Intimebot – A Dialogue Agent for Timekeeping Support

Shoaib Khan
Intapp Labs
shoaib.khan@intapp.com

Alex Samani
Intapp Labs
alex.samani@intapp.com

Rafael E Banchs
Intapp Labs
rafael.banchs@intapp.com

Abstract

This demo paper presents intimebot, an AI-powered timekeeping solution designed to assist with timekeeping. Timekeeping is a fundamental but also overwhelming and complex task in many professional services practices. Our intimebot demo demonstrates how Artificial Intelligence can be utilized to implement a more efficient timekeeping process within a firm. Based on brief work descriptions provided by the timekeeper, intimebot is able to (1) predict the relevant combination of client, matter, and phase, (2) estimate the work effort hours, and (3) rewrite and normalize the provided work description into a compliant narrative. This can save a significant amount of time for busy professionals while ensuring terms of business compliance and best practices.

1 Introduction

Timekeeping constitutes a fundamental process in professional services business operations because, when properly done, it ensures accurate and timely billing, which is a necessary condition for a healthy revenue stream for the firm. However, timekeeping is also typically a very taxing and overwhelming task to the busy professionals, who can see their potential billing hours significantly impacted due to poor timekeeping practices (Boster and Brenan, 2024).

Supporting timekeeping activities with Natural Language Processing technologies is a compelling proposition because of its potential impact on both firms and professionals. On the firm side, more timely and accurate timekeeping reduces revenue losses due to compliance issues, missed work items, and delayed billing cycles. On the professional side, proper timekeeping support can reduce the burden of clerical tasks, improve the quality of reporting, and increase the amount of effective billable hour availability.

This demo paper presents intimebot, a dialogue agent to support professional services timekeeping tasks. The system starts with a brief description of the work to be reported. From there, it interactively guides the user through the process of creating a compliant timecard, including the corresponding narrative following the required guidelines.

2 Related Work

Multiple recommendations and guidelines have been proposed to improve timekeeping practices over time (Henry, 2023; Bill4Time, 2023; Wolf, 2024). Similarly, a myriad of timekeeping solutions and tools are available (Black, 2020; Capterra, 2024; Wikipedia, 2024).

More recently, advancements in generative AI along with the corresponding enablement of agentic frameworks are steering timekeeping automation into a new era of possibilities, specifically for the case of generative AI applications (Trivedi, 2025). Our proposed intimebot demo and experimental framework represents an important incremental effort in that new direction.

3 Problem Statement

The intimebot framework focuses on the creation of a compliant timecard from a brief description of the work done, which is typically a memo entry or personal note provided by the timekeeper in a timely manner.

From such a brief work description, all timecard fields are to be estimated. These are: the client for whom the work is done, the corresponding matter and phase, as well as the estimated amount of worked hours and the narrative. The narrative must be compliant with both firm stylistic and formatting guidelines, as well as to terms of business agreed for the specific client and matter.

The brief description can be provided by means of two different modalities: text or speech.

Proceedings of the 15th International Workshop on Spoken Dialogue Systems Technology, pages 196–199
May 27–30, 2025.

4 System Description

The intimebot system implements an interactive workflow, which integrates different technologies (supervised classification, information retrieval, natural language understanding and generation) along with support data (including matter histories and timekeeper's previous entries), to estimate the complete set of timecard information. Figure 1 depicts the overall workflow of the system.

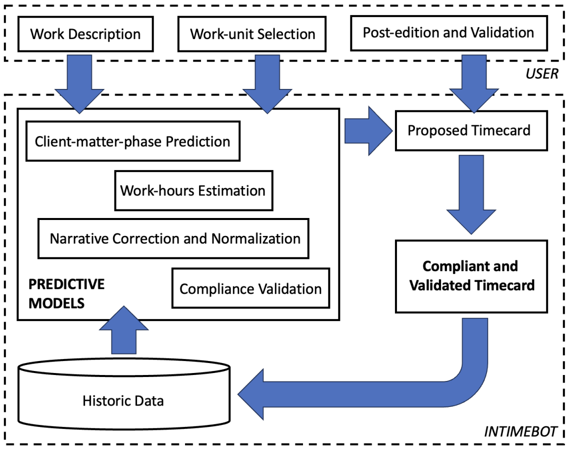

Figure 1: intimebot system components.

The "Predictive Models" block of the diagram in the figure comprises four different components:

Client-matter-phase prediction: for a given timekeeper and work description, it predicts the proper combination of client, matter and phase.

Work hours (effort) estimation: for a given work description, it estimates the expected value and distribution of the corresponding work hours.

Narrative correction and normalization: given the work description, it rewrites it into a proper narrative that follows the firm's formatting and stylistic guidelines.

Compliance validation: this sub-module checks the resulting proposed timecard against contractual commitments for potential compliance issues.

Each of these four components is described in further detail in the following subsections.

4.1 Client-Matter-Phase Prediction

Different firms use matter and phase codes in different ways. In intimebot, we refer to a work-unit as a unique combination of three elements: client, matter, and phase. This work-unit definition will be the contextual unit of analysis for the prediction problem under consideration.

The client-matter-phase prediction problem is approached in intimebot as a binary classification problem. This means that for a given description of work provided by the timekeeper, and the known history of previous work-unit contexts for the same timekeeper, the binary classifier is used to identify the best matching work-unit.

For training the binary classifier, a training data set is gathered across work-units from the historic collection of timecards. Each training data sample consists of a triplet of the form: context-narrative-label. The binary classification system is trained to predict a binary label (1 or 0) depending on whether the narrative matches the context or not.

At inference time the work description provided by the timekeeper is tested against all work-units that timekeeper is working on, and the most probable ones are selected. From the set of all relevant context-description pairs (as many as there are work-units available for the timekeeper under consideration), the model estimates the conditional probabilities of the work-units given the provided description and selects the top candidates, which are then presented to the user.

We have evaluated the client-matter-phase prediction model performance over time against real data, observing a clear need for model updates on a periodic basis to avoid performance degradation. With weekly model updates, the top three work-units selected by the model consistently provided an accuracy of 98% and over.

4.2 Work Effort Estimation

The work description provided by the timekeeper contains useful information for estimating the work effort. The intimebot work effort estimation model is based on the assumption that (1) similar work requires similar effort, and (2) both the length of narratives and the number of worked hours are strongly correlated. Given a work description, the effort is estimated in two steps:

Search: an information retrieval approach is used to retrieve all timecards from the historical data collection with narratives that are similar to the work description provided and rank them by their respective similarity scores.

Inference: using the work efforts of the retrieved timecards and combining that with the similarity scores as weights, a probability distribution of the efforts is computed. This distribution is then used to estimate the minimum, maximum, and average efforts for the given work effort.

In addition to the two-step approach mentioned above, we have explored the use of linear and multi-layer perceptron-based regression methods to predict the work effort. For this, a model needs to be pre-trained on the embedded representation of the historical narratives and their work efforts. At inference time, a given work description is transformed into an embedded vector, which is passed through the model to predict the effort. In the future, this approach can be combined into the intimebot system for effort estimation.

4.3 Narrative Correction and Normalization

In our previous studies we have determined that timecard narrative diversity can be reduced to about 120 basic patterns. These patterns are typically comprised of specific combinations of connectors (functional units from a fixed set of words) and constituents (semantic units with one or more components) that refer to specific entities, properties and/or conditions.

Our current approach to narrative correction is a two-step process. First, we identify the patterns in the input narrative via a combination of rules and vector search. We then use detected patterns and canonical forms of components to rewrite the narrative according to specified standards using LLMs. For this, we use custom prompts that are specific to the identified patterns.

Additional rules, including both grammar and business rules, as well as proprietary formatting and stylistic guidelines, are incorporated as a post-processing step. LLMs and rule-specific prompts are used in this narrative post-edition step.

Examples of grammar rules include applying capitalization and proper usage of punctuation marks. Some of the business rule examples include enforcing verbs to be in past tense, the use of canonical forms for company names and proper formatting for person names.

4.4 Compliance Validation

The compliance validation module uses a hybrid system composed of rules, vector search and LLM-based classifiers to identify potential compliance issues within the generated timecard.

The client and matter information associated to a generated timecard allows for identifying the corresponding terms of business (such as the terms contained in outside counsel guidelines, billing terms, engagement letters, etc.), which should be already indexed and available in consumable form. Some examples of such terms are, for instance, block-billing not being permitted, the definition of specific roles within the organization being able to perform certain tasks, interns not being assigned to research tasks, etc.

Our approach regarding compliance is twofold. We use a rule-based approach to label narratives with common compliance issues. This enables us to rapidly flag violations like certain titles charging time to unpermitted activities. For more complex policy violations or issues specifically tailored to a certain client, we use semantic similarity search with a subsequent LLM-based validation of the potential policy breach.

The more nuanced and client-specific policy breaches are detected by performing semantic similarity search on the narrative against the terms of business repository and asking an LLM-based classifier to validate whether the retrieved potential breaches are actual violations or false alarms.

The vector search can be performed separately for each client and indexed vector database entries can be expanded to use additional metadata, such as combinations of matter, phase, title and work in addition to narratives whenever higher levels of granularity for compliance policies are required.

The hybrid two-tier approach described here provides intimebot with the flexibility to handle common compliance cases rapidly while being adaptable to client-specific needs.

5 User Experience

The user experience of intimebot is designed to be an interactive framework in which the user and the system are able to collaborate, building over time the needed data resources for improving prediction performance, while improving the efficiency and overall experience of the timekeeping process for the user.

The current intimebot user experience is divided in four stages: reporting, selection, validation and submission. All these four stages, which are described next, are illustrated in Figure 2.

Reporting: in this step, the timekeeper enters the brief description of the work conducted. Two input modalities (text and speech) are available.

Selection: after the input is provided, the client-matter-phase prediction model will select the top work-units matching the provided descriptions, for which the timekeeper is required to manually select the correct one.

Validation: after work-unit selection, the effort estimation, narrative correction and normalization, and compliance validation models are used to generate the proposed timecard. At this stage, the timekeeper can revise and edit the timecard.

Submission: after validating the timecard, the timekeeper can submit it, for which the system will provide a submission confirmation message.

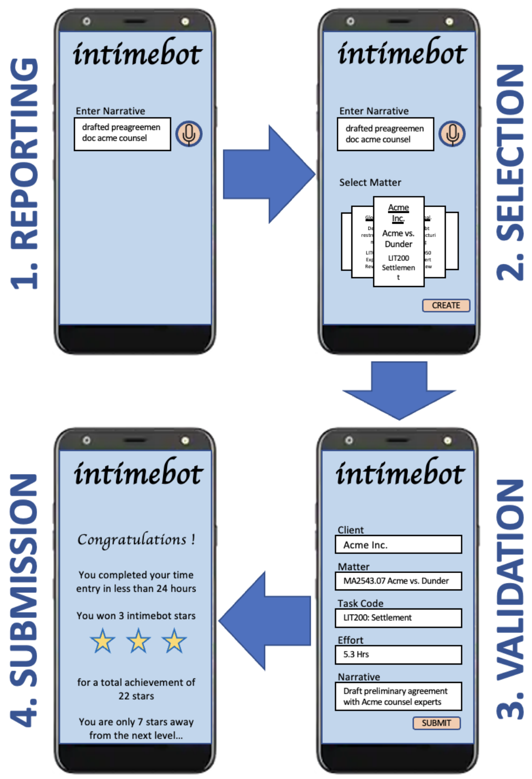

Figure 2: intimebot user experience.

Additionally, the intimebot platform taps on our existent smart-memo feature, integrated via text and/or voice input. Timer functionalities can also be used to prompt the timekeepers for descriptions after, or even before, the timer is activated.

Finally, the framework allows for the implementation and evaluation of gamification ideas to encourage timekeepers to report their time entries as soon as they complete their work.

6 Future Work

The presented intimebot demo system constitutes an experimental framework for showcasing and testing AI and ML capabilities in the timekeeping space. As part of this experimental framework, there are a few novel strategies and features we plan to test. These include:

- Replacing the current work effort estimation model by a better-informed learning-to-rank mechanism able to use adjusted hours inputted by the user to refine the ranking mechanism and similarity metric.
- We have a time capture functionality that collects detailed information on a good proportion of user activities, we can leverage on captured data to improve some of the current models performances.
- There is evidence of code-switching in spoken inputs provided via our smart memo feature (i.e. main description in local language and named entities such as companies and matter names provided in English).
- We need to better understand the value of rewards and explore novel gamification strategies by conducting user studies and other exploratory analyses.
- We plan to develop an evaluation framework for measuring the actual impact of intimebot in timekeeping activities.

Acknowledgments

The authors would like to thank Intapp, and specially the AI & Data unit, for their continuous support and encouragement for experimental and exploratory work like this one.

References

Bill4Time. 2023. A guide to lawyer time tracking. *The National Law Review*.

Nicole Black. 2020. Stay on track with these time-tracking software tools for law firms. *ABA Journal*.

Tom Boster and Scott Brenan. 2024. Why poor time-keeping could be costing your firm (and how to fix it). *Legal Management Magazine*.

Capterra. 2024. Legal timekeeping software.

Justin Henry. 2023. 82% of am law 200 firms enforce weekly or sooner timekeeper deadlines, survey finds. *LAW.com*.

Ayushi Trivedi. 2025. Agentic frameworks for generative ai applications. *Analytics Vidhya*.

Wikipedia. 2024. Comparison of time-tracking software.

Katie Wolf. 2024. Time tracking for lawyers: Methods, tools, and strategies for accurate billing. *Filevine.com online Blog*.