост на вътрешноорганизационните процеси. Тези ограничения затрудняват провеждането на емпирични изследвания и валидирането на нови методологии. В този контекст публично достъпните хранилища на бизнес процеси играят съществена роля, като позволяват на изследователската общност да работи с реални и стандартизирани данни.

За целите на настоящото изследване е избрана колекцията SAP Signavio Academic Models (SAP-SAM), която представлява най-голямата публично достъпна база с бизнес процеси. Тази колекция включва над 1,021,471 бизнес модела, събрани в периода 2011–2021 г., като значителна част от тях – 618,807 (60.58%) – са представени в BPMN 2.0, което осигурява съвместимост със съвременните инструменти за процесен анализ. Освен това, 57.43% от моделите са на английски език, което улеснява автоматизираната обработка и анализа на текстовите описания на процесите.

Включването на SAP-SAM в изследването изисква предварителна филтрация и обработка, с цел да се гарантира, че анализираниите модели отговарят на определени

критерии за качество и релевантност. В този контекст е приложена серия от методологични стъпки, включващи премахване на примерни (vendor-provided) модели, филтриране по нотация, при което са запазени само BPMN 2.0 модели, и филтриране по език, като са включени единствено модели на английски. Освен това, е извършена структурна проверка, гарантираща, че всеки процесен модел съдържа поне един начален и един краен елемент, както и оценка на сложността, при която са запазени само модели със сложност над 80%. След прилагането на тези критерии, финалният набор от анализирани процеси, включени в изследването е 69,263 бизнес модела.

За да бъде изследването по-целенасочено, процесните модели са класифицирани в 12 бизнес категории, като от най-многобройните такива са селектирани четири представителни индустриални сектора:

- Логистика – 15,472 модела
- Информационни технологии – 7,603 модела
- Финанси – 5,555 модела
- Продажби – 5,289 модела

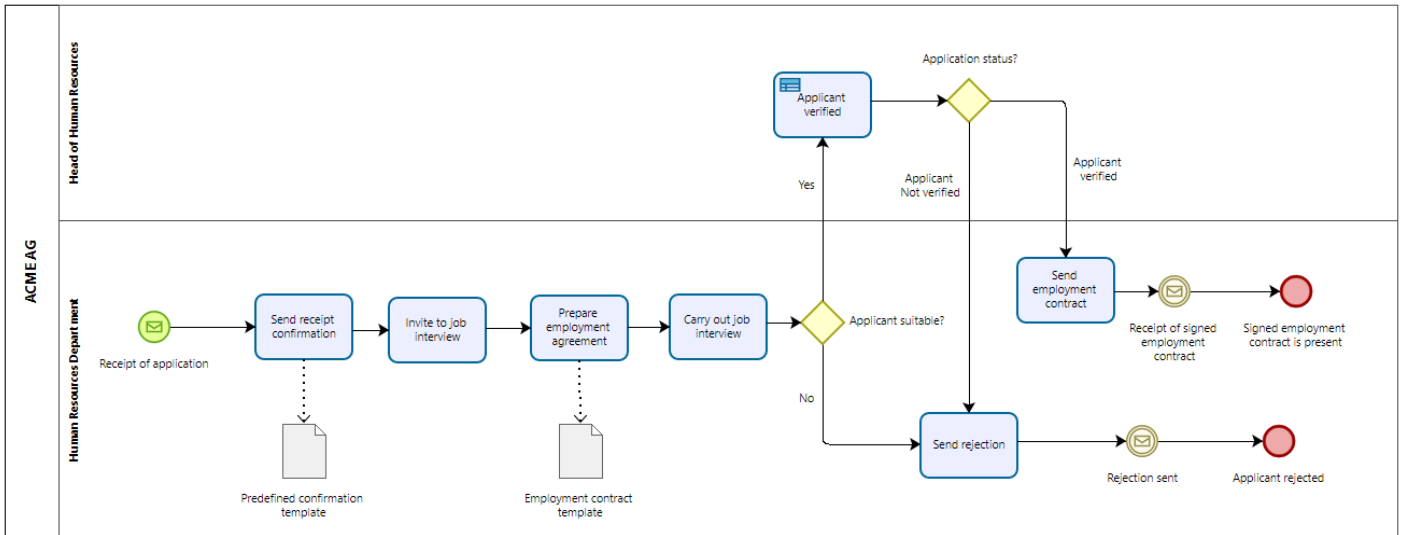
За автоматизираното класифициране на моделите по бизнес категории е използван BART-Large-MNLI – мощен NLP модел, специализиран в текстова категоризация. Неговата архитектура позволява семантичен анализ на имената и описанията на процесите, което улеснява тяхното автоматично класифициране в съответните индустриални сектори. В резултат на този анализ е осигурена по-висока аналитична стойност на изследването, тъй като бизнес процесите са структурирани и организирани по начин, който улеснява тяхното използване за валидиране на предложената методология.

Включените в изследването модели са селектирани на случаен принцип, по един от всяка категория, с цел избягване на предубеденост при избора и гарантиране на неутралност в оценката. Важно е да се отбележи, че някои от тези модели може да не отразяват напълно реални бизнес процеси, тъй като колекцията включва и процеси, разработени от студенти и преподаватели, част от които не са били прилагани в реална бизнес среда.

Важно е да се подчертае също, че начинът, по който са филтрирани, подбрани и категоризирани бизнес моделите, представлява сам по себе си принос към научната

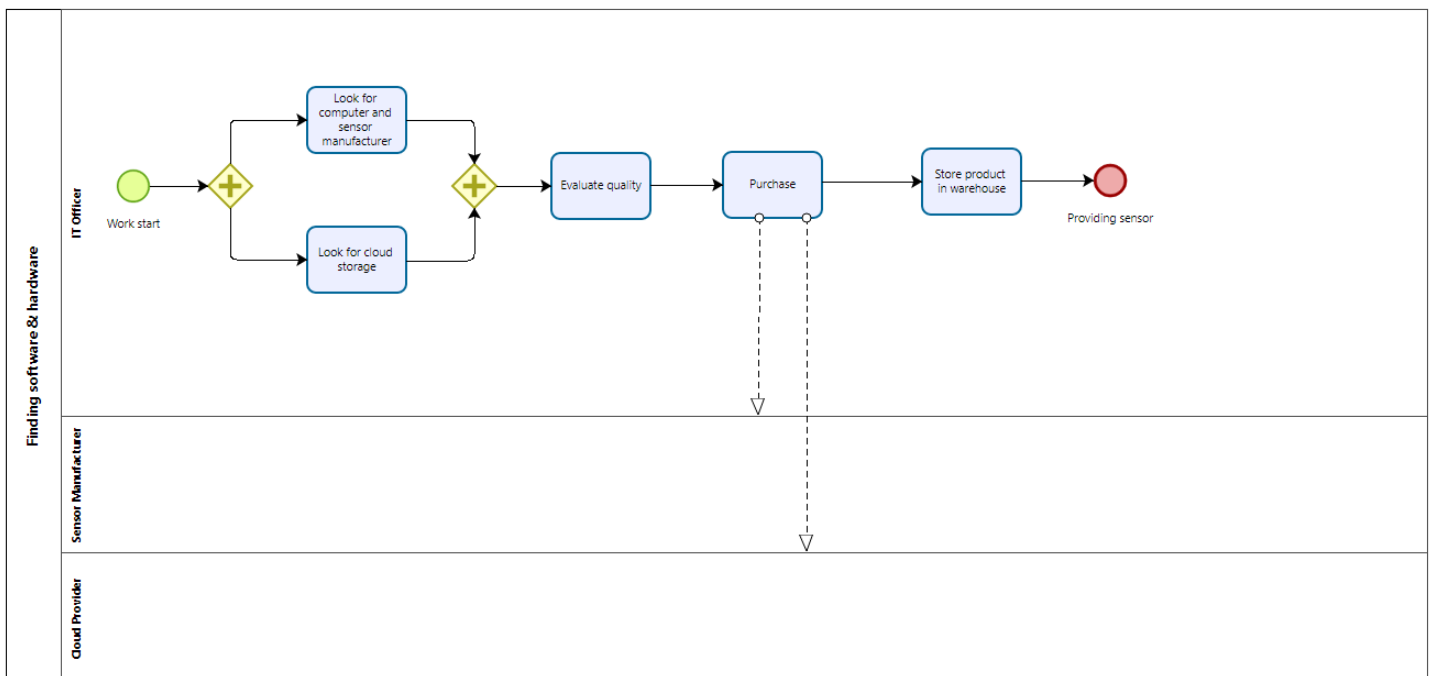
работа. Изградената методология за селекция и класификация на BPMN модели демонстрира подход, който може да бъде приложен и в бъдещи изследвания, свързани с автоматизация на бизнес процеси и използване на NLP за анализ на бизнес модели.

1. Категория „Логистика“



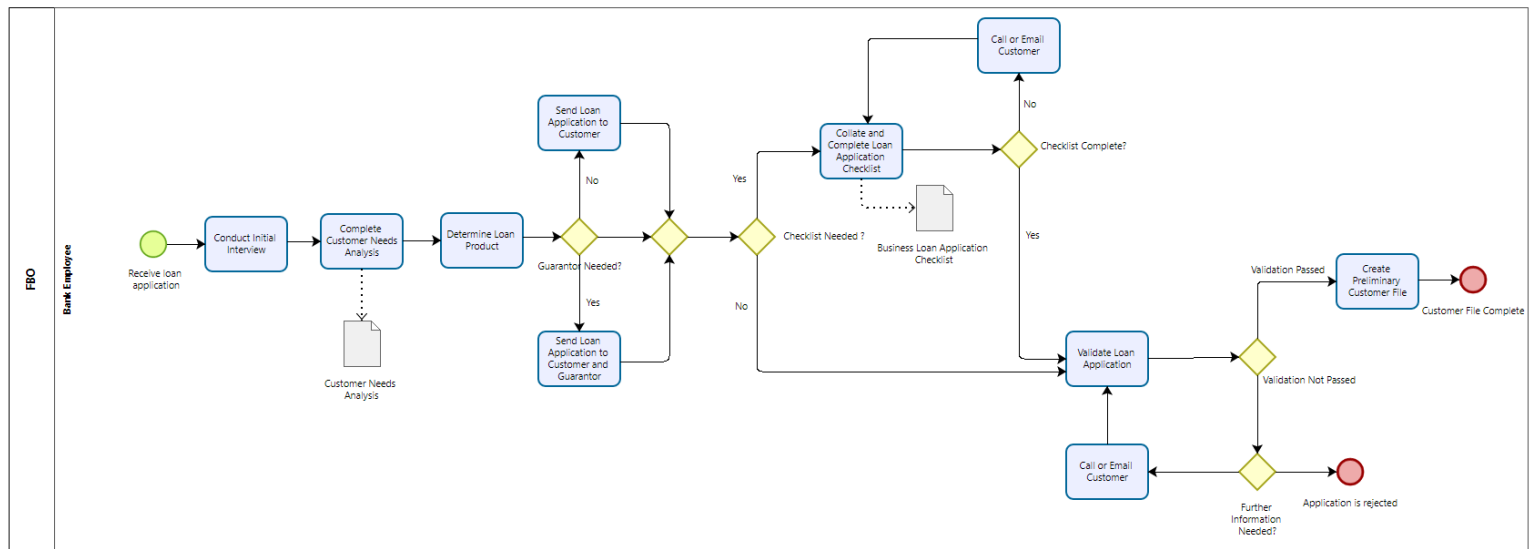
Фигура 7. Бизнес модел "Логистика"

2. Категория „Информационни технологии“



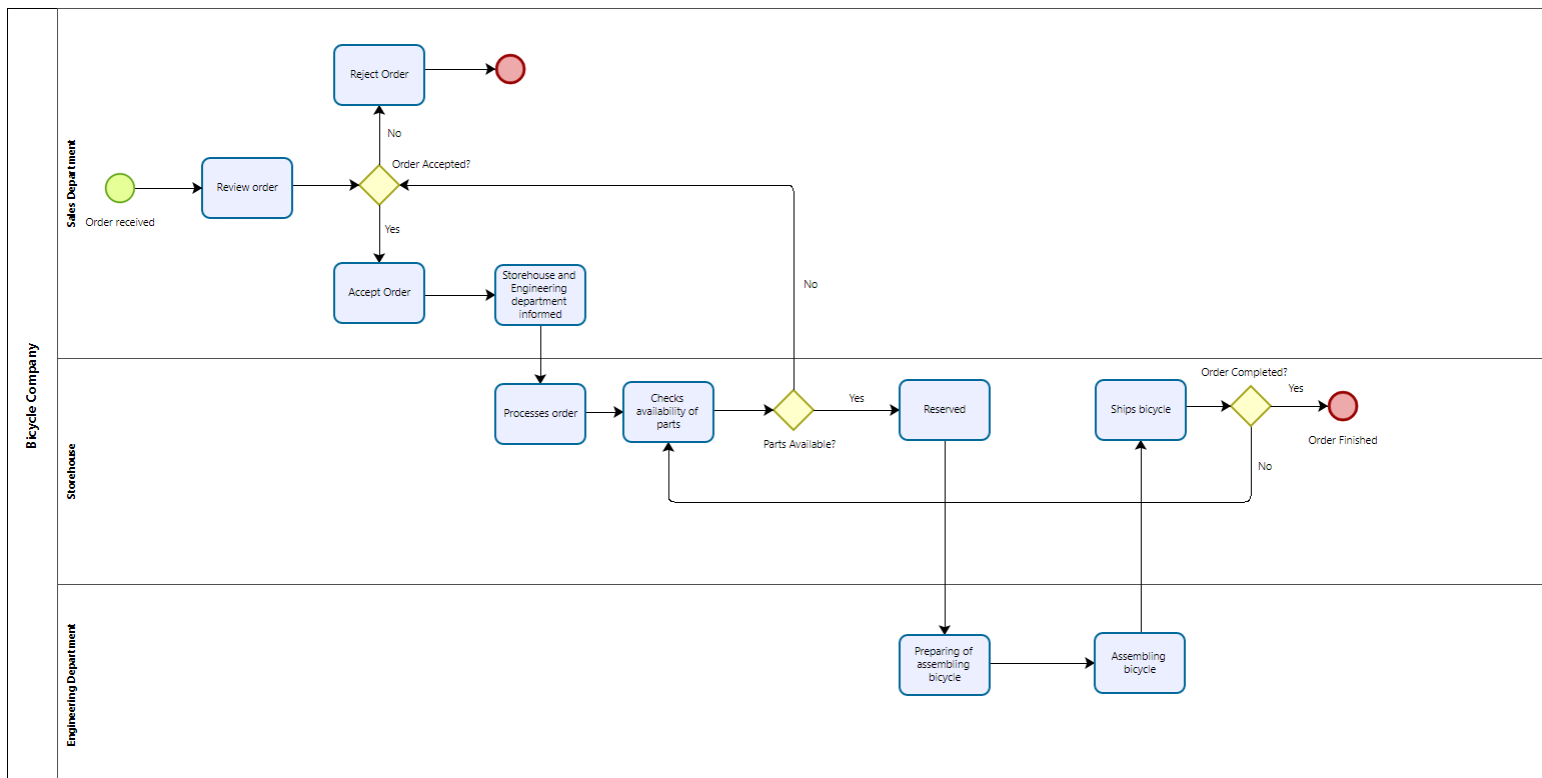
Фигура 8. Бизнес модел "Информационни технологии"

3. Категория „Финанси“



Фигура 9. Бизнес модел "Финанси"

4. Категория „Продажби“



Фигура 10. Бизнес модел "Продажби"

▪ Колекция от нормативни документи

За осигуряване на регулаторна съвместимост при оптимизацията на бизнес процеси, в изследването са включени три ключови нормативни акта, които имат широко

въздействие върху корпоративната дейност и представляват основни източници на бизнес правила:

- Закон за защита на личните данни (GDPR) – регулира обработката, съхранението и защитата на личните данни, като гарантира правата на физическите лица и поставя задължения за организациите, които обработват лична информация.
- Закон за електронната търговия – определя изискванията за дигиталните търговски операции, включително регулации относно електронните договори, защитата на личните данни в онлайн среда и потребителските права в електронната търговия.
- Закон за защита на потребителите – осигурява защита на клиентите при търговски транзакции, като регламентира гаранционните условия, прозрачността на офертите и правото на рекламация.

Изборът на тези нормативни документи е обоснован от тяхната универсалност и широка приложимост в различни индустрии, както и от факта, че те регулират критични аспекти на бизнес операциите, свързани с обработката на лични данни, дигиталните търговски практики и защитата на потребителите. Включването на тези нормативни актове в изследването осигурява валидна правна основа за автоматизираното извличане на бизнес правила и тяхното приложение в BPM средата.

Друга част от тази секция засяга въпроса с поддръжката на представената нормативната база. За да се осигури актуалността ѝ, са разгледани два подхода за поддръжка на правните документи:

- Ръчно въвеждане и управление чрез уеб-базирана система, което осигурява прецизен контрол и юридическа верификация.
- Автоматично извличане на нормативни актове чрез API интеграция, което гарантира висока честота на актуализация.

В настоящото изследване е избран първият подход, тъй като основният фокус е върху разработването и валидирането на методология за автоматизирано извличане на бизнес правила и предложения за оптимизация, а не върху изграждането на цялостна инфраструктура за управление на нормативни актове. Ръчното въвеждане и управление

осигуряват необходимата гъвкавост при обработката на регулаторната информация, без да се налага сложна интеграция с външни източници.

В тази връзка е разработено уеб-базирано приложение, което осигурява интерактивен интерфейс за управление на нормативните документи. Основните му функционалности включват:

- Добавяне на нови нормативни актове – възможност за въвеждане на закони, наредби и вътрешни корпоративни политики.
- Редактиране и актуализиране – поддържане на актуална база чрез ръчна корекция на нормативните текстове.
- Преглед и управление на съдържанието – нормативните актове са категоризирани и организирани, за да улеснят тяхното използване за автоматизирано извличане на бизнес правила.
- Автоматично записване на метаданни – за всеки нормативен документ се съхраняват ключови параметри, включително дата на публикуване, последна актуализация и приложимост в конкретни бизнес сфери.

Макар че настоящото изследване не поставя основен акцент върху инфраструктурата за динамично обновяване на нормативни актове, бъдещи разработки могат да разгледат интеграцията на автоматизирани механизми за актуализация. В този контекст, вторият подход, може да бъде имплементиран в реални бизнес приложения, за да се гарантира непрекъснатата актуалност на регулаторната информация и намаляване на необходимостта от ръчна поддръжка. Това би позволило създаването на напълно автоматизирани системи, които не само анализират и оптимизират бизнес процеси, но и се адаптират в реално време към промените в законодателната рамка.

В заключение, представените набори от данни, както и методологията за тяхната селекция и обработка, осигуряват надеждна входна база за разработеното приложение, като същевременно допринасят за усъвършенстване на изследователските подходи в областта на автоматизираната оптимизация на бизнес процеси.

4. ИЗВЛИЧАНЕ НА ИНФОРМАЦИЯ И ГЕНЕРИРАНЕ НА БИЗНЕС ПРАВИЛА

В **четвърта** глава от дисертационния труд се представя първият компонент от разработената концептуална архитектура, фокусиран върху автоматизираното извличане на информация от нормативни документи и генерирането на бизнес правила чрез големи езикови модели (LLMs). Основната цел на главата е да се опише как чрез интеграцията на RAG подход се осъществява процесът на фрагментиране на нормативни документи, тяхната векторизация, индексирание и последващо използване от генеративен модел за формулиране на релевантни бизнес правила.

Наред с архитектурното описание, в главата се разглеждат четири експериментално изследвани метода за фрагментиране и се представят резултатите от сравнителен анализ на тяхната ефективност.

ОПИСАНИЕ НА АРХИТЕКТУРАТА RAG

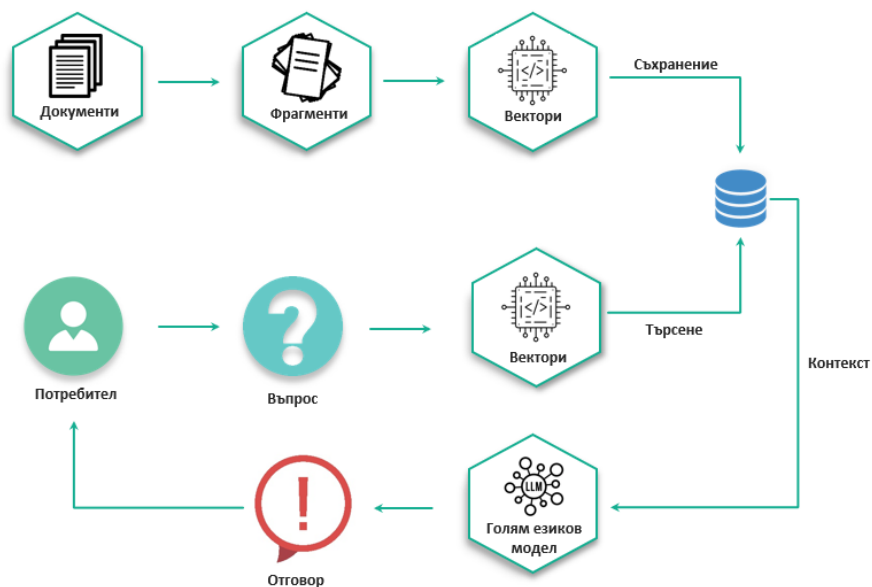
Големите езикови модели са доказали, че съхраняват фактологични знания в своите параметри и постигат съвременни водещи резултати при фина настройка за конкретни задачи в обработката на естествения език. Въпреки че демонстрират силни способности за генериране на текст, те срещат затруднения в сценарии, изискващи достъп до структурирани бази от знания или специфични документи, което ограничава тяхната ефективност в задачи, изискващи интензивно използване на знания.

Съществува подход, който предлага решение, за да се преодолее това ограничение и той включва използването на архитектура с информация от външни бази данни, известни още като Retrieval-Augmented Generation (RAG). Този вид архитектури съчетават модул за извличане на релевантна информация с генеративен модел, като позволяват на езиковия модел да работи върху предварително индексирана база от документи и да генерира отговори, базирани на реални текстови фрагменти, а не само на параметрично знание.

Архитектурата Retrieval-Augmented Generation (RAG) [32] представлява модулна система, изградена от два основни компонента, които работят в тясна взаимовръзка. Първият компонент е модулът за извличане на информация, чиято задача е да идентифицира релевантни текстови фрагменти от предварително индексирана база знания. Вторият

компонент е модулът за генериране на съдържание, който, въз основа на потребителското запитване и извлечената информация, формулира финален отговор.

Тази двукомпонентна архитектура предоставя по-висока степен на контрол върху източниците, използвани от езиковия модел при генерирането на отговор. Тя позволява ясно разграничаване между процесите по извличане на знание и създаване на текст, както и проследимост на използваните фрагменти, което е особено важно при обработка на чувствителна или нормативно обвързана информация. Основно предимство на RAG архитектурата е, че нови документи могат да бъдат динамично добавяни към базата от знания, без да е необходимо повторно преобучаване на езиковия модел.



Фигура 11. Генериране на съдържание, обогатено с извличане на информация (RAG)

Процесът започва с еднократна предварителна обработка, при която текстовите фрагменти от външната база се конвертират във векторни представления (embeddings) и се съхраняват във векторна база данни. При отправяне на заявка, системата създава вектор (embedding) на потребителското запитване чрез същия векторен модел и извършва сравнение между него и всички съхранени вектори, използвайки метрика за семантична близост – най-често скалярно произведение или косинусово сходство (cosine similarity) [33]. В резултат се получава класифициран списък с най-релевантните фрагменти, подредени по степен на сходство със заявката.

След извличането на релевантната информация, генеративният модул съставя отговора. За целта се конструира вход (prompt) към големия езиков модел, който съдържа:

- Потребителското запитване
- Извлечените фрагменти, представени в естествен език
- Допълнителни инструкции относно формата и съдържанието на очаквания отговор

Използвайки този структуриран вход, езиковият модел генерира отговор, който е контекстуално обоснован, логически свързан и базиран на документи с проследим произход.

По този начин генерирането на отговор се „обогатява“ чрез извлеченото знание, а ролята на езиковия модел се трансформира от носител на „вградено“ знание към синтезатор на външно подадена информация.

Основното предимство на този подход е, че позволява използването на външна база от знания, което намалява необходимостта цялата информация да бъде кодирана в параметрите на езиковия модел. По този начин генеративният модел може да се фокусира върху синтезиране на извлечените данни в логически свързан и последователен отговор, без да разчита изцяло на предварително научените параметрични зависимости.

За да се приложи подхода в рамките на настоящото научно изследване, е необходимо да се разгледат няколко основополагащи въпроса, свързани с ефективността на информационното извличане и генерирането на текст. Тези въпроси включват:

- Определяне на метод за фрагментиране на текст – изборът на подходящ метод за разделяне на документите на отделни текстови сегменти е от съществено значение за прецизността на извличането, като оптималният размер и начин на сегментация могат значително да повлияят върху резултатите от модела.
- Избор на модел за векторни представяния (embeddings) – необходимостта от подходяща техника за векторизиране на текстовите фрагменти, която осигурява висока семантична съпоставимост между заявките на потребителя и наличните документи.

- Избор на векторна база данни за съхранение и търсене на векторни представяния – определяне на база данни, която предоставя ефективни механизми за бързо извличане на релевантни текстови фрагменти въз основа на сходството на векторите, като се вземат предвид фактори като мащабируемост, производителност и поддръжка на търсене по близост между вектори.

Анализът на тези въпроси е от решаващо значение за ефективността и надеждността на предложената концептуална архитектура, като всеки от тези компоненти играе ключова роля за подобряване на качеството на генерираните предложения.

За да се отговори на първият въпрос, а именно избора на подходящ метод за фрагментиране, научното изследване представя и описва набор от такива, като ги разделя в две основни групи:

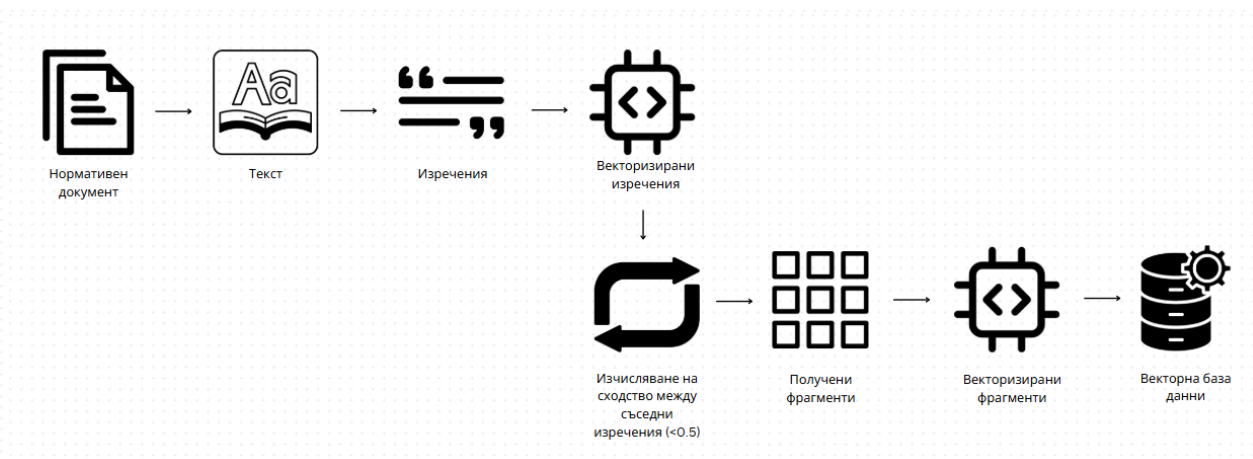
1. Синтактично или структурно фрагментиране - тази група включва методи, които разделят текста на базата на предварително зададени граматически или форматни правила, като разделяне по изречения, параграфи, заглавия, секции, подточки и други формални маркери. Тези методи са по-лесни за имплементация, но могат да загубят важни семантични връзки. Примери за такива методи са: метод за фрагментиране на базата на фиксиран брой изречения, метод за фрагментиране на базата на тематични единици и метод за фрагментиране на базата на брой думи или символи.

2. Семантично фрагментиране – тази група включва методи, които се основават на семантичната структура и контекста на обработваното съдържание по време на фрагментирането. Тези методи анализират логическите и смислови връзки в текста, осигурявайки по-прецизно разделяне на информацията.

В този контекст, изследването заключава, че традиционните методи за механично разделяне на текст на изречения или параграфи не са достатъчни, тъй като могат да разкъсат логически свързана информация и да доведат до загуба на смисъл при последващото ѝ извличане. Това поставя основата, изследването да предложи четири различни метода за семантично фрагментиране, като всеки от тях има за цел да подобри точността на откриването на релевантни правни текстове.

МЕТОД ЗА ФРАГМЕНТИРАНЕ И ВЕКТОРИЗАЦИЯ ЧРЕЗ СЕМАНТИЧНА БЛИЗОСТ НА ИЗРЕЧЕНИЯ

При този метод текстът се разделя на отделни изречения, които след това се преобразуват във векторни представления (embeddings). Чрез изчисляване на косинусово сходство (cosine similarity) между последователни изречения се оценява семантичната им близост. Когато стойността на косинусовото сходство е по-ниско от определен праг (в текущото изследване -0.5), се приема, че има значителна тематична промяна, и на това място се извършва разделяне на текста на сегменти. Този метод позволява динамично фрагментиране, съобразено със смисловите преходи в текста.



Фигура 12. Метод за фрагментиране и векторизация чрез семантична близост на изречения

Методът за фрагментиране на текст чрез семантична близост на изреченията най-общо следва няколко стъпки:

1. Извличане на текст от нормативен документ (pdf файлове)

За тази цел се използва библиотеката PyPDF2, която преминава през всяка страница на документа и извлича съдържанието ѝ в текстови формат.

2. Разделяне на текста на изречения

За разделянето на текста на изречения е избрана библиотеката Stanza NLP. Stanza е мощна Python библиотека за обработка на естествен език, предназначена за анализ и структуриране на текстови данни и разработена от университета Станфорд. Тя предлага пълен набор от инструменти, които автоматично анализират текста, като го разделят на изречения и думи, определят основната форма на думите (лематизация¹), разпознават

¹ Лематизация - процес в обработката на естествен език (NLP), при който дума се преобразува в своята основна (лема) форма. Лематизацията използва лингвистични правила и речници, за да върне граматически коректната основна форма на думата.

частите на речта, анализират граматичните зависимости между думите и идентифицират именувани обекти, като организации, дати, места и личности.

Едно от ключовите предимства на библиотеката е, че поддържа български език, което я прави изключително подходяща за обработката на нормативни документи, като закони и наредби. Това позволява по-точен анализ и извличане на информация от правни текстове, което е важно за автоматизираната обработка на регулаторни данни. [35]

3. Векторизация на изреченията

Всяко изречение се преобразува във вектор чрез предварително избрания векторен модел за целта, а именно Alibaba-NLP/gte-multilingual-base. Подробна обосновка за избора на този модел може да бъде открита в секция 5.2.3.

4. Семантично фрагментиране

Стъпка, при която последователно се обхождат изреченията и се изчислява косинусово сходност между векторните представяния на изреченията. След проведени тестове е избрана граница от 0.5, което означава, че ако сходността между две съседни изречения е под 0.5, тогава се създава нов фрагмент. По този начин текстът се разделя динамично, без фиксирани граници, което повишава точността на информацията при последващо извличане.

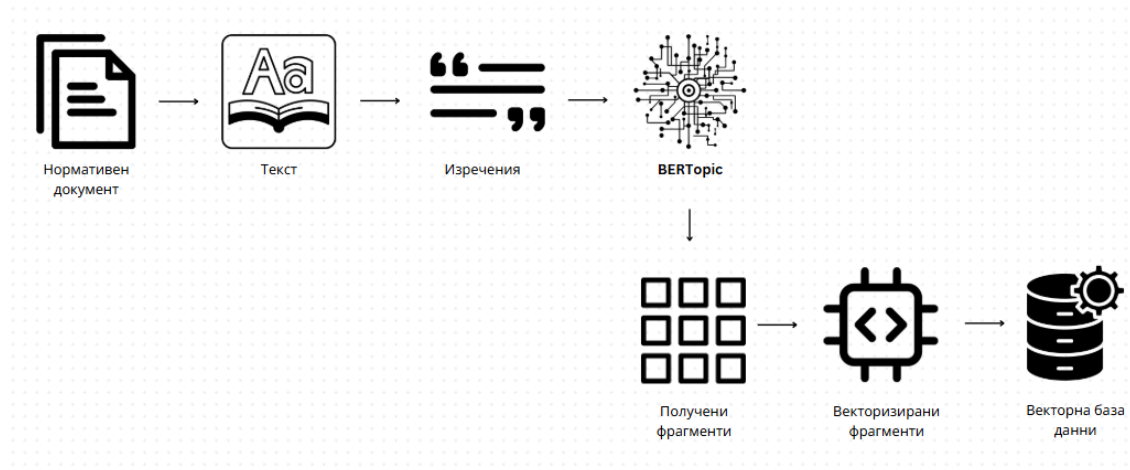
5. Векторизация на получените фрагменти и съхраняване в базата данни

Всеки един от получените фрагменти се векторизира, отново с помощта на модела Alibaba-NLP/gte-multilingual-base, след което получените вектори се съхраняват във векторната база данни (Pinecone).

МЕТОД ЗА ФРАГМЕНТИРАНЕ И ВЕКТОРИЗАЦИЯ ЧРЕЗ СЕМАНТИЧНА КЛЪСТЕРИЗАЦИЯ

Семантичната клъстеризация представлява метод за групиране на текстови фрагменти въз основа на тяхната смислова близост. Този подход комбинира трансформър-базирани модели за извличане на векторни представяния (embeddings) на текста и статистически техники за анализ на честотата и значимостта на думите. В резултат на това текстовите фрагменти се организират в тематично свързани групи (клъстър), което

позволява идентифициране на основните теми и съответно разделянето на текста на сегменти, отразяващи тяхното съдържание.



Фигура 13. Метод за фрагментиране и векторизация чрез семантична клъстеризация

При предложеният метод за фрагментиране чрез семантична клъстеризация, първите две стъпки са идентични с тези на по-горе представения метод. Текстът на документа отново бива извлечен и разделен на изречения чрез посочените библиотеки. Основната разлика между двата метода е, че тук масивът от получените изречения се подава към библиотека наречена BERTopic. BERTopic е съвременна техника за тематично моделиране, която използва семантични векторни представяния на текст за извличане и класифициране на дискурсивни теми. Разработена върху клъстеризиращи алгоритми и дълбочинни езикови модели, BERTopic се отличава с гъвкавостта си при анализ на текстове с различна структура и контекст, като е особено полезна при работа с големи текстови колекции. За разлика от традиционните методи, които разчитат основно на честотния анализ на думи, BERTopic използва трансформър² модели за извличане на богати и контекстно осмислени теми. Основните функции, които изпълнява библиотеката са две:

² Трансформър модели - невронно-мрежови архитектури, използвани основно за обработка на естествен език (NLP) и задачи с изкуствен интелект. Те са в основата на модерните NLP приложения, като модели за машинен превод, текстова генерация и търсачки, благодарение на способността си да разбират дългосрочни зависимости в текста.

1. Използва предварително обучени трансформър модели за извличане на смислови представяния на текстовите фрагменти (векторизация).
2. Използва клъстериращи алгоритми като HDBSCAN³ (Hierarchical Density-Based Spatial Clustering of Applications with Noise) за групиране на текстовите изречения в теми, като позволява идентифициране на теми, които не са предварително дефинирани в текстовите масиви.

По този начин BERTopic се доказва като по-ефективен от стандартните методи, защото позволява по-гъвкаво и адаптивно извличане на информация въз основа на контекстна семантика, а не просто на честотния анализ на думите.

В настоящото изследване BERTopic се използва като един от методите за семантично фрагментиране, като позволява оптимално групиране на тематично свързани изречения, което е от съществено значение за подобряване на крайните резултати.

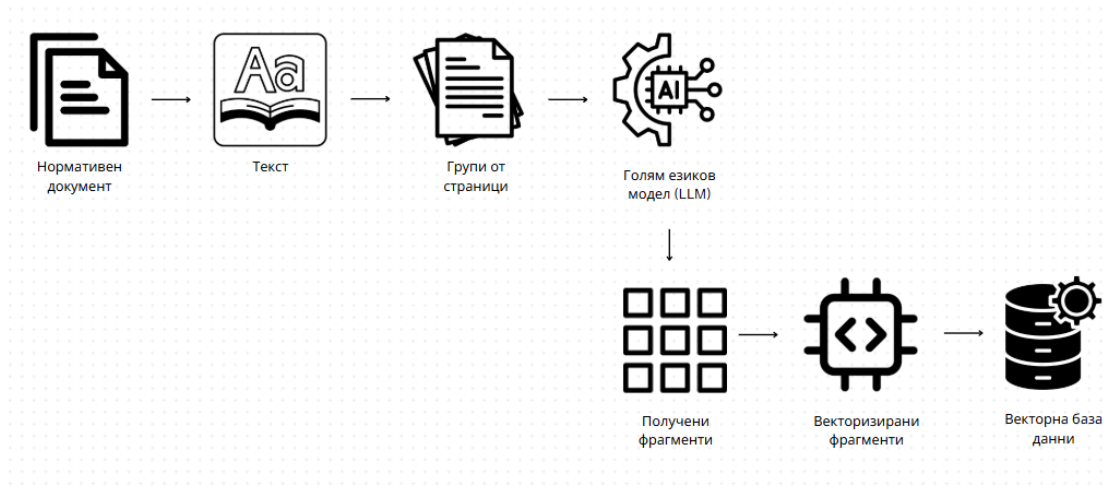
След получаването на клъстери от изречения (фрагменти) процесът отново следва стъпките от първия метод. Фрагментите биват векторизирани и съхранени във векторната база данни.

МЕТОД ЗА ФРАГМЕНТИРАНЕ И ВЕКТОРИЗАЦИЯ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ

Този метод включва използването на голям езиков модел (LLM), който получава последователно групи от страници или параграфи от текста. Моделът е инструктиран да анализира съдържанието и да идентифицира оптималните точки за разделяне на текста на сегменти въз основа на смисловите и тематични преходи. Това позволява създаването на контекстуално съобразени сегменти, които отразяват логическата структура на оригиналния текст.

Фрагментирането на текста при този метод следва малко по-различни стъпки, в сравнение с предходните два, като изключим първата стъпка за извличане на текста от документа.

³ HDBSCAN (Йерархично пространствено клъстериране на приложения с шум, базирано на плътност) – метод, който е разширение на DBSCAN. Характеризира се с това, че автоматично определя броя на клъстерите, като използва йерархична структура и плътност на данните. Този метод е особено ефективен при работа с данни с различна гъстота, без необходимост от предварително задаване на броя на клъстерите.



Фигура 15. Метод за фрагментиране и векторизация чрез голям езиков модел

Стъпките, които се включват в този метод са следните:

1. Разделяне на текста по страници и създаване на групи

На тази стъпка текстът бива извлечен от документа и разделен на групи от страници, като към всяка от образувалите се групи се добавя предходната и следващата страница. Този метод за фрагментиране налага използването на голям езиков модел, а от там произлиза и ограничение за размера на текста, респективно на броя на токените, които могат да се подават като вход към модела. Това ограничение всъщност определя размера на групите от страници. По време на процеса на изпълнение се изчислява броя токени, които се съдържат на всяка страница и по този начин се създават групи с максимална големина от 3200 токена. Броят на токените е изчислен на базата на проведени тестове с различни нормативни документи.

Условието за добавяне на предходна и следваща страница произтича от съображението за прекъсване на смисловото значение на текста. По този начин се гарантира, че разделението по страници няма да наруши смисъла на отделните фрагменти.

2. Фрагментиране на текста чрез голям езиков модел

Всяка от получените групи от страници се подава като вход към голям езиков модел, заедно с инструкции за фрагментиране на текста и генериране на кратко описание (контекст) за съдържанието на фрагмента.

Големият езиков модел, използван в тази стъпка е Gemini 2.0 Flash Experimental (gemini-2.0-flash-exp). Изборът на този модел е основан на два ключови критерия, които го правят оптимален спрямо алтернативите:

✓ **Контекстен прозорец и ефективност при обработка на дълги документи**

Gemini 2.0 Flash Experimental предлага най-големия контекстен прозорец сред съвременните големи езикови модели (LLMs) – 1,048,576 токена.

✓ **Оптимален баланс между цена и производителност**

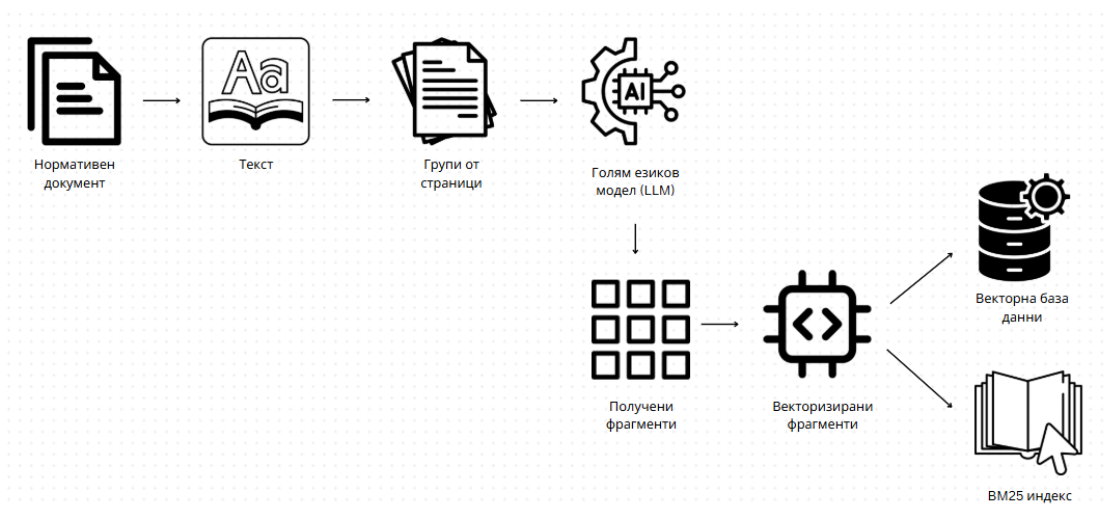
Съществува още един модел, който поддържа контекстен прозорец с такъв размер (1,048,576) и това е Gemini 2.0 Flash Thinking Experimental. Въпреки идентичния контекстен прозорец, изборът на Gemini 2.0 Flash Experimental е обоснован от разликите в предназначението и изчислителните разходи между двата модела.

3. Векторизация на получените фрагменти и съхраняване в базата данни

Както при предходните два метода, така и тук получените фрагменти биват векторизирани посредством модела Alibaba-NLP/gte-multilingual-base, след което получените вектори се съхраняват във векторната база данни (Pinecone).

МЕТОД ЗА ФРАГМЕНТИРАНЕ И ВЕКТОРИЗАЦИЯ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ И ИНДЕКСИРАНЕ С BM25

Този метод включва две стъпки - фрагментиране чрез LLM модел и индексирание с BM25. Както при предходния подход, голям езиков модел обработва документа последователно и го сегментира на смислово обособени части, като отчита контекста и вътрешната свързаност на съдържанието. Разликата при този подход, е че при съхранението във векторната база данни, не се разчита единствено на векторните представяния на фрагментите (embeddings), а се създава и т.нар. BM25 (Best Matching 25) индекс за ефективно търсене.



Фигура 16. Метод за фрагментиране и векторизация чрез голям езиков модел и BM25

BM25 (Best Matching 25) е алгоритъм за търсене и ранжиране на документи, използван широко в информационното извличане. Той е разширена версия на TF-IDF (Term Frequency-Inverse Document Frequency) индекса и се базира на статистическия модел на вероятно извличане на информация.

BM25 индексът изчислява релеванността на даден текст спрямо потребителска заявка. Докато векторните представяния чрез embeddings се фокусират само върху семантичната близост между фрагментите, BM25 осигурява прецизно извличане, като взема предвид терминологичната точност и честота. Използването му в допълнение към векторните представяния (embeddings) позволява хибриден подход за търсене, при който се съчетава традиционно индексирание с лексикално търсене и семантично извличане на информация. Това дава възможност за по-добра точност и по-широк обхват на извлечените резултати.

В резултат на извършеното фрагментиране, векторизация и индексирание, нормативните текстове са трансформирани в семантично търсима база от знания, която служи като основа за последващо генериране на съдържание. Следващата стъпка от концептуалната методология се фокусира върху механизма за генериране на бизнес правила, който използва тази база, заедно с потребителска заявка и ясно зададени инструкции, за да създаде регулаторно обосновани и приложими правила за оптимизация на бизнес процеси.

Тази стъпка се явява вторият основен модул от Retrieval-Augmented Generation (RAG) архитектурата. Този модул разчита на възможностите на голям езиков модел (LLM), който, базирайки се на предоставените фрагменти, може да:

- ✓ Идентифицира най-подходящите опорни точки за формулиране на отговор;
- ✓ Минимизира риска от т.нар. „халюцинации“ (генериране на недостоверни факти), като се позовава на реално извлечена и проверима информация;
- ✓ Гарантира актуалност, тъй като работи върху динамично обновявана база от данни, а не само върху предварително обучен корпус.

Като комбинира извлеченото знание с формата на потребителското запитване, езиковият модел изгражда отговор, който е едновременно контекстуално подходящ, терминологично прецизен и нормативно обоснован. По този начин се надграждат стандартните техники за текстово търсене чрез автоматично и кохерентно формулиране на резултатите.

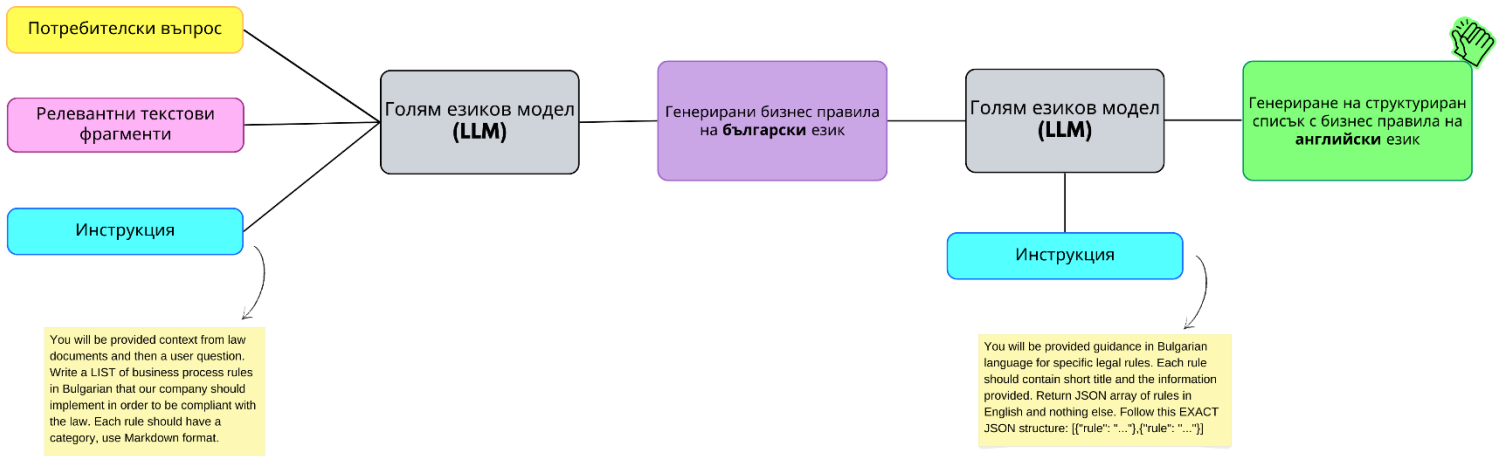
На този етап изследването предлага т.нар. метод за генериране на бизнес правила.

МЕТОД ЗА ГЕНЕРИРАНЕ НА БИЗНЕС ПРАВИЛА

След като извлечените текстови фрагменти са подготвени и индексирани в семантична база от знания, следващата логическа стъпка в архитектурата от тип Retrieval-Augmented Generation (RAG) е генерирането на релевантно съдържание въз основа на тези фрагменти и потребителското запитване. В рамките на настоящото изследване, този процес се реализира чрез метод за автоматично генериране на бизнес правила, който използва голям езиков модел (LLM) за синтезиране на текст, отговарящ на нормативния контекст и специфичните нужди на бизнес процеса.

Методът цели да формулира приложими, нормативно обосновани и контекстуално адекватни правила, които могат да бъдат използвани за последваща оптимизация на съществуващи бизнес процеси. Основен фокус е поставен върху проектирането на входните параметри, селекцията на релевантни фрагменти и управлението на инструкциите към модела, така че да се осигури максимална точност и интерпретируемост на генерирания резултат.

Настоящият метод описва двуетапен процес, при който извлечените текстови фрагменти от първия модул се комбинират с потребителско запитване и специализирани инструкции за формиране на структурирани правила на български, а впоследствие и на английски език. Методът е базиран на използването на голям езиков модел, който е използван за генериране на бизнес правила и включва следните две стъпки:



Фигура 17. Метод за генериране на бизнес правила

1. Генериране на бизнес правила на български език

Входните параметри към големия езиков модел са:

- Потребителско запитване - формулирано на български език и отразяващо конкретен контекст или проблемна област, свързана с нормативни документи.
- Извлечени текстови фрагменти - първият модул на RAG архитектурата идентифицира и подава десетте най-релевантни (тематично близки) фрагменти от законови и регулаторни документи.
- Инструкция за генериране на бизнес правила - формулирана е като насока към големия езиков модел (LLM).

Големият езиков модел анализира предоставените текстови фрагменти съвместно с потребителското запитване. На тази основа той формулира набор от бизнес правила на български език, всяко придружено от категория. По този начин се осигурява по-четивна и структурирана форма на резултатите, отговаряща на конкретните нормативни изисквания и организационни цели.

2. Преобразуване на генерираните правила от български на английски език

Отново се използват възможностите на същия голям езиков модел за генериране на съдържание. Входните параметри за тази стъпка в метода са следните:

- Бизнес правила на български език - получени от предходната стъпка, съдържащи основни насоки за съответствие с конкретните законови документи.
- Инструкция за преобразуване на бизнес правилата - формулирана е отново за същия голям езиков модел, но изисква резултатът да бъде представен в англоезична структура.

Големият езиков модел интерпретира бизнес правилата на български език и превежда съдържанието им на английски, като спазва зададения JSON формат. Всеки елемент в JSON масива съдържа ключ „rule“ и стойност, описваща конкретното правило, включително кратко заглавие и обяснителен текст. Изискването за строга JSON структура гарантира лесна интеграция с последващи системи за обработка и анализ, които оперират с форматирана информация.

В резултат от двуетапния процес се формира набор от структурирани бизнес правила на английски език, представени под формата на JSON масив. Благодарение на метода се:

- Запазва смисловата свързаност с първоначалните нормативни изисквания, извлечени в модула за извличане.
- Осигурява висока пригодност за по-нататъшна автоматизирана обработка, тъй като данните са в широко приложим JSON формат.
- Гарантира непротиворечивост и адекватност на генерираните правила, защото те отразяват както законовите текстове, така и потребителските нужди, определени в запитването.

Предложеният метод за генериране на бизнес правила и последващо превеждане на английски език добавя съществена стойност в контекста на RAG архитектурата, позволявайки:

- Динамично адаптиране към специфичните нормативни изисквания, които могат да варират по време и юрисдикция.

- Улеснено многоезично разширяване, тъй като техническият процес е пригоден към използване на различни целеви езици, при условие че LLM поддържа съответните езикови ресурси.
- Оптимално взаимодействие с външни системи (например системи за управление на бизнес процеси, правни платформи, AI анализатори), където структурираният JSON формат улеснява интеграцията и автоматизацията.

Съчетанието между генериране на законово съобразени правила и адаптиране на езиковата форма разширява приложимостта на RAG в различни индустриални вертикали и правни режими, като същевременно осигурява висока степен на стандартизация и гъвкавост в бизнес среда.

В рамките на настоящото изследване при приложението на метода за генериране на бизнес правила е избран големият езиков модел **BgGPT**, разработен от INSAIT. Изборът е продиктуван от високата езикова компетентност на модела на български език, както и от специализацията му в синтактичния и семантичен анализ на текстове на български. Сред наличните версии, BgGPT 27B се отличава с най-добри резултати при обработка на юридически и административни текстове, като значително превъзхожда други мултиезикови LLM модели в задачите по генериране на съдържание на български език.

ЕКСПЕРИМЕНТАЛЕН АНАЛИЗ

Семантичното фрагментиране е ключов етап в процеса на извличане и използване на релевантна информация. Основната му цел е да разграничи отделни логически и смислово свързани части от текстовите данни, така че те да могат да бъдат ефективно индексирани и впоследствие извлечени при зададено потребителско запитване. Този раздел от научното изследване представя експериментално тестване на четирите предложени метода за семантично фрагментиране и векторизация на текст, като всеки от тях е тестван върху четири примерни запитвания, свързани с вече векторизираните и съхранени нормативни документи.

Получените резултати включват релевантните фрагменти, които са извлечени от базата данни според съответния метод на базата на потребителското запитване, също така метаданни на файла, от който са извлечени, както и т.нар. метрика – score. Това е т.нар. коефициент на „подобност“ на потребителската заявка и намерения фрагмент.

Метриката може да варира от 0 до 1, в зависимост от това каква семантична близост има между двата текста. По-висок резултат (score) означава по-голяма релевантност на получения фрагмент към потребителското запитване.

При експерименталното тестване на всички методи се извличат първите 10 от най-подходящите фрагменти, които да послужат като вход за големия езиков модел (BgGPT), за генериране на бизнес правила.

Получените текстови резултати са представени в подробности в основния текст на дисертационния труд. В контекста на настоящия реферат, с цел синтезирано изложение, ще бъде включен само примерен резултат, илюстриращ основните аспекти на анализа.

Представеният резултат е от метода за фрагментиране чрез семантична близост на изречения, при следното зададено потребителско запитване: *"Какви действия трябва да предприеме моята компания във връзка с обработката на биометрични данни?"*.

Получените релевантни фрагменти са следните:

Закон	Съдържание	Резултат
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	Предаването на лични данни на получатели в други държави - членки на Европейския съюз, или на агенции, служби и органи на Европейския съюз, създадени съгласно дал V, глави 4 и 5 от Договора за функционирането на Европейския съюз, се извършва при същите условия, които се прилагат при подобно предаване в Република България. Чл. 51. (Нов - ДВ, бр. 17 от 2019 г.) (1) Обработването на лични данни, разкриващо расов или етнически произход, политически възгледи, религиозни или философски убеждения, членство в професионални съюзи, обработването на генетични данни, биометрични данни с цел уникално идентифициране на физическото лице, данни, свързани със здравословното състояние или сексуалния живот и сексуалната ориентация на лицето, е разрешено, когато това е абсолютно необходимо, съществуват подходящи гаранции за правата и свободите на субекта на данни и е предвидено в правото на Европейския съюз или в законодателството на Република България.	0.650064588
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	Чрез мерки по ал. 1 администраторът осигурява защита на личните данни и на етапа на проектирането, като отчита достиженията на техническия прогрес, разходите за прилагане и естеството, обхвата, контекста и целите на обработването на лични данни, както и рисковете за правата и свободите на физическите лица при обработването. Мерките трябва да са съобразени с изискванията на чл. 45, планират се към момента на определяне на средствата за обработването на лични данни и се прилагат при самото обработване. Мерките може да включват псевдонимизация, съвещане на данните до минимум и въвеждане на необходими гаранции в процеса на обработване на лични данни.	0.643237591
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	При обработването на лични данни за целите по чл. 42, ал. 1 личните данни трябва да: 1. се обработват законно, обосноваемо и добросъвестно; 2. се събират за конкретни, изрично указани и законни цели и да не се обработват по начин, който е несъвместим с тези цели; 3. са подходящи, относими и да не надхвърлят необходимото във връзка с целите, за които данните се обработват; 4. са точни и при необходимост да са поддържани в актуален вид; трябва да се предприемат всички необходими мерки, за да се гарантира своевременното изтриване или коригиране на неточни лични данни, като се имат предвид целите, за които те се обработват; 5. се съхраняват във вид, който позволява идентифицирането на субекта на данните за период, не по-дълъг от необходимия за целите, за които те се обработват; 6. се обработват по начин, който гарантира подходящо ниво на сигурност на личните данни и, включително защита срещу неразрешено или незаконно обработване и срещу случайна загуба, унищожаване или повреждане, като се прилагат подходящи технически или организационни мерки. (2) Обработването на лични данни от администратор, който първоначално ги е събрал, или от друг администратор за която и да е от целите по чл. 42, ал. 1, различна от целта, за която личните данни са събрани, се разрешава, при условие че: 1. администраторът е оправомощен да обработва лични данни за такава цел в съответствие с правото на Европейския съюз или законодателството на Република България, и 2. обработването е необходимо и пропорционално на тази различна цел в съответствие с правото на Европейския съюз или със законодателството на Република България. (3)	0.631484032
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	При извършване на справка или разкриване на данни дневниците по ал. 1 трябва да дават възможност за установяване на основание, датата и часа на тези операции и доколкото е възможно - идентификацията на лицето, което е направило справка и или е разкрило личните данни, както и данни, идентифициращи получателите на тези лични данни.	0.625713646
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	При обработването на лични данни за целите на създаване на фотографско или аудио-визуално произведение чрез заснемане на лице в хода на обществената му дейност или на обществено място не се прилагат чл. 6, чл. 12 - 21, чл. 30 и 34 от Регламент (ЕС) 2016/679. Чл. 25к. (Нов - ДВ, бр. 17 от 2019 г.) (1) Работодател или орган по назначаването, в качеството си на администратор на лични данни, приема правила и процедури при: 1. използване на система за докладване на нарушения; 2. ограничения при използване на вътрешнофирмени ресурси; 3. въвеждане на системи за контрол на достъпа, работното време и трудовата дисциплина. (2)	0.620340049
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	Компетентният орган предприема необходимите мерки лични данни, които са неточни, непълни или вече не са актуални, да не се предават. За тази цел всеки компетентен орган, доколкото това е възможно, проверява качеството на личните данни преди тяхното предаване. Доколкото е възможно, при всяко предаване на лични данни се добавя необходимата информация, позволяваща на получаващия компетентен орган да оцени степента на точност, пълнотата и надеждността на личните данни и до каква степен те са актуални.	0.613881588
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	(3) Обработването от страна на обработващия лични данни се урежда с друг правен акт съгласно правото на Европейския съюз или законодателството на Република България, който обвързва обработващия лични данни с администратора по ал. 1 и регламентира предмета и срока на обработването, естеството и целта на обработката, вида лични данни и категориите субекти на данни, задълженията и правата на администратора. Посоченият договор или друг правен акт предвижда по-специално, че обработващия лични данни: 1. действат единствено по указания на администратора; 2. гарантира, че лицата, оправомощени да обработват личните данни, са пълни задължени за поверителност или са задължени по закон да спазват поверителност; 3. подпомага администратора с всички подходящи средства, за да се гарантира спазването на правата на субекта на данни; 4. по избор на администратора изтрива или връща на администратора всички лични данни след приключване на предоставяне на услуги по обработване на данни и изтрива съществуващите копия, освен ако правото на Европейския съюз или законодателството на Република България не изисква съхранение на личните данни; 5. предоставя на администратора цялата информация, необходима за доказване на спазването на този член; 6. спазва условията по т. 1 - 5 и по ал. 2 за включване на друг обработващ лични данни. (4) Договорът или другият правен акт, посочен в ал. 3, се изготвя в писмена, включително в електронна форма. (5)	0.6090554
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	(6) Когато администраторът има основателни съмнения относно самоличността на физическото лице, което подава искане по чл. 55 или 56, той може да поиска да се предостави допълнителна информация, необходима за потвърждаване на самоличността на субекта на данните. Срокът по ал. 3 започва да тече от получаването на тази допълнителна информация. Чл. 54. (Нов - ДВ, бр. 17 от 2019 г.) (1)	0.602767169
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	Правилата и процедурите по ал. 1 съдържат информация относно обхвата, задълженията и методите за прилагането им на практика. С тях се отчитат предметът на дейност на работодателя или органа по назначаването и свързаното с него естество на работата и не може да се ограничават правата на субектите на данните по Регламент (ЕС) 2016/679 и по този закон. (3) Работниците и служителите се уведомяват за правилата и процедурите по ал. 1. Чл. 25к. (Нов - ДВ, бр. 17 от 2019 г.) (1) Работодател или орган по назначаването, в качеството си на администратор на лични данни, определя срок за съхранение на лични данни на участници в процедури по набиране и подбор на персонала, който не може да е по-дълъг от 6 месеца, освен ако кандидатът е дал своето съгласие за съхранение за по-дълъг срок. След изтичането на този срок работодателят или органът по назначаването изтрива или унищожава съхраняваните документи с лични данни, освен ако специален закон предвижда друго.	0.600753725
ЗАКОН ЗА ЗАЩИТА НА ЛИЧНИТЕ ДАННИ	Обработването на лични данни за целите на Националния архивен фонд на Република България е обработване в обществен интерес.	0.599012434

Генерираните бизнес правила на български език са както следва:

1. ****Определяне на целите и обхвата на обработката:**** Компанията трябва ясно да определи какви са целите за събирането, използването и съхранението на биометричните данни. Това включва идентифициране на конкретните нужди и причини за обработката на тези данни.

2. ****Законосъобразност и добросъвестност:**** Обработката на биометрични данни трябва да бъде в съответствие с приложимото законодателство и да се извършва по начин, който е честен и прозрачен спрямо субектите на данните.
3. ****Минимизиране на данните:**** Компанията трябва да събира само необходимите биометрични данни за постигане на своите цели. Излишното или ненужно събиране на такива чувствителни данни е недопустимо.
4. ****Точност и актуализация:**** Биометричните данни трябва да бъдат точни и поддържани в актуален вид. Компанията трябва да предприеме всички необходими мерки за своевременно изтриване или коригиране на неточни биометрични данни.
5. ****Съхранение с ограничен срок:**** Биометричните данни трябва да се съхраняват само за периода, който е необходим за постигане на целите, за които са събрани. След изтичането на този период те трябва да бъдат унищожени или изтрити по подходящ начин.
6. ****Сигурност:**** Компанията трябва да внедри подходящи технически и организационни мерки за защита на биометричните данни от неразрешен достъп, промяна, разкриване или унищожаване. Това включва криптиране на данните, контрол на достъпа и редовно тестване на сигурността.
7. ****Информираност и съгласие:**** Субектите на данни трябва да бъдат информирани за обработката на техните биометрични данни по ясен и разбираем начин. Трябва да бъде получено тяхното изрично съгласие, освен ако не е предвидено друго в закона.
8. ****Права на субектите на данни:**** Компанията трябва да осигури спазването на правата на субектите на данни, включително правото на достъп, коригиране и изтриване на лични данни, както и правото на възражение срещу обработката.
9. ****Документация:**** Компанията трябва да поддържа документация за всички процеси, свързани с обработката на биометрични данни, включително целите, мерките за сигурност и процедурите за управление на рисковете.
10. ****Обучение и осведоменост:**** Персоналът на компанията трябва да бъде обучен относно значението на защитата на биометричните данни и да е запознат с вътрешните правила и процедури, свързани с тяхната обработка.

След провеждането на експерименталните тестове, изследването разработва метод за оценка на получените резултати, който позволява обективен анализ на различните подходи за фрагментиране. Чрез този метод се извършва сравнителна оценка на ефективността на разглежданите техники, което впоследствие позволява идентифицирането и избора на най-оптималния метод за интегриране в цялостната методологична рамка на изследването.

МЕТОД ЗА ОЦЕНКА НА ФРАГМЕНТИРАНЕ И ВЕКТОРИЗАЦИЯ

За дефинирането на метод за оценка, изследването предлага подхода LLM-as-Judge като автоматизирана методология за оценка на генерираното съдържание, с оглед на неговата висока ефективност, скалируемост и съответствие с експертните оценки. Основната концепция се основава на способността на големите езикови модели (LLMs) не само да генерират текст, но и да извършват обективна оценка на съдържанието спрямо предварително дефинирани критерии, като релевантност, точност, пълнота и последователност. Приложението на LLM-as-Judge в рамките на изследването е мотивирано от няколко ключови фактора:

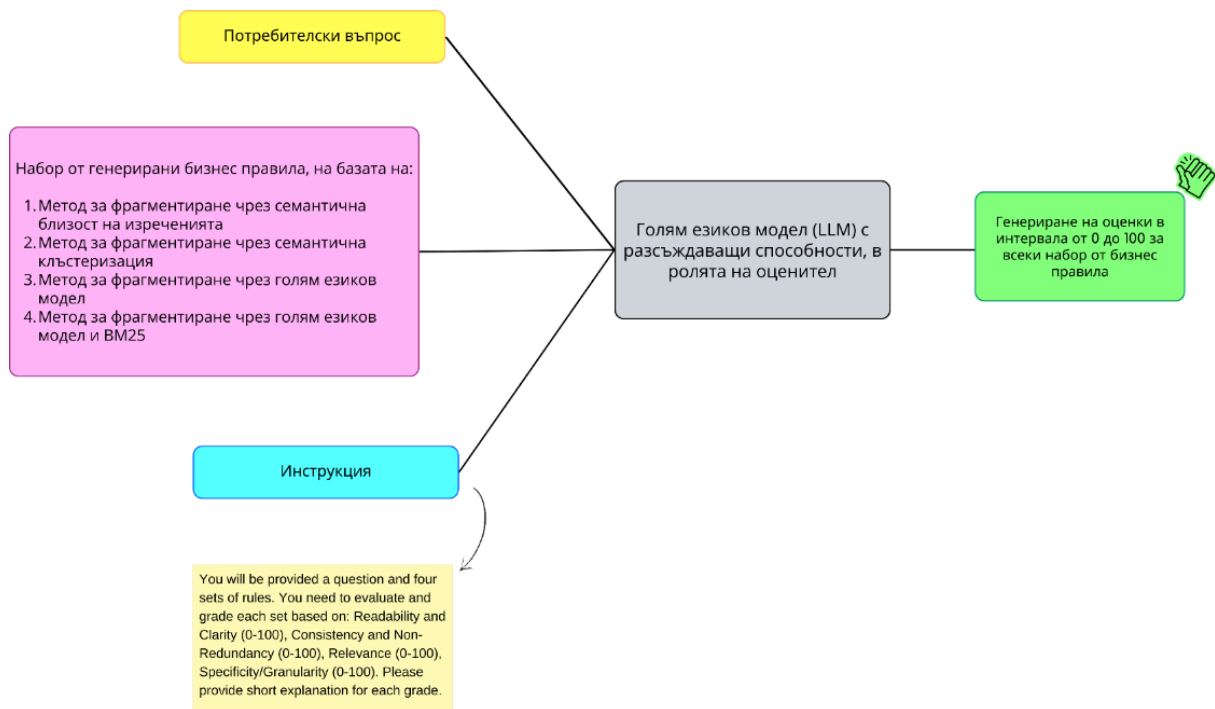
1. **Обективност и консистентност** – анализите показват, че LLM-базираните оценители демонстрират стабилност на резултатите и ниска стандартна девиация, което ги прави надеждни за систематични оценки.
2. **Скалируемост и автоматизация** – моделите могат да обработват и оценяват големи обеми текстови данни в реално време, елиминирайки нуждата от ръчна експертна оценка, което води до значително намаляване на разходите.
3. **Съпоставимост с човешките оценки** – експериментални проучвания показват висока степен на съответствие (над 80%) между оценките, предоставени от LLM модели, и тези на експертите, като в някои случаи автоматизираните оценки дори надминават консистентността между самите експерти.
4. **Ограничаване на разходите и ресурсоемкостта** – ръчната експертна оценка изисква значителни ресурси и време, което я прави трудно приложима в мащабни анализи. В този контекст LLM оценителите предоставят икономически ефективна алтернатива.

Относно избора на големи езикови модели в ролята на оценители, изследването предлага използването на три независими AI оценители (AI judges), които извършват качествена и количествена оценка на резултатите от методите, не само за генериране на бизнес правила, но и за генериране на предложения за оптимизация, които е дефиниран по-късно в изследването. Цялостния процес на оценка е разделен на две фази:

1. **Първа фаза** – оценка на генерираните бизнес правила, базирани на нормативни документи, с цел идентифициране на най-подходящия метод за фрагментиране.
2. **Втора фаза** – анализ на предложенията за оптимизация, генерирани от различни LLM модели, въз основа на предварително зададени критерии, което позволява сравнителна оценка на моделите и извеждане на заключения за избор.

За да се гарантира обективност и надеждност, изборът на LLM модели, изпълняващи ролята на оценители, е извършен чрез механизъм, който също така е използван за подбора на модели във втората част от изследването, а именно модели за генериране на предложения за оптимизация. В случая при избор на оценители, обаче, критерият за

подбор е избраните модели да принадлежат към категорията „Thinking“, която включва модели, доказали своите способности за дълбоко логическо разсъждение, аргументиран анализ, критично мислене и оценка на сложни текстове. На тази основа, трите езикови модела, избрани за AI оценители са: Gemini-2.0-Flash-Thinking-Exp-01-21, GPT-o1-2024-12-17 и DeepSeek-R1.



Фигура 18. Метод за оценка на фрагментиране и векторизация

Методът за оценка получава следните входни параметри:

1. **Потребителски въпрос** - зададения от потребителя въпрос или проблем, към който са формулирани бизнес правила. Това осигурява контекста, в който се извършва анализът и оценката на качеството на фрагментираните текстове.
2. **Набор от генерираните бизнес правила** - всички генерирани бизнес правила, които са генерирани от големия езиков модел, на базата различните методи за фрагментиране. Целта е да се анализира доколко генерираните правила отговарят на изискванията за яснота, уместност и конкретност.
3. **Инструкция за оценка** - специализирана инструкция към оценяващия модел, която задава критериите за оценка и формулира изискванията за обобщаващ

коментар. Инструкцията дефинира четири основни критерия и изисква кратка обосновка за всяка оценка, като по този начин се гарантира прозрачност и проследимост на резултатите.

Предложеният метод за оценка се прилага трикратно, като при всяко изпълнение се изменя моделът, изпълняващ ролята на оценител. Получените резултати се събират, анализират и представят в табличен вид в текста на дисертацията. Оценяването се извършва на ниво потребителски въпрос, като са разгледани четири различни запитвания. В следващата част са представени примерни резултати за първото потребителско запитване ("Какви действия трябва да предприеме моята компания във връзка с обработката на биометрични данни?"):

Модел OPEN AI					
Подход	Четивност и яснота	Последователност и липса на повторения	Релевантност	Конкретност / Детайлност	Средна оценка
ФРАГМЕНТИРАНЕ ЧРЕЗ СЕМАНТИЧНА БЛИЗОСТ НА ИЗРЕЧЕНИЯ	95	90	100	80	91,25
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ СЕМАНТИЧНА КЛЪСТЕРИЗАЦИЯ	95	85	100	90	92,5
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ	90	80	100	95	91,25
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ И ИНДЕКС BM25	85	80	95	85	86,25
Модел DEEP SEEK					
ФРАГМЕНТИРАНЕ ЧРЕЗ СЕМАНТИЧНА БЛИЗОСТ НА ИЗРЕЧЕНИЯ	90	90	95	90	91,25
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ СЕМАНТИЧНА КЛЪСТЕРИЗАЦИЯ	90	85	95	95	91,25
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ	85	80	95	95	88,75
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ И ИНДЕКС BM25	80	75	85	70	77,5
Модел GEMINI					
ФРАГМЕНТИРАНЕ ЧРЕЗ СЕМАНТИЧНА БЛИЗОСТ НА ИЗРЕЧЕНИЯ	90	95	95	70	87,5
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ СЕМАНТИЧНА КЛЪСТЕРИЗАЦИЯ	85	90	98	80	88,25
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ	80	85	96	80	85,25
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ И ИНДЕКС BM25	75	90	95	65	81,25

Таблица 1. Резултати от подходите за фрагментиране – първо запитване

Получените резултати се агрегират и за тях се изчислява стандартното отклонение на оценките, предоставени от различните AI оценители. Допълнително е проведен анализ на стандартното отклонение на оценките за различните методи за фрагментиране на текст, обхващащ всички разглеждани потребителски въпроси.

Подход	Първи въпрос	Втори въпрос	Трети въпрос	Четвърти въпрос	Средна оценка
ФРАГМЕНТИРАНЕ ЧРЕЗ СЕМАНТИЧНА БЛИЗОСТ НА ИЗРЕЧЕНИЯ	90,00	62,92	88,75	79,17	80,21
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ СЕМАНТИЧНА КЛЪСТЕРИЗАЦИЯ	90,67	93,75	74,58	72,08	82,77
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ	88,42	90,00	81,67	90,42	87,63
ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ И ИНДЕКС BM25	81,67	83,75	80,83	89,58	83,96

Таблица 2. Обща оценка на резултатите от фрагментиране

МЕТРИКА	ФРАГМЕНТИРАНЕ ЧРЕЗ СЕМАНТИЧНА БЛИЗОСТ НА ИЗРЕЧЕНИЯ	ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ СЕМАНТИЧНА КЛЪСТЕРИЗАЦИЯ	ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ	ФРАГМЕНТИРАНЕ НА ТЕКСТ ЧРЕЗ ГОЛЯМ ЕЗИКОВ МОДЕЛ И ИНДЕКС ВМ25
Стандартно Отклонение	12,50	11,02	4,06	3,95

Таблица 3. Стандартно отклонение при различните подходи за фрагментиране

На базата на проведените експерименти и анализи на резултатите, изследването идентифицира най-подходящия метод за фрагментиране на нормативни документи в контекста на управление на бизнес процеси (BPM). Резултатите показват, че методът, базиран на голям езиков модел, демонстрира най-висока ефективност, както по отношение на консистентността на оценките (ниско стандартно отклонение), така и спрямо качеството на получените резултати.

Допълнителният анализ разкрива, че използването на ВМ25 не води до съществено подобрение при обработката на нормативни текстове. Това предполага, че спецификата на текстовия корпус и необходимостта от дълбоко семантично разбиране са определящи фактори при избора на метод за фрагментиране. Въз основа на тези наблюдения, големите езикови модели се открояват като най-ефективното и надеждно решение за осигуряване на високо качество и релевантност при фрагментиране на текстове на нормативни документи.

5. ГЕНЕРИРАНЕ И ОЦЕНКА НА ПРЕДЛОЖЕНИЯ ЗА ОПТИМИЗАЦИЯ

Пета глава от дисертационния труд се фокусира върху разработването на метод за генериране и оценка на предложения за оптимизация на бизнес процеси, експерименталното тестване на метода в различни категории бизнес процеси, както и оценка на получените резултати.

МЕТОД ЗА ГЕНЕРИРАНЕ НА ПРЕДЛОЖЕНИЯ ЗА ОПТИМИЗАЦИЯ

Предложеният метод (фиг.19) в изследването за автоматично генериране на предложения за оптимизация на бизнес процеси, комбинира бизнес правилата, генерирани от предходната стъпка, BPMN моделите, които са съхранени в реляционна база данни и представени в JSON формат и инструкции за дефиниране на задачата.

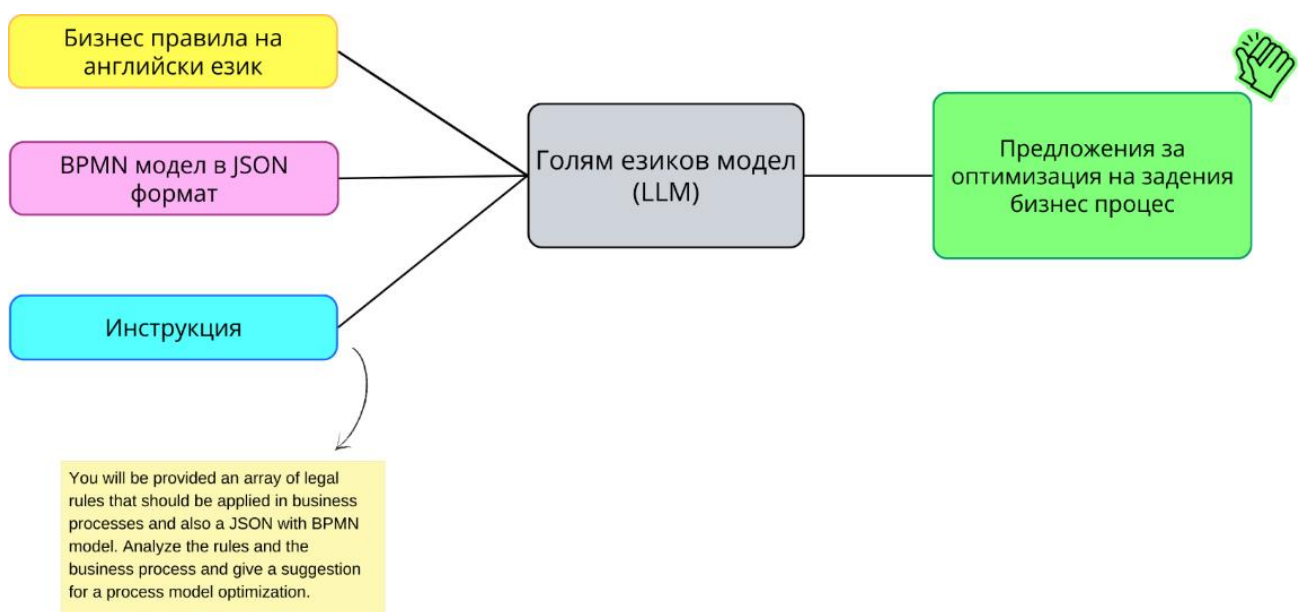
Методът се реализира чрез интегриране на голям езиков модел (LLM), който анализира входните данни и формулира предложения за усъвършенстване на съответния бизнес процес. Целта на разработения метод е да предостави структуриран и мащабируем начин за прилагане на нормативни изисквания към реални бизнес процеси, като в същото време предложи конкретни стъпки за оптимизация на разглежданите процесни модели. За постигането на тази цел методът включва няколко компонента:

1. **Бизнес правила**, базирани на нормативни документи - както бе описано в предходните стъпки, бизнес правилата се събират и формулират така, че да отразяват съответните правни и регулаторни изисквания, приложими към бизнес процесите. Тези правила се обобщават във вид на списък, който служи като основен вход на метода.

2. **BPMN модели в JSON формат** - реляционната база данни съдържа множество BPMN модели, описващи различни бизнес процеси. За нуждите на настоящия метод се извличат селектираните процесни модели, които са представени в JSON формат. Тази структура улеснява предаването на информацията към големия езиков модел и осигурява стандартизиран начин за машинна интерпретация на процесите.

3. **Инструкция към големия езиков модел** - критичен елемент в метода е формулирането на указания към големия езиков модел. Тази инструкция задава ясно контекста на задачата, като изисква от езиковия модел да отчете нормативната рамка и конкретния BPMN модел, за да идентифицира потенциални несъответствия или възможности за усъвършенстване.

4. **Голям езиков модел (LLM)** - за анализ и синтез на информацията се прилага LLM, който притежава разширени семантични и синтактични способности за обработка на



Фигура 19. Метод за генериране на предложения за оптимизация

С цел да бъде идентифициран най-подходящият голям езиков модел (LLM) за автоматично генериране на предложения за оптимизация на бизнес процеси, в рамките на настоящото изследване прилага дефинирания метод с шест различни „кандидат-модела“. За осигуряване на обективност и надеждност, подборът на моделите е извършен на базата на класации публикувани в Chatbot Arena Leaderboard, като са разгледани две основни категории – платени модели и модели с отворен код. С цел систематичен подбор, в изследването са включени трите най-високо класирани модела във всяка категория, като критериите за селекция са базирани на филтър по категория „Изпълнение на инструкции“ (Instruction Following). Моделите, използвани в анализа, са:

- **Платени модели:**

1. ChatGPT-4o-latest (2024-11-20)
2. Gemini-2.0-Flash-Exp-01-21
3. Claude 3.5 Sonnet (20241022)

- **Модели с отворен код:**

1. DeepSeek-V3
2. Meta-Llama-3.1-405B-Instruct-bf16
3. Mistral-Large-2407

При селекцията са изключени моделите от категорията „Thinking“, тъй като те са използвани в ролята на оценители в предходни етапи на изследването.

ПРИЛОЖЕНИЕ НА МЕТОДА ЗА ГЕНЕРИРАНЕ НА ПРЕДЛОЖЕНИЯ ЗА ОПТИМИЗАЦИЯ

В дисертационния труд са представени резултатите от всички шест приложения на метода за генериране на предложения за оптимизация, като данните са детайлно структурирани с цел осигуряване на яснота и възпроизводимост на анализа. Входните параметри и резултатите са организирани както следва:

- Входни параметри: бизнес модел, потребителски въпрос, генерирани бизнес правила на английски език

- Резултат: генерирани предложения за оптимизация от съответния голям езиков модел

В контекста на настоящия реферат с цел синтезиране на изложението, ще бъде включен само един потребителски въпрос, приложен върху само един бизнес модел, разгледан в рамките на едно от приложенията на метода с един от изследваните езикови модели. Шестте езикови модела са общо приложени върху четири бизнес процеса, но този ограничен обхват има за цел да илюстрира основните характеристики и ефективността на предложената методология, без да се навлиза в детайлния анализ, представен в дисертацията.

Входни параметри:

- Бизнес модел: категория „Логистика“
- Потребителски въпрос: Зададеният потребителски въпрос, насочен към модела, е *„Какви действия трябва да предприеме моята компания във връзка с обработката на биометрични данни?“*.
- Генерираните бизнес правила представени на английски език, включват:

№	Правило
1	Conduct a Data Protection Impact Assessment (DPIA) before processing biometric data to identify risks and implement appropriate measures.
2	Consult with the Commission or Inspectorate if the processing will be part of a new biometric data registry that is about to be created.
3	Obtain explicit, informed consent from the data subject before processing their biometric data.
4	Apply appropriate technical and organizational security measures to protect biometric data, taking into account the risks for individuals' rights and freedoms.
5	Store data only for the necessary period, allowing identification of the data subject for no longer than required for the purposes they are processed.
6	Periodically review and update security measures as needed to ensure lawful processing of biometric data.
7	Maintain registers of activities related to biometric data processing and provide access to these upon request from the Commission or Inspectorate.
8	Train personnel who have access to biometric data on how to process them in accordance with legal requirements and applied security measures.

Резултат от модел Gemini-2.0-Flash-Exp-01-21:

****Analysis of the BPMN Model****

The provided BPMN model describes a hiring process, starting from receiving an application, and ending with either a signed employment contract or a rejection. Here's a breakdown:

1. **Application Receipt:** The process starts with receiving an application.
2. **Initial Verification:** The application is verified against predefined criteria.
3. **Decision Point:** Based on the verification, the applicant is either rejected or moved to the next stage.
4. **Job Interview:** If accepted, the applicant is invited to an interview.
5. **Interview Evaluation:** After the interview, a decision is made on the applicant's suitability.
6. **Contract Preparation:** If suitable, an employment contract is prepared.
7. **Contract Sending & Signing:** The contract is sent to the applicant, and the process waits for the signed contract.
8. **Rejection:** If the applicant is rejected at any point, a rejection is sent.

Missing Elements Related to Biometric Data Processing

The current model **does not explicitly include any activities related to biometric data processing**. Therefore, it does not comply with any of the provided legal rules.

Recommendations for Process Model Optimization

To incorporate the legal requirements for biometric data processing, the following changes are recommended:

1. Identify Biometric Data Collection Points:

- * Determine at which point in the hiring process biometric data might be collected. This could be during the application process itself (e.g., facial recognition for online applications), during the interview (e.g., voice analysis), or during onboarding (e.g., fingerprint scanning for access).

- * For this example, let's assume that there is a potential for facial recognition during the online application process.

2. Incorporate DPIA:

- * **Add a "Conduct DPIA" task:** Before any biometric data processing, a Data Protection Impact Assessment (DPIA) must be conducted. This task should be placed before the biometric data collection point.

- * The DPIA task should document the potential risks and the measures to mitigate them.

3. Consultation with Authorities:

- * **Add a "Consult with Commission/Inspectorate" task:** If the biometric data processing is part of a new registry, a consultation with the relevant authorities is needed. This task should be placed after the DPIA and before the biometric data collection.

- * This task might be conditional, based on whether a new registry is being created.

4. Obtain Explicit Consent:

- * **Add a "Obtain Consent" task:** Before processing any biometric data, explicit, informed consent must be obtained from the data subject. This task should be placed immediately before any biometric data is collected.

- * This task should include a clear explanation of the processing purpose, data retention period, security measures, and the data subject's rights.

5. Implement Security Measures:

- * **Add a "Implement Security Measures" task:** Before biometric data collection, a task to implement technical and organizational security measures should be added.

* This task should be linked to the DPIA findings and should include measures to protect the data against unauthorized access, loss, or damage.

6. ****Data Storage and Retention:****

* ****Add a "Set Data Retention Policy" task:**** After the biometric data is collected, a task to set the data retention policy should be added.

* This task should define the period for which the data will be stored, ensuring it is only kept for the necessary period.

7. ****Periodic Review and Update:****

* ****Add a "Review Security Measures" timer event:**** A timer event should be added to trigger a periodic review of the security measures.

* This should be placed at the end of the process or as a separate sub-process.

8. ****Maintain Registers:****

* ****Add a "Maintain Register" task:**** A task to maintain registers of biometric data processing activities should be added.

* This task should be placed at the end of the process or as a separate sub-process.

9. ****Personnel Training:****

* ****Add a "Train Personnel" task:**** A task to train personnel on the lawful processing of biometric data should be included.

* This task should be placed at the beginning of the process or as a separate sub-process.

****Revised BPMN Model (Conceptual)****

Here's how the model could be revised conceptually:

1. ****Start Event:**** "Receipt of application"
2. ****Task:**** "Train Personnel on Biometric Data Processing" (separate sub-process, if needed)
3. ****Task:**** "Conduct DPIA for Biometric Data Collection"
4. ****Gateway:**** "New Biometric Registry?" (Yes/No)
 - * ****Yes:**** "Consult with Commission/Inspectorate"
5. ****Task:**** "Obtain Explicit Consent for Biometric Data Processing"
6. ****Task:**** "Implement Security Measures for Biometric Data"
7. ****Task:**** "Collect Biometric Data" (e.g., Facial Recognition)
8. ****Task:**** "Set Data Retention Policy"
9. ****Task:**** "Maintain Register of Biometric Data Processing"
10. ****Existing Tasks:**** "Applicant Verified," "Invite to Job Interview," etc. (as per the original model)
11. ****Timer Event:**** "Review Security Measures for Biometric Data" (periodic)
12. ****End Events:**** "Signed employment contract is present" or "Applicant rejected"

****Implementation Notes****

* ****Sub-processes:**** Some of the added tasks (e.g., "Train Personnel," "Maintain Register") can be implemented as separate sub-processes to keep the main process clean.

* **Data Objects:** Data objects should be used to represent the DPIA report, consent forms, security measures documentation, and data registers.

* **Gateways:** Gateways should be used to handle conditional logic, such as whether a new biometric registry is being created.

* **Documentation:** Each task should have clear documentation describing its purpose and the legal requirements it fulfills.

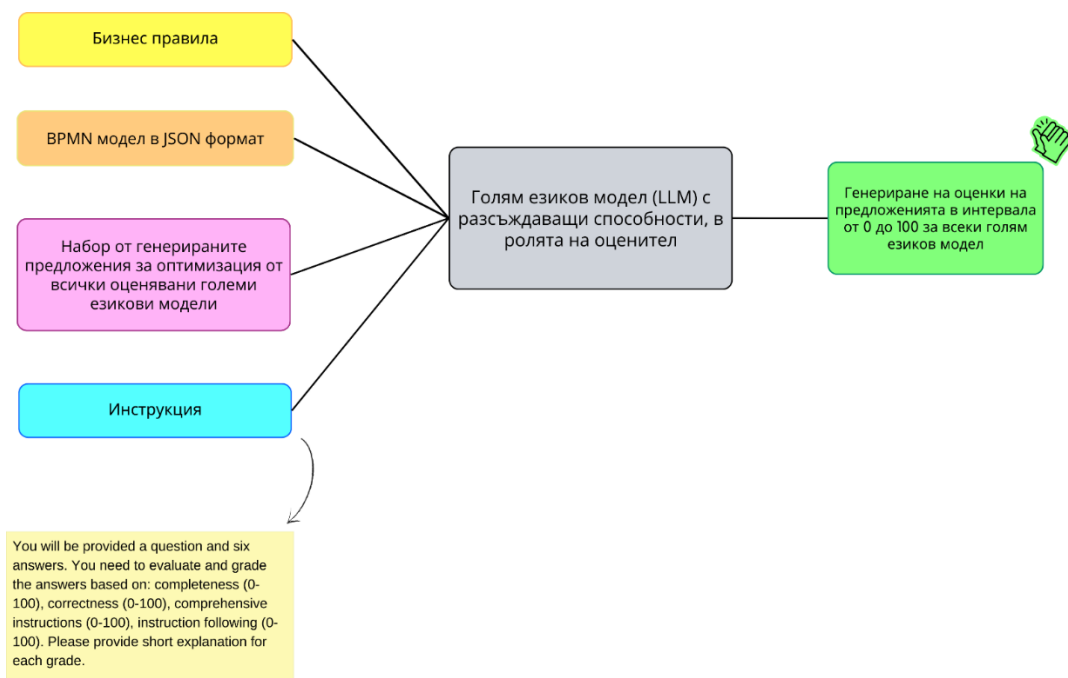
Conclusion: By incorporating these changes, the BPMN model will be compliant with the provided legal rules for biometric data processing. The revised model will ensure that biometric data is processed lawfully, with appropriate safeguards for the rights and freedoms of individuals.

Important: This is a conceptual optimization. The specific implementation will depend on the details of your organization's processes and the specific biometric data being processed. It is highly recommended to consult with legal experts and data protection officers to ensure full compliance.

Като една от последните стъпки в концептуалната рамка, изследването предлага метод за оценка на генерираните предложения за оптимизация на бизнес процеси. Методът прилага отново концепцията LLM-as-Judge, при която голям езиков модел изпълнява ролята на оценител на генерираното съдържание.

МЕТОД ЗА ОЦЕНКА НА ГЕНЕРИРАНИТЕ ПРЕДЛОЖЕНИЯ ЗА ОПТИМИЗАЦИЯ

Методът разчита на предварително дефинираните бизнес правила, извлечени от нормативни документи, и на набора от генерираните предложения, представени от съответните големи езикови модели. Оценяването се извършва посредством голям езиков модел, конфигуриран да изпълнява ролята на „оценител“ (judge).



Фигура 20. Метод за оценка на генерираните предложения за оптимизация

Дефинираният метод за оценка има следните входни компоненти:

1. **Бизнес правила** – набора от бизнес правила, които е подаден към големия езиков модел първоначално за генериране на предложенията за оптимизация на бизнес процес.
2. **Бизнес модел** - бизнес моделът, за който са генерирани предложенията за оптимизация
3. **Генерираните предложения за оптимизация** - конкретните препоръки за подобряване на бизнес процесите, получени от големите езикови модел.
4. **Инструкция за оценка** - специализирана инструкция, която насочва AI оценителя да оцени предложенията въз основа на четири критерия и да присъди количествена оценка във всеки от тях от 0 до 100, придружена от кратък обосноваващ коментар.

Оценката на моделите се извършва въз основа на четири ключови критерия, които гарантират обективност и измеримост на резултатите:

- **Пълнота** – оценява степента, до която генерираните предложения обхващат всички релевантни аспекти на въпроса или задачата.
- **Коректност** – измерва точността и съответствието на генерирания текст с изискванията, произтичащи от зададените бизнес правила.
- **Изчерпателност на инструкциите** – анализира яснотата и детайлността на предоставените указания, необходими за изпълнението на предложените оптимизации.
- **Следване на инструкциите** – оценява доколко моделът се придържа към зададените инструкции и формулира отговори, съответстващи на предварително зададените параметри.

Предложеният метод демонстрира иновационен подход за оценка на системи, включващи големи езикови модели, при който втори езиков модел изпълнява ролята на обективен оценител. Чрез прилагане на четирите ключови критерия се осигурява многоаспектен анализ на генерираните предложения, който адресира не само техническото им качество, но и съответствието им с нормативната рамка и

практическата приложимост. По този начин оценяването е не само количествено (чрез числови стойности), но и качествено (посредством обосноваващи коментари и анализ на коректността спрямо нормативните документи).

Изборът на критериите за оценка е обоснован на утвърдени теоретични рамки в областта на управление на бизнес процеси (BPM), бизнес процесното реинженерство и автоматизираното оценяване чрез AI-базирани оценители. Въведените показатели отразяват ключови аспекти на качеството на генерираните предложения, които гарантират тяхната приложимост, коректност и съответствие с предварително зададените цели и изисквания.

- Пълнота - критерият оценява степента, до която генерираното предложение обхваща всички съществени аспекти на даден бизнес процес. В литературата по BPM [3] пълнотата се разглежда като критичен фактор за точността на процесните модели, като липсата на ключови компоненти може да компрометира оптимизационния ефект.
- Коректност - измерва точността и съответствието на предложените оптимизации спрямо наложените бизнес правила и логическата консистентност на процеса. В контекста на бизнес процесното реинженерство [7] коректността е ключов принцип, тъй като неправилно дефинираните оптимизации могат да доведат до нежелани оперативни последици.
- Изчерпателност - свързана с детайлността и дълбочината на генерираните предложения. Според изследванията в областта на AI-асистираното управление на бизнес процеси [77], предложенията, които са повърхностни или непълни, не могат да бъдат ефективно приложени в реални бизнес сценарии.
- Следване на инструкции - следването на инструкции е от съществено значение за ефективността на системите за автоматизирано генериране на съдържание и бизнес оптимизации. Неспазването на предварително зададените указания може да доведе до нерелевантни или неточни резултати. В тази връзка, изследването от 2023 година показва, че използването на GPT-4 за генериране на данни за следване на инструкции подобрява способността на големите езикови модели да изпълняват нови задачи без допълнителни инструкции от хора. Освен това, друго научно изследване [79] представя рамката Self-Instruct, която подобрява

способността на езиковите модели да следват инструкции чрез самогенерирани данни, демонстрирайки значителни подобрения в изпълнението на различни задачи. Тези изследвания подчертават важноста на следването на инструкции за постигане на точни и надеждни резултати в системите с изкуствен интелект.

Чрез използването на тези критерии, изследването прилага структуриран, многопластов подход за обективна оценка на генерираните предложения за оптимизация. Избраните показатели се основават на доказани принципи в BPM и бизнес процесното реинженерство и същевременно интегрират съвременните практики за оценяване.

АНАЛИЗ НА РЕЗУЛТАТИТЕ

На базата на получените отговори от шестте големи езикови модела и техните предложения за оптимизация на различните бизнес процеси е извършена оценка за всеки един от бизнес процесите в различните категории.

За целите на настоящия реферат ще бъдат представени само обобщените оценки, докато в дисертационния труд са включени детайлни оценки на всеки голям езиков модел, приложен върху четирите изследвани бизнес модела.

С цел яснота и структурираност на резултатите, е изготвена обобщена таблица, която визуализира цялостната оценка на всеки от шестте анализирани LLM модела. Таблицата съдържа средните стойности на получените оценки от всички AI оценители, предоставяйки комплексен поглед върху представянето на различните модели.

Голям Езиков Модел	Обща средна оценка на бизнес модел от категория Логистика	Обща средна оценка на бизнес модел от категория Инф. технологии	Обща средна оценка на бизнес модел от категория Финанси	Обща средна оценка на бизнес модел от категория Продажби	Средна оценка	Класиране
Gemini	93,75	95,42	93,33	95,83	94,58	1
Open AI	82,08	83,75	82,08	87,08	83,75	5
Claude	88,33	90,42	80,00	87,50	86,56	3
DeepSeek	92,50	94,17	92,17	95,83	93,67	2
Llama	71,67	58,75	70,42	63,75	66,46	6
Mistral	96,08	77,50	85,42	83,33	85,58	4

Таблица 4. Обща оценка на предложенията и класиране на моделите

В допълнение към обобщената таблица с резултатите, в изследването са представени и изчисления на стандартната девиация на оценките, предоставени от различните AI

оценители за съответните модели. Анализът на стандартната девиация предоставя информация за разпределението на оценките и нивото на съгласуваност между оценителите. По-ниски стойности на стандартната девиация показват по-висока консистентност, докато по-високите стойности сигнализират за по-големи разлики в интерпретацията на резултатите. Тази метрика допринася за по-обективна оценка на представянето на езиковите модели.

Метрика	Модел Gemini	Модел Open AI	Модел Claude	Модел DeepSeek	Модел Llama	Модел Mistral
Ст. девиация между оценители - категория Логистика	1,25	10,48	7,32	5,73	15,83	2,32
Ст. девиация между оценители - категория Инф. Технологии	4,02	3,75	2,60	7,11	5,73	6,25
Ст. девиация между оценители - категория Финанси	1,44	6,41	3,31	3,54	14,49	1,91
Ст. девиация между оценители - категория Продажби	3,61	4,39	5,00	0,72	8,75	12,52
Обща девиация за модел	10,32	25,04	18,23	17,10	41,25	23,00

Таблица 5. Стандартно отклонение между AI оценителите на предложенията

Анализът на експерименталните резултати позволява извеждането на няколко ключови изводи относно ефективността на шестте разгледани големи езикови модела (LLMs) при генериране на предложения за оптимизация на бизнес процеси в различни категории:

- Водещите модели демонстрират стабилност в представянето си в различни бизнес контексти. По-конкретно, Gemini, DeepSeek и Mistral последователно получават високи оценки, което показва способността им да генерират добре структурирани и приложими предложения за оптимизация на бизнес процеси.
- Gemini и DeepSeek имат най-ниска стандартна девиация (съответно 10.32 и 17.10), което свидетелства за висока степен на консистентност и предвидимост на резултатите, независимо от използвания оценител.
- OpenAI и Llama показват най-висока вариация в оценките (25.04 и 41.25), което предполага по-непостоянно представяне, вероятно в зависимост от спецификата на бизнес процеса или критериите, зададени от оценителите. Тази разлика може да е резултат от различния стил на генериране на оптимизационни предложения, което затруднява осигуряването на консистентност в оценяването.

На база на тези резултати, изследването прави важен практически извод: Gemini и DeepSeek са моделите с най-голяма стабилност в оценките, но съществува ключова разлика в тяхната достъпност. Gemini е платен модел, докато DeepSeek е безплатен модел с отворен код. Това разграничение е съществено при реалната имплементация на предложената методология в бизнес среда, тъй като организациите могат да направят избор между платен модел с потенциално по-висока производителност и безплатна алтернатива с ниска вариация в оценките. Следователно, изборът на езиков модел за интеграция в BPM системи ще зависи не само от неговата ефективност при генериране на оптимизационни предложения, но и от финансовите възможности и стратегическите цели на организацията.

6. МЕТОДОЛОГИЯ ЗА АВТОМАТИЗИРАНО ГЕНЕРИРАНЕ НА ПРЕДЛОЖЕНИЯ ЗА ОПТИМИЗАЦИЯ НА БИЗНЕС ПРОЦЕСИ

В **шеста** глава на дисертационния труд е разработена методологична рамка, предназначена за автоматизирано извличане на бизнес правила и генериране на предложения за оптимизация на бизнес процеси чрез използване на големи езикови модели (LLMs). Предложената методология интегрира достижения от областта на изкуствения интелект, обработката на естествен език и управлението на бизнес процеси (BPM), като акцентира върху възпроизводимостта, мащабируемостта и обективността на резултатите.

Целта на разработената методология е да трансформира регулаторни текстове и процесни модели в приложими, проверими и контекстово релевантни предложения за подобрене на бизнес процеси. Това се постига чрез структуриране на процеса в шест функционално обособени и логически свързани етапа, които обхващат пълния жизнен цикъл на обработка – от събирането на източници до оценката на генерираните предложения.

Методологията е представена схематично на Фигура 22 и включва следните етапи:

Етап 1: Събиране и структуриране на данни

Цел: Изграждане на унифициран корпус от нормативни документи и BPMN модели, който да послужи като вход към последващите процеси по извличане и генериране на предложения.

Дейности:

- Ръчен подбор на нормативни документи от отворени регистри и институционални източници.
- Импортиране на BPMN модели под формата на CSV файлове.
- Съхранение на метаданни и съдържание в релационна база и хранилище за данни.

Инструменти: Релационна база данни с REST API достъп, Централизирано хранилище за файлове, Уеб-базирана платформа за въвеждане, обработка и управление на текстови документи

Резултат: Подготвен корпус от нормативни документи и BPMN модели, съхранявани в структурирана среда, подходяща за автоматизирана обработка.

Етап 2: Фрагментиране и векторизация на нормативни текстове

Цел: Подготовка на нормативните документи за семантично търсене чрез прилагане на техники за фрагментиране и векторно представяне.

Дейности:

- Прилагане на различни методи за фрагментиране на документи.
- Векторизация чрез мултиезиков векторен модел.
- Съхранение на векторите във векторна база данни.

Инструменти: Мултиезиков модел за векторизация на текстови фрагменти, Облачна векторна база данни за индексирани и търсене по семантична близост, Изчислителна среда, предоставяща възможност за използване на GPU ресурси.

Резултат: Векторно индексирани фрагменти от нормативни документи, готови за извличане чрез семантични заявки.

Етап 3: Извличане и формулиране на бизнес правила

Цел: Формулиране на бизнес правила, отговарящи на нормативни изисквания, чрез комбиниране на векторно извлечени фрагменти и потребителски заявки.

Дейности:

- Създаване на потребителски запитвания.
- Извличане на релевантни фрагменти чрез косинусова семантична близост.
- Генериране на бизнес правила чрез голям езиков модел.
- Структуриране и превод на английски на генерираните бизнес правила чрез голям езиков модел.

Инструменти: Голям езиков модел с генеративни възможности, Изчислителна среда, предоставяща възможност за използване на GPU ресурси, Векторна база данни.

Резултат: Структуриран списък от бизнес правила на български и английски език.

Етап 4: Оценка на бизнес правилата

Цел: Автоматизирана верификация на качеството на генерираните правила чрез AI-базирана оценка.

Дейности:

- Формулиране на критерии: точност, пълнота, съответствие с контекста.
- Оценка от три големи езикови модела, с разсъждаващи способности.
- Представяне на резултатите чрез числова скала (0–100) и стандартно отклонение.

Инструменти: Големи езикови модели с генеративни и оценъчни възможности, Изчислителна среда, предоставяща възможност за използване на GPU ресурси.

Резултат: Количествена и качествена оценка на всеки от методите за фрагментиране и векторизация чрез изведените бизнес правила.

Етап 5: Генериране на предложения за оптимизация

Цел: Създаване на персонализирани предложения за подобрене на бизнес процеси чрез анализ на BPMN модели и бизнес правила, базирани на нормативни документи.

Дейности:

- Комбиниране на BPMN модели, правила и инструкции.
- Подаване на вход към шест LLM модела.
- Генериране на оптимизационни предложения.

Инструменти: Големи езикови модели с генеративни възможности, Изчислителна среда, предоставяща възможност за използване на GPU ресурси.

Резултат: Набор от предложения за оптимизация, съобразени с нормативния и процесния контекст.

Етап 6: Оценка на предложенията и избор на оптимален модел

Цел: Идентифициране на най-надеждния генеративен модел за реална имплементация чрез сравнителен анализ.

Дейности:

- AI-базирана оценка чрез три големи езикови модела, с разсъждаващи способности.
- Изчисляване на средни стойности и стандартни отклонения.
- Крайна класификация на моделите според стабилност и ефективност.

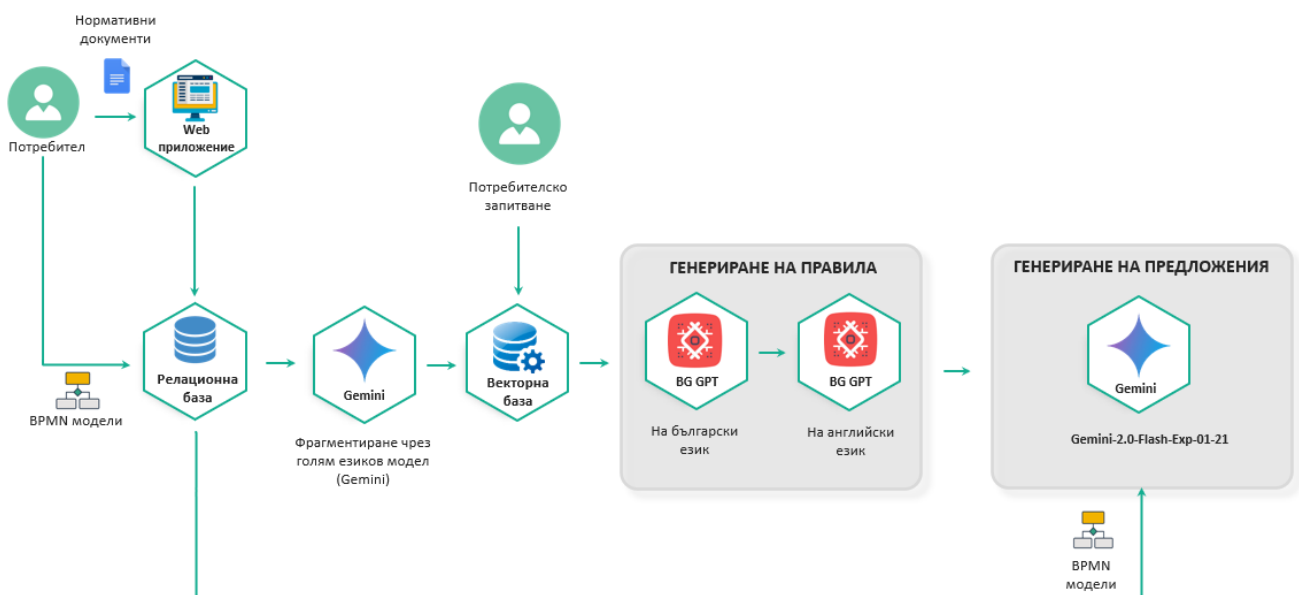
Инструменти: Големи езикови модели с генеративни и оценъчни възможности, Изчислителна среда, предоставяща възможност за използване на GPU ресурси.

Резултат: Избор на LLM модел за бъдеща интеграция в BPM система.



Фигура 22. Методологични етапи

В допълнение към методологията, в рамките на изследването е предложена интегрирана архитектура за реална имплементация, визуализирана на Фигура 23.



Фигура 23. Интегрирана архитектура за реална имплементация

Архитектурата следва принципите на Retrieval-Augmented Generation (RAG) и е съставена от четири взаимосвързани модула:

- **Модул за въвеждане на данни**, който позволява управление на нормативни документи и BPMN модели чрез уеб интерфейс;
- **Модул за фрагментиране и векторизация**, в който се използва най-добре представеният се метод – семантично фрагментиране чрез LLM (напр. Gemini);
- **Модул за генериране на бизнес правила**, който комбинира потребителски заявки с релевантни фрагменти от нормативната база и генерира съдържание на български и английски език;
- **Модул за генериране на предложения за оптимизация**, който, чрез използване на избран езиков модел (напр. Gemini 2.0 Flash), създава контекстуализирани предложения, съобразени с конкретния процесен модел и регулаторна рамка.

Интегрираният модел демонстрира както функционалната свързаност между компонентите, така и възможностите за внедряване в съществуващи BPM системи. Сред основните предимства се открояват:

- възможност за адаптация към нови нормативни изисквания без нужда от преобучение;
- проследимост и прозрачност на трансформациите;
- висока степен на повторяемост, стабилност и скалируемост.

Заклучителната част на главата обобщава научната стойност на предложената методология и архитектура. Те представляват интеграция на няколко съвременни парадигми – семантично извличане на знание, векторно търсене, генеративно съдържание, и AI-базирана оценка (LLM-as-Judge). Предложената рамка създава надеждна основа както за експериментални изследвания, така и за реални приложения в домейна на интелигентното управление на бизнес процеси.

7. ЗАКЛЮЧЕНИЕ

Настоящото изследване представя методологична рамка за автоматизирана оптимизация на бизнес процеси, базирана на комбинирането на методи, имплементиращи големи езикови модели (LLMs), техники за семантично фрагментиране

на нормативни документи и AI-базирана оценка на резултатите. Проведените експерименти и анализи потвърждават валидността на формулираните хипотези и демонстрират, че предложените методи, осигуряват надеждни и последователни резултати.

По отношение на Хипотеза I, резултатите от изследването показаха, че интегрирането на големи езикови модели с векторни и релационни бази данни създава ефективна среда за автоматизирано извличане и анализ на нормативни документи с цел генериране на бизнес правила и предложения за оптимизация. Постигнатата висока консистентност в оценките на AI моделите демонстрира, че такава методологична рамка може да бъде успешно приложена за подпомагане на бизнес процесите, като подобрява разбирането на контекста и предлага смислени оптимизации.

Относно Хипотеза II, изследването потвърди значението на метода за фрагментиране на нормативните документи върху качеството на генерираните предложения за оптимизация. Резултатите ясно показаха, че използването на голям езиков модел за контекстуално фрагментиране е най-ефективният метод, осигуряващ минимална загуба на контекст и по-добра структурираност на текстовата информация. Въпреки предходни изследвания, които подчертават ефективността на BM25 индекса в информационното извличане, настоящите резултати не показаха значително подобрене при неговото прилагане в контекста на нормативни документи, което подчертава необходимостта от адаптиране на методите спрямо спецификата на обработваните текстови източници.

Що се отнася до Хипотеза III, изследването успешно демонстрира, че нормативната рамка оказва измеримо въздействие върху бизнес процесите и тяхната оптимизация. Използването на AI оценители предостави обективна оценка на генерираните бизнес правила, като позволи анализ на последиците от законодателните промени върху бизнес процесите. Получените резултати доказаха, че автоматизираната обработка на нормативни документи може не само да улесни разбирането на законодателните изисквания, но и да подпомогне адаптацията на бизнес процесите спрямо тях.

В заключение, настоящото изследване не само потвърди валидността на формулираните хипотези, но и предостави ясна методологична рамка, която може да бъде използвана за интеграция в реални BPM системи. Получените резултати демонстрират, че

комбинирането на големи езикови модели, семантично фрагментирани и AI оценка може значително да подпомогне адаптацията и оптимизацията на бизнес процесите в контекста на динамично променящата се нормативна среда.

ЛИТЕРАТУРА

- [1] J. Jeston and J. Nelis, *Business Process Management: Practical Guidelines to Successful Implementations*, 2nd ed., Routledge, 2008.
- [2] С. Димков, *Бизнес процеси*, Изд. СофтПрес, 2017.
- [3] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*, 2nd ed. Springer, 2018.
- [4] M. Hammer, "Reengineering Work: Don't Automate, Obliterate," *Harvard Business Review*, vol. 68, no. 4, pp. 104–112, 1990.
- [5] T. H. Davenport, *Process Innovation: Reengineering Work Through Information Technology*, Harvard Business School Press, 1993.
- [6] M. Rosemann and J. vom Brocke, "The Six Core Elements of Business Process Management," in *Handbook on Business Process Management 1: Introduction, Methods, and Information Systems*, 2nd ed., J. vom Brocke and M. Rosemann, Eds. Springer, 2015, pp. 105–122.
- [7] M. Hammer and J. Champy, *Reengineering the Corporation: A Manifesto for Business Revolution*, HarperBusiness, 1993.
- [8] P. Harmon, *Business Process Change: A Business Process Management Guide for Managers and Process Professionals*, 3rd ed., Morgan Kaufmann, 2014.
- [9] P. Harmon, *Business Process Change: A Manager's Guide to Improving, Redesigning, and Automating Processes*, 3rd ed., Morgan Kaufmann, 2014.
- [10] M. Weske, *Business Process Management: Concepts, Languages, Architectures*, 2nd ed., Springer, 2012.
- [11] T. H. Davenport, "Putting the Enterprise into the Enterprise System," *Harvard Business Review*, vol. 76, no. 4, pp. 121–131, 1998.
- [12] Wil M. P. van der Aalst, *Process Mining: Data Science in Action*, 2-ро издание, Springer, 2016.
- [13] W.M.P. van der Aalst, H.A. Reijers, A.J.M.M. Weijters, B.F. van Dongen, A.K. Alves de Medeiros, M. Song, and H.M.W. Verbeek, "Business Process Mining: An Industrial Application," *Information Systems*, vol. 32, no. 5, pp. 713–732, 2007.
- [14] Mathias Weske, *Business Process Management: Concepts, Languages, Architectures*, 2-ро издание, Springer, 2012.

- [15] Gary Marcus и Ernest Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust*, 2019.
- [16] Haenlein, M., & Kaplan, A. (2021). "Artificial intelligence and robotics: Shaking up the business world and society at large." *Journal of Business Research*, 124, 405–407.
- [17] Luciana Barbieri, Edmundo Madeira, Kleber Stroeh и Wil van der Aalst, "A Natural Language Querying Interface for Process Mining," *Journal of Intelligent Information Systems*, том 61, брой 1, страницы 113–142, 2023.
- [18] Paul Harmon, *Business Process Change: A Business Process Management Guide for Managers and Process Professionals*, 4-то издание, 2019.
- [19] Kotter, J. P. (1996). *Leading Change*. Boston: Harvard Business School Press.
- [20] Davenport, T. H., & Short, J. E. (1990). "The New Industrial Engineering: Information Technology and Business Process Redesign." *Sloan Management Review*, 31(4), 11–27.
- [21] Hammer, M. (2010). What is Business Process Management?. In J. vom Brocke & M. Rosemann (Eds.), *Handbook on Business Process Management 1* (pp. 3–16). Springer.
- [22] Jan Mendling, Matthias Weidlich и Mathias Weske (ред.), *Business Process Modeling Notation: Second International Workshop, BPMN 2010, Potsdam, Germany, October 13-14, 2010, Proceedings*, Springer, 2010.
- [23] Howard Smith и Peter Fingar, *Business Process Management: The Third Wave*, Meghan-Kiffer Press, 2003.
- [24] Hollingsworth, D. (2004). "The Workflow Reference Model: 10 Years On." In Fischer, L. (Ed.), *Workflow Handbook 2004* (pp. 295–312). Lighthouse Point, FL: Future Strategies Inc.
- [25] Sayal, M., Casati, F., Dayal, U., & Shan, M. C. (2002). "Business Process Management Systems." In *Lecture Notes in Computer Science*, vol. 2469, pp. 10–16. Springer
- [26] Basu, A., & Kumar, A. (2002). "Research Commentary: Workflow Management Issues in e-Business." *Information Systems Research*, 13(1), 1–14.
- [27] Aldin, L., & de Cesare, S. (2011). "A literature review on business process modelling: New frontiers of reusability." *Enterprise Information Systems*, 5(3), 359–383.
- [28] Kassim, S. A., Gartner, J.-B., Labbé, L., Landa, P., Paquet, C., Bergeron, F., Lemaire, C., & Côté, A. (2022). "Benefits and limitations of business process model notation in modelling patient healthcare trajectory: a scoping review protocol." *BMJ Open*, 12(5), e060357.
- [29] Smirnov, S., Weidlich, M., Mendling, J., & Weske, M. (2012). "Action Patterns in Business Process Model Repositories." *Computers in Industry*, 63(2), 98–111.
- [30] Polyvyanyy, A., Smirnov, S., & Weske, M. (2019). "Process Model Abstraction: A Slider Approach." In *Conceptual Modeling – ER 2019* (pp. 119–135). Springer.
- [31] Leopold, H., Mendling, J., & Günther, O. (2018). "Learning from Quality Issues of BPMN Models from Industry." *IEEE Software*, 35(2), 26–33.

- [32] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv preprint arXiv:2005.11401.
- [33] Karpukhin, V., Oguz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., & Yih, W.-t. (2020). Dense Passage Retrieval for Open-Domain Question Answering. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 6769–6781.
- [34] Robertson, S., & Zaragoza, H. (2009). The Probabilistic Relevance Framework: BM25 and Beyond. Foundations and Trends® in Information Retrieval, 3(4), 333–389
- [35] Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). "Stanza: A Python Natural Language Processing Toolkit for Many Human Languages." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 101–108.
- [36] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171–4186). Association for Computational Linguistics.
- [37] Terčon, Luka, and Nikola Ljubešić. CLASSLA-Stanza: The Next Step for Linguistic Processing of South Slavic Languages. Preprint, August 14, 2023. University of Ljubljana, Jožef Stefan Institute, arXiv:2308.07321.
- [38] Anton Alexandrov, Veselin Raychev, Mark Niklas Mueller, Ce Zhang, Martin Vechev, and Kristina Toutanova. 2024. Mitigating Catastrophic Forgetting in Language Transfer via Model Merging. In Findings of the Association for Computational Linguistics: EMNLP 2024, pages 17167–17186, Miami, Florida, USA. Association for Computational Linguistics.
- [39] Ivan Georgiev, Nikolay Petrov, Elena Dimitrova, and Stoyan Nikolov. 2024. BG GPT: Building Bulgarian Large Language Models at INSAIT. In Proceedings of the International Conference on Natural Language and Speech Processing (ICNLSP 2024), pages 235–249, Sofia, Bulgaria. Association for Computational Linguistics.
- [40] Alexandrov, A., Raychev, V., Dimitrov, D. I., Zhang, C., Vechev, M., & Toutanova, K. BgGPT1.0: Extending English-centric LLMs to other languages, arXiv:2412.10893, 2024.
- [41] Klouček, Bozhidar, and Riza Batista-Navarro. Bulgarian Grammar Error Correction with Data Augmentation and Machine Translation Techniques. In Proceedings of the 7th International Conference on Natural Language and Speech Processing (ICNLSP 2024), edited by Mourad Abbas and Abed Alhakim Freihat, 365–376. Trento: Association for Computational Linguistics, 2024.
- [42] Sola, Diana, Christian Warmuth, Bernhard Schäfer, Peyman Badakhshan, Jana-Rebecca Rehse, and Timotheus Kampik. SAP Signavio Academic Models: a large process model dataset. International Conference on Process Mining, 453–465. Springer.

- [43] A. Rebmann, F. D. Schmidt, G. Glavaš and H. van Der Aa, Evaluating the Ability of LLMs to Solve Semantics-Aware Process Mining Tasks, 2024 6th International Conference on Process Mining (ICPM), Kgs. Lyngby, Denmark, 2024, pp. 9-16, doi: 10.1109/ICPM63005.2024.10680677.
- [44] Guan, Wei, Jian Cao, Jianqi Gao, Haiyan Zhao, and Shiyong Qian. Dabl: Detecting semantic anomalies in business processes using large language models. arXiv preprint arXiv:2406.15781 (2024).
- [45] Kourani, Humam, Alessandro Bertj, Daniel Schuster, and Wil MP van der Aalst. Evaluating Large Language Models on Business Process Modeling: Framework, Benchmark, and Self-Improvement Analysis, arXiv:2412.00023 (2024).
- [46] Busch, Kiran, and Henrik Leopold. Towards a Benchmark for Large Language Models for Business Process Management Tasks, arXiv:2410.03255 (2024).
- [47] Van der Aa, H., Carmona Vargas, J., Leopold, H., Mendling, J., Padró, L.: Challenges and opportunities of applying natural language processing in business process management. In: COLING 2018: The 27th International Conference on Computational Linguistics: Proceedings of the Conference: August 20–26, 2018 Santa Fe, New Mexico, USA. pp. 2791–2801. Association for Computational Linguistics (2018)
- [48] Grohs, Michael, Luka Abb, Nourhan Elsayed, and Jana-Rebecca Rehse. Large language models can accomplish business process management tasks, International Conference on Business Process Management, pp. 453-465, Springer Nature Switzerland, 2023.
- [49] Vidgof, M., Bachhofner, S., Mendling, J. (2023). Large Language Models for Business Process Management: Opportunities and Challenges. In: Di Francescomarino, C., Burattin, A., Janiesch, C., Sadiq, S. (eds) Business Process Management Forum. BPM 2023. Lecture Notes in Business Information Processing, vol 490. Springer, Cham. https://doi.org/10.1007/978-3-031-41623-1_7
- [50] Estrada-Torres, B., del-Río-Ortega, A., Resinas, M. (2024). Mapping the Landscape: Exploring Large Language Model Applications in Business Process Management. In: van der Aa, H., Bork, D., Schmidt, R., Sturm, A. (eds) Enterprise, Business-Process and Information Systems Modeling. BPMDS EMMSAD 2024 2024. Lecture Notes in Business Information Processing, vol 511. Springer, Cham. https://doi.org/10.1007/978-3-031-61007-3_3
- [51] Beheshti, A., Yang, J., et al.: ProcessGPT: transforming business process management with generative artificial intelligence. IEEE ICWS, pp. 731–739 (2023). <https://doi.org/10.1109/ICWS60048.2023.00099>
- [52] van der Aa, H., Di Ciccio, C., Leopold, H., Reijers, H.A.: Extracting declarative process models from natural language. In: Giorgini, P., Weber, B. (eds.) CAiSE 2019. LNCS, vol. 11483, pp. 365–382. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21290-2_23
- [53] Lashkevich, K., Milani, F., Avramenko, M., et al. (2024). Llm-assisted optimization of waiting time in business processes: A prompting method. Business Process Management, Springer Nature, pp. 474–492

- [54] Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. Int. Conf. on Neural Information Processing Systems, Proceedings, NIPS'20
- [55] Naveed, H., Khan, AU., Qiu, S., et al. (2023). A comprehensive overview of large language models. <https://doi.org/10.48550/ARXIV.2307.06435>
- [56] Object Management Group (2011) Business Process Model and Notation (BPMN), Version 2.0. <http://www.omg.org/spec/BPMN/2.0>
- [57] Agarwal, P., Gao, B. et al (2022) A Process-Aware Decision Support System for Business Processes. In: ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Proceedings, p. 2673–2681. <https://doi.org/10.1145/3534678.3539088>
- [58] Agostinelli, S., De Luzi, F., Di Canito, U. et al (2022) A data-centric approach to design resilient-aware process models in BPMN. In: Business Process Management Forum - BPM 2022 Forum, Proceedings, Springer, pp. 38–54. https://doi.org/10.1007/978-3-031-16171-1_3
- [59] Fahland, D., Fournier, F., Limonad, L., et al. (2024) How well can large language models explain business processes? CoRR. <https://doi.org/10.48550/ARXIV.2401.12846>
- [60] Fontenla-Seco, Y., Lama, M., Bugarín, A. (2020) Process-To-Text: A Framework for the Quantitative Description of Processes in Natural Language. Trustworthy AI - Integrating Learning, Optimization and Reasoning - Workshop, TAILOR 2020, LNCS, vol. 12641. Springer, pp. 212–219
- [61] Goel, Shashwat, Joschka Struber, Ilze Amanda Auzina, Karuna K. Chandra, Ponnurangam Kumaraguru, Douwe Kiela, Ameya Prabhu, Matthias Bethge, and Jonas Geiping. Great Models Think Alike and this Undermines AI Oversight., arXiv:2502.04313 (2025).
- [62] Liu, Y., Moosavi, N., and Lin, C. LLMs as narcissistic evaluators: When ego inflates evaluation scores. In ACL, 2024. URL <https://aclanthology.org/2024.findings-acl.753/>.
- [63] Bedemariam, Rewina, Natalie Perez, Sreyoshi Bhaduri, Satya Kapoor, Alex Gil, Elizabeth Conjar, Ikkei Itoku, David Theil, Aman Chadha, and Naumaan Nayyar. Potential and perils of large language models as judges of unstructured textual data. arXiv:2501.08167 (2025).
- [64] Thakur, A. S., Choudhary, K., Ramayapally, V. S., Vaidyanathan, S., & Hupkes, D. (2024). Judging the Judges: Evaluating Alignment and Vulnerabilities in LLMs-as-Judges. arXiv preprint arXiv:2406.12624.
- [65] Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E. P., Zhang, H., Gonzalez, J. E., & Stoica, I. (2023). *Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena*. arXiv preprint arXiv:2306.05685.
- [66] Gu, Jiawei, Xuhui Jiang, Zhichao Shi, Hexiang Tan, Xuehao Zhai, Chengjin Xu, Wei Li et al. A Survey on LLM-as-a-Judge. arXiv:2411.15594 (2024).
- [67] Zhu, L., Wang, X., & Wang, X. (2023). *JudgeLM: Fine-tuned Large Language Models are Scalable Judges*. arXiv preprint arXiv:2310.17631.

- [68] Shi, L., Ma, W., & Vosoughi, S. (2024). *Judging the Judges: A Systematic Investigation of Position Bias in Pairwise Comparative Assessments by LLMs*. arXiv preprint arXiv:2406.07791.
- [69] Yao, S., Cao, S., Yu, A. W., Narang, S., & Abbeel, P. (2023). Tree of Thoughts: Deliberate Problem Solving with Large Language Models. arXiv preprint arXiv:2305.10601.
- [70] Nye, M., Tessler, M., Wu, J., Wang, X., & Tenenbaum, J. (2021). Improving Language Model Reasoning with Self-Reflection. NeurIPS 2021.
- [71] Wei, J., Wang, X., Schuurmans, D., Bosma, M., & Chi, E. (2022). Chain of Thought Prompting Elicits Reasoning in Large Language Models. arXiv preprint arXiv:2201.11903.
- [72] Ouyang, L., Wu, J., Jiang, J., Chen, D., & Zhang, X. (2023). Reinforcement Learning from Human Feedback: Current Trends and Future Directions. Proceedings of NeurIPS.
- [73] Bubeck, S., Chandrasekaran, V., Eldan, R., & Gehrke, J. (2023). Sparks of Artificial General Intelligence: Early Experiments with GPT-4. arXiv preprint arXiv:2303.12712.
- [74] Bommasani, R., Hudson, D. A., Adeli, E., & Bai, Y. (2022). On the Opportunities and Risks of Foundation Models. Journal of Artificial Intelligence Research.
- [75] Marcus, G., & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon Books.
- [76] Anthropic, 2024, Introducing Contextual Retrieval, <https://anthropic.com/news/contextual-retrieval>
- [77] Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. J.-F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research, 70, 356–365.
- [78] Peng, B., Li, S., He, X., Galley, M., Zhang, W., Gao, J., & Liu, J. (2023). Instruction tuning with GPT-4. arXiv preprint arXiv:2304.03277. Retrieved from <https://arxiv.org/abs/2304.03277>
- [79] Wang, Y., Kordi, Y., Mishra, S., Liu, P., Smith, N. A., & Hajishirzi, H. (2022). Self-instruct: Aligning language models with self-generated instructions. arXiv preprint arXiv:2212.10560. Retrieved from <https://arxiv.org/abs/2212.10560>

СПИСЪК ФИГУРИ

Фигура 1. Съставни части на бизнес процес	17
Фигура 2. Жизнен цикъл при управлението на бизнес процеси	19
Фигура 3. Концептуална рамка	23
Фигура 4. Концептуална рамка – първи компонент	26
Фигура 5. Концептуална рамка – втори компонент	28
Фигура 6. Технологична рамка	31
Фигура 7. Бизнес модел "Логистика"	39
Фигура 8. Бизнес модел "Информационни технологии"	39
Фигура 9. Бизнес модел "Финанси"	40

Фигура 10. Бизнес модел "Продажби"	40
Фигура 11. Генериране на съдържание, обогатено с извличане на информация (RAG) ...	44
Фигура 12. Метод за фрагментиране и векторизация чрез семантична близост на изречения	47
Фигура 13. Метод за фрагментиране и векторизация чрез семантична клъстеризация ...	49
Фигура 14. Метод за фрагментиране и векторизация чрез голям езиков модел	50
Фигура 15. Метод за фрагментиране и векторизация чрез голям езиков модел	51
Фигура 16. Метод за фрагментиране и векторизация чрез голям езиков модел и BM25	53
Фигура 17. Метод за генериране на бизнес правила	55
Фигура 18. Метод за оценка на фрагментиране и векторизация	61
Фигура 19. Метод за генериране на предложения за оптимизация	64
Фигура 20. Метод за оценка на генерираните предложения за оптимизация	69
Фигура 21. Методологични етапи	77
Фигура 22. Методологични етапи	78
Фигура 23. Интегрирана архитектура за реална имплементация	78

СПИСЪК ТАБЛИЦИ

Таблица 1. Резултати от подходите за фрагментиране – първо запитване	62
Таблица 2. Обща оценка на резултатите от фрагментиране	62
Таблица 3. Стандартно отклонение при различните подходи за фрагментиране	63
Таблица 4. Обща оценка на предложенията и класиране на моделите	72
Таблица 5. Стандартно отклонение между AI оценителите на предложенията	73