Riri Ai Chatbot 要件定義書

プロジェクトの目的

riri システムは法律系のチャットボットである。
現在流通しているチャットボットには争いなどの分析には弱く、その折衷案や過失割合などを評価し、中立的に和解する機能は存在していない。また法律や価値観などは国によって違うため、他の国は利用することはできないため、本プロジェクトで開発することとする。
本アプリケーションはユーザーからの応答を分析し、中立的な立場から折衷案や過失割合を分析する。

利用者運用環境

・利用者の種類：一般人、一般社員

・利用場所・利用端末：スマートフォン、タブレット、パソコンなどアプリケーション、Web サイトでの運用

・システム稼働時間：365 日 24 時間

機能要件

1.**ユーザー入力関連機能：**
・ユーザーが法律相談やトラブル内容を自然言語で入力できる機能

・入力内容から争点（例：契約違反、交通事故、損害賠償など）を自動抽出する機能

・複数の当事者（例：A さん・B さん）を識別し、それぞれの主張を整理する機能

2.**法律知識ベース連携機能**
・日本の法律データベース（民法、商法、労働法など）を参照し、関連条文を提示する機能

・最新の法改正情報を反映できる更新機能

・法的な根拠を自然言語で説明する補助機能（例：「民法第 709 条に基づき～」など）

3.**分析/評価機能**
・各当事者の主張を分析し、過失割合を推定する機能

・判例データや過去事例をもとに、類似ケース分析を行う機能

・双方の主張をもとに、*中立的な折衷案(和解案)*を自動生成する機能

・分析の根拠を説明(Explainable AI)として出力する機能

4.**応答生成機能**

・ユーザーに対して自然な日本語で回答を提示する機能

・法律的判断と中立的コメントを分けて表示(例:「法的観点」+「調停案」)

・シナリオに応じた複数回答パターンを生成(例:「厳格な法解釈」/「柔軟な和解提案」)

5.**設定パーソナライズ機能**

・分野(交通、契約、労務、家族など)をユーザーが選択できる機能

・応答トーン(フォーマル/カジュアル)を選択できる機能

・利用者属性(一般ユーザー、弁護士、企業法務担当など)によって出力レベルを切り替える機能

6.**学習/改善機能**

・ユーザーからのフィードバックを蓄積し、分析精度を向上させる機能

・Chat ログを匿名化して学習データに反映する機能

・専門家監修によるモデル修正が可能な管理者向けインターフェース

7.**システム管理機能**

・ユーザー認証(メール・Google アカウントなど)

・利用履歴・相談履歴の保存と検索機能

・管理者画面で相談内容の統計分析や改善指標を確認できる機能

8.**非機能要件との関連機能**

・セキュリティ：個人情報を含む相談内容を暗号化保存

・スケーラビリティ：アクセス数増加にも対応可能な構成

・法的適合性：弁護士法などの規制に抵触しない設計

非機能要件

1.**性能**__must__

目的：ユーザにストレスなく応答させる

指標・目標値（例）：

・テキスト応答レイテンシ（初期応答）：平均 ≤ 800 ms、95 パーセンタイル ≤ 1500 ms

・音声応答（音声認識＋生成→テキスト→返答→TTS）：平均 ≤ 3.0 s、95 パーセンタイル ≤ 5.0 s

・同時接続数（想定初期負荷）：1,000 同時ユーザーに対してターゲット応答を維持（スケール計画は別途）

・スループット（API）：100 req/s（ピーク時設計例。実運用で再評価）

受け入れ基準／テスト方法：負荷試験（JMeter/Gatling 等）で上記指標を満たすこと。

2.**可用性/信頼性**__must__

目的：サービス稼働を確保し、ダウン時の影響を最小化する

指標・目標値：

・SLA 目標：稼働率 99.9%（年間ダウンタイム ≤ 8.76 時間）※要調整

・フェールオーバー：プライマリ障害時の切替 ≤ 60 秒

・バックエンド（モデル推論レイヤー）は冗長化（複数 AZ/ノード）

受け入れ基準：障害注入（Chaos）テストで自動切替が機能すること。

3.**セキュリティ・プライバシー**__shuld__

目的：機密相談情報を適切に保護し、法的リスクを回避する。

要件：

・通信：TLS 1.2+ 強制。外部 API 連携も TLS 必須。

・保存データ：個人情報含む相談は暗号化保存（AES-256 at rest）および転送時暗号化。

・アクセス管理：RBAC（ロールベースアクセス制御）、管理操作は多要素認証（MFA）。

・ログの保護：監査ログは改竄防止（WORM、ハッシュチェーン等）で保存。

・データ持ち出し制御：エクスポートは管理者承認とログ記録。

・準拠規格：国内法（個人情報保護法）準拠。機密度高ければオンプレまたは専用 VPC／プライベートクラウドを検討。

・匿名化：学習用に利用するデータは自動匿名化／個人識別子除去のパイプラインを必須とする。
受け入れ基準：セキュリティ監査（外部脆弱性診断）、ペネトレーションテスト合格。

4.**法的適合性・コンプライアンス**__must__
目的：弁護士法や広告規制等に抵触しない運用。
要件：

・本システムは「法的助言の代替」ではなく「一般的情報・中立案提示」である旨を明示（利用規約／UI）

・必要に応じて「弁護士への相談推奨フロー」を組み込む。

・出力には*出典・根拠（条文・判例）*を自動添付し、引用元をトレース可能にする。
受け入れ基準：法務レビュー済みの文言・フローでローンチ。

5.**モデルガバナンス（Model Governance）**__must__
目的：モデルの振る舞いを管理し、説明責任を確保する。
要件：

・モデルバージョン管理（モデル ID・デプロイ履歴の保存）

・推論結果に対する説明（説明可能性：why の説明）：根拠条文、類似判例、推定過失の根拠スコア等の提示

・モニタリング：誤応答率、低信頼回答の割合、ユーザーフィードバック率を監視

・ドリフト検知：データ分布/予測分布の変化を検出したらアラート
指標例：ある期間内で信頼度＜閾値の応答は 1%未満（目標）
受け入れ基準：モデル切替時の AB テストで性能/公平性が悪化しないこと確認。

6.**公平性・バイアス対策（Fairness）（法律領域では特に重要）**__must__
目的：特定属性に不利な推定を避ける。
要件：

・性別・年齢・人種と思われる属性で不当に偏る出力がないことをテスト。

・バイアス検出ルールと是正プロセス（モデル再学習・ルールベース修正）を定義。
受け入れ基準：主要な属性群で公平性指標（例：差分が閾値内）を満たす。

7.**説明可能性（Explainability）**__must__
目的：ユーザー／管理者が判断根拠を理解できるようにする。
要件：

・各分析（過失割合、和解案）に根拠のスコアリング（例：証拠重み、類似判例の一致度）を付与

・出典の明示（「参照条文」「類似判例 ID」「参照データセット名」）
受け入れ基準：サンプルケースで人間レビューが説明を理解・追認可能であること。

8.**ログ・監査（Logging & Auditability）**__must__
目的：後追い調査とコンプライアンス対応を可能にする。
要件：

・入力・出力・モデルバージョン・タイムスタンプを保存（一定期間）

・ログは改竄耐性を持ち、管理者操作も全てログ化

・プライバシーのため保存ログは自動マスキング/匿名化のオプション
受け入れ基準：監査時に直近 6 ヶ月分の証跡が提示できること（期間は法要件に合わせ可変）。

9.**音声認識(ASR)・音声合成(TTS)**__must__

目的:音声での相談を可能にする(対話の幅を拡げる)。

ASR 指標:

・日本語での WER(Word Error Rate)目標: ≤ 10%(雑音環境での目標は要調整)

・リアルタイム処理:ストリーミング入力での部分認識レスポンス ≤ 300 ms(部分応答)

TTS 指標:

・応答開始遅延 ≤ 500 ms(テキスト→音声)

・音質は人間に近い自然度(MOS 評価で ≥ 4.0／5.0 を目安)

受け入れ基準:代表的コーパス(会話式、騒音あり)での WER 評価、ユーザーテスト。

10.**感情システム(Emotion / Sentiment)**__should__

目的:ユーザーの感情状態を推定し、回答のトーンや推奨フロー(例:緊急性の判定)を調整する。

要件:

・感情分類:ポジティブ／ニュートラル／ネガティブ＋怒り・悲しみ・恐れなどの細分類

・精度目標:F1 スコア ≥ 0.80(代表的テストセット)※要データ準備

・感情推定は「機械推定(confidence)」を明示し、低信頼時は保守的な応答に変更

受け入れ基準:テスト会話データで F1≥0.8／誤判定が致命的影響を与えない設計。

11.**形態素解析・言語解析**__must__

目的:日本語の法律相談文を正確に解析して争点抽出を行う。

要件:

・日本語形態素解析の精度(名詞抽出・固有表現認識)を高めるため、ドメイン固有辞書を導入(法律用語、判例表現)

・固有表現抽出(当事者名、日付、金額、条文名)の抽出精度目標:F1 ≥ 0.90(重要)

・統合:形態素解析→依存構造解析→意味役割ラベリング(SRL)を通じて事件要素を構築

受け入れ基準：法的相談コーパスでの NER/固有表現抽出評価を満たすこと。

12.**ローカライゼーション（Localization）**__must__
目的：法律・価値観が国ごとに違うため、初期は日本限定。将来拡張可能に設計。
要件：

・ローカル規則（日本版）を独立モジュール化（法律データを国別で分離）

・将来の多言語対応はモジュール追加で可能な設計（i18n 対応）
受け入れ基準：法データは国別に入れ替えられること（テストで確認）。

13.**運用性・保守性（Operability & Maintainability）** __must__
目的：モデル更新・規則変更を現場で容易に反映できるようにする。
要件：

モ・デル・ルールのデプロイは CI/CD パイプラインで自動化（ステージング→本番）

・管理画面で学習データの追加・黒リスト/白リスト管理・ルール編集が可能

・ドキュメント（API 仕様、運用手順）を自動生成・バージョン管理
受け入れ基準：非エンジニア管理者が管理画面で簡単なルール修正を実施できること。

14.**テスト可能性（Testability）**__must__
目的：要件を自動テストで検証できるようにする。
要件：

・回帰テストスイート（ユースケースごとの期待出力）を整備し、デプロイごとに実行

・自動評価用のゴールドスタンダード（正解データ）を用意
受け入れ基準：主要ユースケース（例：交通事故・契約違反）で回帰テスト合格率 ≥ 95%。

15.**スケーラビリティ（Scalability）**__should__
目的：利用増に合わせて水平スケール可能にする。
要件：

・推論サービスはコンテナ化（Kubernetes 等）でスケーリング可能

・モデル推論は GPU/TPU のスケールアウト設計を考慮
受け入れ基準：ロード増加時にオートスケールでレイテンシが許容値内に収まること。

16.**可観測性（Observability）**__must__
目的：運用中の問題を速やかに検知・対応する。
要件：

・メトリクス（リクエスト数、レイテンシ、エラー率、低信頼回答率）を蓄積

・アラートルール（閾値超過で Ops に通知）とダッシュボードを用意
受け入れ基準：主要メトリクスのダッシュボードを設置しアラートが検証可能であること。

17.**可搬性・インターフェース**__should__
目的：他システムと連携しやすくする。
要件：

・REST/GraphQL の公開 API 仕様を定義（認証は OAuth2/JWT）

・Webhook で外部システムへイベント通知可能（相談完了、重大リスク検出など）
受け入れ基準：サンプルクライアントで API 連携ができること。

18.**データ保持・削除ポリシー（Retention & Right-to-Delete）** __must__

目的：個人情報保護法等への準拠とユーザー信頼確保。
要件：

・ータ保持期間：初期は相談ログを最大 3 年保持（要法務の確認で調整）

・ユーザーが「削除要求」した場合、該当データを速やかに削除・匿名化（ログ含む）する手順を実装
受け入れ基準：削除要求から 72 時間以内に実施できるプロセスを確認。

19.**コスト制約(Cost)** __should__

目的:予算内での運用を確保。
要件:

・推論コストの見積もり(推論回数あたり)を提示し、コスト最適化(キャッシュや軽量モデルの導入)を計画
受け入れ基準:運用コスト試算を提示し、ROI の基礎を作る。

20.**障害時のフォールバック(Fallback) **__should__

目的:モデルが利用不可時でも最低限の案内ができるようにする。
要件:

モデル障害時は「定型テンプレ回答＋弁護士紹介案内」などのルートを用意
受け入れ基準:障害状態でのユーザー案内が動作すること。