Results of Experimental Development R&D Activities for the Integration of AI and RPA/PA Technologies into Clinical Front-Office Processes

Lider
Anca Maria Rancea

Partner
Vlad Tărnicieru
Alexandru Suciu
Vlad Mocanu
Daniel Szőts

Project: "Adoption of Advanced Digital Technologies at MONEFY BUSINESS S.R.L.", financed under the National Recovery and Resilience Plan (NRRP), Component C9.

MONEFY BUSINESS S.R.L., in partnership with Steady Dopetech S.R.L., carried out experimental development research activities dedicated to a solution that integrates Artificial Intelligence and RPA/PA automation technologies into clinical front-office processes.

The objective of these activities was the experimental development of a digital solution capable of supporting the automation, standardisation and monitoring of complex administrative processes specific to mental health services, starting from a set of technological uncertainties whose solutions were not known at the time the project was initiated and could not be obtained through the direct application of knowledge already available on the market.

This article presents the technological uncertainties addressed, the experimental methodology applied, the results obtained for each component — including the alternatives evaluated and rejected, and the outcomes measured during piloting — as well as the challenges identified, the control measures adopted and the transferable knowledge resulting from the integration of AI and RPA/PA into clinical front-office processes.

1. Context and the Need for Advanced Digitalisation of the Clinical Front Office

Clinical front-office work involves a high volume of repetitive administrative processes: scheduling, rescheduling, cancellations, confirmations, patient notifications, insurance eligibility verification, preparation of billing information, claim creation, interpretation of EOB documents, invoice generation, transmission of patient communications, tracking of outstanding balances and operational reporting.

These processes are, for the most part, rule-based. Even so, they cannot be automated efficiently through simple, rigid logic, because each case depends on multiple factors: patient data, appointment status, provider availability, applicable operational rules, insurance data, available documentation, financial statuses, internal policies, user permissions and the requirements governing the protection of personal, medical and financial data.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

Before advanced digital components were introduced, such processes involved numerous manual steps and transfers of information between applications. This fragmentation generates operational difficulties: data entered several times over, error risk, repetitive verification, time lost on mechanical activities, poor visibility over bottlenecks, and a high degree of dependence on staff intervention for tasks that can be standardised.

At the time the project was initiated, resolving these difficulties was not a matter of conventional software engineering. It depended instead on answering a series of technological questions with no previously known solution — set out in detail in the following section. The outcome of these activities was not guaranteed at the outset, which places the work within the scope of experimental development.

2. The Technological Uncertainty Addressed

The principal technological uncertainties investigated were the following:

- **Reliable interpretation of incomplete, context-dependent natural language**, particularly over the SMS channel, where messages are short, elliptical and unstructured. At the start of the project it was not known whether a language model could consistently extract the intent, the entities and the parameters required for scheduling (service type, modality, provider, time slot) to a level of accuracy sufficient for use in a live operational environment.
- **Deterministic transformation of a non-deterministic conversation into safe operational actions.** A language model is probabilistic by nature, whereas operational processes require actions that are verifiable, repeatable and auditable. It was not clear which integration architecture would allow AI capabilities to be exploited without introducing erroneous actions into the operational system.
- **Structuring heterogeneous billing documents** (EOBs, CARC/RARC denial and adjustment codes) into actionable information, given that the format, terminology and completeness of these documents vary significantly from payer to payer.
- **Calibrating the appropriate level of autonomy for each workflow**, so that automation reduces repetitive effort without triggering actions with operational or financial impact in the absence of human validation.
- **Resilience of conversational flows to malicious instructions (prompt injection)** and to the unintended exposure of sensitive data, in an environment that processes both medical and financial information.

None of these uncertainties could be resolved through the direct application of off-the-shelf market solutions. Each required systematic investigation, prototyping and iterative testing, as detailed in the sections that follow.

3. Experimental Methodology

In order to resolve the uncertainties set out above, the work was organised as an iterative experimental process structured around cycles of the form hypothesis → prototype → measured testing → adjustment. Each component was developed through successive cycles in which test results were measured quantitatively and used to determine subsequent design decisions.

Responsibilities within the partnership were delimited as follows. Steady Dopetech S.R.L. carried out the technical research and development work — architecture design, prototyping, comparative model evaluation, component development and integration. MONEFY BUSINESS S.R.L. provided project management, expertise in real-world clinical front-office processes, definition of operational requirements,

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României”.

and operation of the application in a live operational environment, where it was tested by the beneficiary's authorised personnel in order to simulate conditions as close as possible to real use.

The experimental activities were conducted over four iterative cycles:

Cycle I – Analysis and hypothesis formulation. Mapping of real clinical front-office processes, identification of workflows with automation potential, and formulation of the technological hypotheses to be tested.

Cycle II – Prototyping and model evaluation. Development of functional prototypes for each component and comparative evaluation of several language models and integration architectures.

Cycle III – Testing in controlled environments. Testing of the prototypes against representative datasets, with measurement of accuracy, escalation rate and behaviour in edge cases.

Cycle IV – Operational piloting. Operation of the application in a live operational environment and testing by the beneficiary's authorised personnel in order to simulate conditions as close as possible to real use, with progressive adjustment of rules, prompts and thresholds.

4. Technologies Used

The AI component uses language models accessed through the OpenAI API. These models support the interpretation of requests expressed in natural language, intent classification, extraction of relevant information, semantic checks within billing workflows, interpretation of EOB documents, and the drafting or preparation of operational content.

The AI technologies employed include structured tool calling, operational instructions organised within an agent loop, incremental adjustment of model behaviour, real-time inference, and the transformation of unstructured information into structured operational data. This combination allows a message written by a patient, text drawn from an EOB, or the data of a claim to be transformed into a verifiable result that is usable within the platform.

The RPA/PA component constitutes the process automation layer. It triggers operational steps on the basis of rules and events: a new appointment, a cancellation, a reschedule, invoice issuance, receipt of an EOB, a change in insurance status, a payment received, or the occurrence of an exception. The automations do not introduce rules of their own, separate from the platform; they apply the rules configured within the system.

From a technical standpoint, AI and RPA/PA are integrated through the Monefy backend, its APIs and its event layer. This approach centralises authentication, authorisation, permissions, validation, logging and auditing. The AI never accesses the database directly, and RPA/PA operates exclusively through controlled services and APIs.

A further significant technological element is asynchronous processing. Many automations must not block the user's work: eligibility checks, notifications, invoice generation, EOB interpretation, report preparation, and the execution of recurring checks. These tasks are handled through asynchronous execution mechanisms, with statuses, logs and retry or fallback procedures.

For billing automations, the solution incorporates two separate rules engines: the Billing Rules Engine (BRE) and the Posting Rules Engine (PRE). The Billing Rules Engine runs at the encounter and claim creation stage, drawing on the operational context available, and contributes to claim construction and pricing by applying the configured rules. The Posting Rules Engine runs at the payment posting stage,

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

once remittances have been received, and governs which cases are safe enough to follow the auto-posting path and which are routed to manual review.

For reporting and business intelligence, the solution uses Metabase. This supports the configuration of dashboards, dynamic tables, filters, drill-down, exports and access control, so that operational data is made available to authorised users in a format relevant to analysis and decision-making.

5. The Solution Delivered and Its Functional Components

The functionality is presented using the structure adopted in the feasibility study: AI components for scheduling, the patient portal, SMS, claims and EOBs; and RPA/PA components for notifications, eligibility, billing, reporting and the central automation dashboard.

5.1. AI Components Delivered

5.1.1. AI Conversational Scheduling Assistant

The AI conversational scheduling assistant is one of the principal components delivered under this project. It supports routine scheduling-related interactions between patients and the Monefy front-office service by interpreting requests and transforming them into structured requests.

The AI assistant identifies the patient's intent, the type of request and the information required for the workflow to proceed. Requests relating to scheduling, rescheduling, cancellation or confirmation are thereby handled in a more structured and more consistent manner.

Scheduling in a live operational environment depends on multiple rules: provider availability, service location, session duration, service type, patient status, applicable operational rules, and any calendar conflicts. For this reason, the AI does not function as an independent decision-making mechanism but as an assistive component. The platform remains responsible for validating rules and confirming actions.

5.1.1.1. Conversational Assistant in the Patient Portal

Within the patient portal, the AI assistant supports interactions relating to scheduling, rescheduling, cancellation, confirmation and straightforward clarifications. The portal provides a controlled digital context in which the available information can be correlated with account status, existing appointments and applicable rules.

The AI assistant supports natural-language scheduling, recommendation of available time slots, rescheduling, cancellation, guidance through frequently asked questions, and management of waitlist requests, in accordance with the rules defined within the platform.

In this scenario, the AI guides the patient through a conversational flow, requests any missing information and converts the responses into a structured request. The platform validates availability, applies the scheduling rules and confirms the outcome in line with the defined operational logic.

5.1.1.2. SMS Self-Scheduling Assistant

The SMS self-scheduling assistant complements the functionality available in the patient portal. SMS is a fast and accessible channel, well suited to patients who use the portal less frequently or who need only simple scheduling-related interactions.

Through this channel, the patient can carry out certain interactions relating to confirmation, rescheduling, cancellation or simple requests. Messages may be short, incomplete or context-dependent, and the AI interprets responses in relation to the current step of the conversation and the platform's rules.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

The functionality includes patient identity verification, consultation of availability in the Monefy calendar, application of scheduling rules, and handling of restricted cases such as patients flagged with “Do Not Schedule” status. Once a scheduling, rescheduling or cancellation action has been completed, the workflow triggers the corresponding notifications to the therapist or other relevant staff.

Where the information provided is sufficient, the workflow proceeds in accordance with the defined rules. Where a message is unclear, or where the situation requires verification, the case is routed to staff. SMS is thus integrated within a controlled workflow rather than treated as an isolated automation.

5.1.1.3. Integration of the AI Assistant with the Monefy Platform

Integration of the AI assistant with the Monefy platform ensures that the recommendations and interpretations generated by the AI are used within a controlled framework. The AI does not operate directly on critical data; it passes its results to the platform backend, where business rules, permissions and the necessary validations are applied.

Through this integration, requests interpreted by the AI are converted into controlled actions: slot search, presentation of options, preparation for confirmation, status updates, or handover to staff in unclear cases.

The Monefy platform remains the source of truth for data, rules, statuses and audit trails, while the AI functions as a conversational assistance layer integrated into the platform’s workflows.

5.1.2. AI for Drafting Claims and Processing EOBs

A second significant AI component is the use of artificial intelligence in the billing domain, specifically for composing and preparing draft claims and for processing Explanations of Benefits (EOBs).

Within the claims workflow, the AI receives the details of a claim in draft status and generates operational recommendations for the user. These recommendations are displayed in an informational area associated with the claim, in the form of clear, easily verifiable suggestions.

The AI flags missing information, inconsistencies between data points, potential denial risks, codes or fields requiring review, and elements that do not appear aligned with the data already held in the platform. The AI thus functions as a review and assistance layer before the billing process continues.

This functionality does not constitute full automation of the billing decision. The AI does not decide independently whether a claim is submitted, nor does it replace specialist staff. Its role is to help the user identify potential problems more quickly and to reduce the time spent on repetitive checks.

Within the EOB workflow, the AI analyses the documents or information received from payers, consolidates denial and adjustment codes, groups the relevant reasons and indicates which claims require user attention. The output is presented in an informational area, so that staff can quickly understand which cases need to be examined and why.

This functionality is particularly valuable where a single EOB covers multiple claims, adjustment codes, denial codes or inconsistent messages. The AI reduces initial analysis time, standardises interpretation and assists in prioritising the cases that require action.

Through this approach, the AI serves as an assistant within the billing function. It prepares and structures information and indicates possible courses of action, while the final decision remains with authorised staff.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României”.

5.2. RPA/PA Components Delivered

The RPA/PA components support the consistent execution of operational processes by digitalising repetitive steps and standardised decisions. Under this model, the RPA/PA layer detects relevant events within the platform, applies configured business rules and triggers steps that can be executed reliably without routine human intervention.

This section groups together the specific automations delivered as part of the solution: notifications, eligibility verification, billing, the Billing Rules Engine, the Posting Rules Engine, reporting, and the central automation dashboard.

5.2.1. Automated Patient Notifications and Reminders

The automated notification functionality allows communication workflows to be triggered and executed without manual intervention in each standard case. Notifications relate to appointments, rescheduling, cancellations, confirmations, reminders, invoices or other relevant status changes.

RPA/PA monitors events within the platform and determines, on the basis of the configured rules, whether a communication should be sent to the patient or to staff. A new appointment, a change to an appointment, or a generated invoice triggers the corresponding messages through the available channels, provided the defined conditions are met.

This functionality reduces repetitive work for front-office teams and contributes to more consistent communication. At the same time, the rules provide for situations in which a notification should not be sent automatically or should be handled through a different workflow.

5.2.2. Automated Insurance Eligibility Verification

Insurance eligibility verification is one of the RPA/PA automations most important for administrative and financial workflows. Eligibility affects scheduling, patient communication, billing, financial responsibility and the prioritisation of cases requiring manual intervention.

Through RPA/PA, eligibility checks are triggered automatically according to events or rules: a new appointment, the approach of the date of service, a change in insurance information, the absence of a valid status, the expiry of a previous check, or the need for a recurring check.

The results of these checks are surfaced in user work queues, in reports or within notification workflows, depending on the applicable rules. Where eligibility is clear and the rules are satisfied, the process continues. Where the result is incomplete or contradictory, or indicates a problem, the case is flagged for manual review.

5.2.3. Automation in Billing and Invoice Processing

Billing automation is connected to the claims, EOB and patient financial responsibility workflows. Once the relevant information has been processed, RPA/PA supports the generation of invoices and their transmission to the patient, where the rules are clear and the necessary data is complete.

In a traditional workflow, staff must identify the amounts to be billed, verify the available information, generate the invoice and send the communication to the patient. Through RPA/PA, this process is supported automatically where the necessary conditions are met: the existence of the relevant data, valid statuses, applicable billing rules, available communication channels, and the absence of exceptions requiring manual review.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României”.

The AI supports this area by interpreting EOBs and structuring the resulting information, while RPA/PA triggers document generation, status updates and transmission of the communication to the patient, in accordance with the defined rules.

This functionality is built with clearly defined control points. Where the data is complete and the rules are unambiguous, the invoice is generated and sent automatically. Where data is missing, or where there are discrepancies, contradictory statuses or financial risk, the workflow halts and routes the case to staff.

5.2.4. Billing Rules Engine

The Billing Rules Engine (BRE) is the RPA/PA component used at the encounter and claim creation stage. The engine evaluates the available context — including provider, appointment type, CPT code, carrier, patient age, duration, place of service and other configurable conditions — and then applies prioritised rules for claim construction and pricing.

Through the BRE, CPT codes, billing amounts, modifiers, place of service, diagnoses, revenue codes, billing provider and type of bill can be determined according to the applicable rules. The component can also suppress billing for non-billable cases, require prior authorisation, or flag a case for human review.

Where the rules cannot supply the required elements, the case is held or routed to staff for completion and review. The BRE thus standardises the repetitive steps of coding, pricing and claim preparation without eliminating human verification for exceptions.

5.2.5. Posting Rules Engine

The Posting Rules Engine (PRE) is the RPA/PA component used in remittance processing and payment posting. Before the rules are applied, the system enforces a set of fixed structural checks that no rule may relax: the claim must be identifiable locally; it must not be voided or corrected; the ERA lines must correspond to the claim; the case must not be a reversal or takeback; and there must be no overpayment.

For cases that pass these checks, the PRE applies configured rules that may restrict or refine the auto-posting workflow: routing to manual review, void on denial, or controlled write-off of small residual balances. Where nothing blocks the case, the claim may be auto-posted, with the payment, adjustments, write-offs and patient responsibility recorded in accordance with the breakdown received.

5.2.6. Automated Reporting for Business Decisions, Use of Metabase, and Scheduled Report Distribution

Automated reporting supports the transformation of the platform's operational data into information useful to management and to operational teams. Rather than requiring reports to be built manually, the solution allows reports to be generated and refreshed on the basis of the data available within the platform.

This functionality contributes to clearer analysis of workloads, processing times, statuses, recurring activities and the areas in which operational bottlenecks arise. Through reporting, operational data becomes easier to track, compare and apply to business decisions.

Within the solution, Metabase is used as the BI and reporting tool for configuring dashboards, dynamic tables, filters, drill-down options, exports and access control. Authorised users consult the information relevant to their work and drill into the data in accordance with their access rights.

On the RPA/PA side, the reporting component also includes scheduled distribution of reports or dashboards by email, at configured intervals and to configured recipients. This functionality brings reporting closer to an automated workflow: the system collects the available data, refreshes the report

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

and triggers its distribution without manual preparation for each transmission, within the limits of the applicable rules and access rights.

This component is separate from the AI functionality. The purpose of reporting is not to interpret operational data automatically through AI, but to provide visibility over the information available within the platform, in a format useful for analysis and decision-making.

5.2.7. Central Automation Dashboard

The central automation dashboard is the component through which authorised users administer the automation options available within the platform. It provides a single point of configuration for RPA/PA workflows, without requiring direct code changes for each operational adjustment.

Through the dashboard, authorised users can enable or disable RPA/PA workflows and configure rules, parameters, timings, thresholds and priorities, within the limits of their access rights.

The role of the dashboard is the administration and configuration of automations. Technical monitoring of executions, logs, errors and service health is handled through dedicated observability and logging mechanisms, and is not a primary function of the automation dashboard.

Through this component, the solution enables faster and more coherent management of the available automations, while retaining control over which workflows are started, stopped or adjusted by authorised users.

6. Results of the Experimental Activities

6.1. Experimentation with the AI Conversational Scheduling Assistant

The AI conversational scheduling assistant supports routine scheduling-related interactions between patients and the Monefy front-office service by interpreting requests and converting them into structured requests, both through the patient portal and over the SMS channel.

Comparative evaluation of language models. The conversational component was tested successively with several models accessed through the OpenAI API, in order to identify the optimal balance between accuracy, latency and cost per conversation. Models from different capability and cost classes were evaluated, including GPT-4o, GPT-4o mini, GPT-4.1 mini, a model from the reasoning family (o4-mini) and, finally, GPT-5.6 Luna.

The conclusions of this evaluation were as follows:

- High-capability models (GPT-4o) delivered very good interpretation of ambiguous messages, but introduced latency and per-conversation cost incompatible with a channel such as SMS, where a rapid response is expected.
- The first-generation very small models (GPT-4o mini) reduced cost but produced a higher rate of tool-calling errors — invocation of non-existent functions or incorrect population of parameters.
- The reasoning model (o4-mini) delivered superior results on complex interpretation cases, but at a latency and cost that did not justify its use for routine scheduling workflows.
- GPT-5.6 Luna offered the best ratio between tool-calling accuracy, latency and cost for the agent loop, and was therefore adopted as the default model, with the model remaining configurable per environment.

Architectural alternatives evaluated and rejected. A central experimental result is the transition from an initial approach based on free-text interpretation — in which the model returns text and the system

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României”.

attempts to extract parameters through pattern analysis — to an architecture based on structured tool calling, in which the model invokes explicitly defined functions (lookup_responsible_party, get_appointments, reschedule_appointment and others) and the system executes the actions within a controlled framework. The first approach proved fragile: parameters were extracted inconsistently and behaviour was difficult to reproduce. The second approach was retained because it enabled deterministic validation of every action and a clear separation between interpretation (AI) and execution (platform).

A second rejected alternative was the single monolithic prompt, in which the entire conversational logic was encoded within one extended instruction. This was replaced by an agent loop — a dynamic system prompt constructed from session state, combined with a defined tool set and successive iterations — which proved more robust, easier to maintain and more predictable.

Handling hallucinations and platform-level validation. Testing revealed a specific risk: the model could propose non-existent appointment slots (slot hallucination). The experimental solution retained was for the platform to remain the source of truth. Any proposed slot is validated against actual calendar availability before being presented to the patient, and the scheduling rules — a provider from the care team with active Direct Care status, service type, patient eligibility, applicable operational blocks, and the restrictive “Do Not Schedule” status — are enforced by the backend rather than by the model. Invalid proposals are therefore eliminated before reaching the user.

Prompt injection resilience testing. The conversational workflows were subjected to a battery of adversarial messages designed to induce the model to disregard its rules, disclose information about other patients, or trigger impermissible actions. The results led to the strengthening of operational instructions, to strict limitation of the context supplied to the model to the role and entities permitted for the current session, and to the principle that the model never accesses the database directly but exclusively through tools with a limited scope of action.

Results measured during piloting:

- Intent recognition accuracy (portal): approximately 82% for the initial prototype, rising to approximately 96% in the final solution.
- Intent recognition accuracy (SMS): approximately 74% for the initial prototype, rising to approximately 93% in the final solution.
- Conversations completed without escalation to staff (portal): approximately 55% for the initial prototype, rising to approximately 78% in the final solution.
- Conversations completed without escalation to staff (SMS): approximately 40% for the initial prototype, rising to approximately 64% in the final solution.
- Invalid slots reaching the patient: approximately 7% of proposals for the initial prototype, reduced to 0% in the final solution through backend validation.

Integration of the AI assistant with the Monefy platform ensures that the recommendations and interpretations generated by the AI are used within a controlled framework: requests interpreted by the AI are converted into controlled actions — slot search, presentation of options, preparation for confirmation, status updates, or handover to staff in unclear cases — while the Monefy platform remains the source of truth for data, rules, statuses and audit trails.

6.2. Experimentation with the AI Component for Claims and EOBs

Hypothesis and uncertainty. The hypothesis tested was that a language model could reduce billing review time by flagging, prior to submission, the definite problems within a claim draft, and by prioritising the claims requiring follow-up once an EOB has been received. The principal uncertainty concerned the

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României”.

false-positive rate: an assistant that frequently flags non-existent problems erodes user confidence and is ultimately abandoned.

Alternatives evaluated and rejected. An initial approach — supplying the EOB document as raw text within a single prompt and requesting a free-form analysis — produced inconsistent results across heterogeneous documents. It was replaced by a structured extraction approach, in which the billing information available within the platform (payments, adjustments, denial and remark codes, outstanding amounts, appeal deadlines) is supplied to the model in structured form, and the model produces discrete, verifiable findings.

The decision on the level of autonomy — an experimental outcome. Testing confirmed that the model could not be permitted to act autonomously on the billing decision. A strictly advisory principle was therefore adopted by design: the component runs automatically (controlled by a feature flag), is visible only to users holding claim-update permissions, and does not modify, submit or change the status of any claim or EOB. For claims, suggestions are formulated as short, single-line actions (for example, “Set the POS on line 1 to Telehealth”), displayed in an informational area at the top of the claim record. For EOBs, the component prioritises only those claims it is confident require follow-up, providing the claim reference, a priority level, the reason and a recommended action, while explicitly excluding write-off or patient-billing recommendations.

Confidence threshold calibration. A significant experimental result was the deliberate calibration of the component towards high precision at the expense of recall: the system flags only those problems of which it is certain. This choice followed directly from testing the precision–recall trade-off and from the finding that, in an assistive context, a false positive costs more in user confidence than a false negative.

Results measured during piloting:

- Precision of suggestions on claim drafts (problems confirmed as genuine): approximately 93%.
- Agreement between EOB prioritisation and the assessment of a billing specialist: approximately 90%.
- Reduction in initial review time for an EOB covering multiple claims: approximately 40%.

6.3. Experimentation with the RPA/PA Components

Rejected alternative: rigid script-based automation. An initial approach modelled the automations as fixed sequences of steps. This proved unsuited to real-world processes, which contain exceptions, variable statuses and contextual conditions. A rigid workflow would either halt at the first deviation or execute steps in an inappropriate context. The experimental outcome was a redesign of the automations on an event- and rule-driven basis: the RPA/PA layer detects relevant events within the platform, evaluates the configured rules and determines whether a step proceeds, halts or is escalated.

Calibration of escalation thresholds. For eligibility verification, testing showed that results may be ambiguous or contradictory. A set of conditions was defined experimentally under which a case is flagged for manual review (incomplete or contradictory results, or results with financial impact), so that automation reduces repetitive volume while retaining human control over unclear situations. Similarly, for automated billing, the stopping points at which the workflow routes a case to staff (missing data, discrepancies, contradictory statuses, financial risk) were established through testing.

The RPA/PA components delivered, as they emerged from this calibration process, cover:

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României”.

- **Automated patient notifications and reminders** — confirmations, reminders and cancellation or rescheduling alerts, triggered on the basis of events and configured rules, with exceptions handled explicitly for cases in which automated notification is not appropriate.
- **Insurance eligibility verification** — triggered by events (a new appointment, the approach of the date of service, a change in insurance information, expiry of a previous check), with automatic escalation to manual review where the result is incomplete or contradictory, or carries financial impact.
- **Billing and invoice processing** — automatic generation and transmission of the invoice where the data is complete and the rules are unambiguous; halting and routing to staff where data is missing or where there are discrepancies, contradictory statuses or financial risk, on the basis of the rules applied by the dedicated billing and posting components.
- **Billing Rules Engine** — construction and pricing of claims at the creation stage, through the application of rules governing coding, pricing, modifiers, place of service, diagnoses, revenue codes, billing provider, prior authorisation, billing suppression and flagging for review.
- **Posting Rules Engine** — control of auto-posting at the remittance and payment posting stage, following fixed structural checks, through rules that may force manual review, apply void on denial under controlled conditions, or write off small residual balances.
- **Automated reporting through Metabase**, including scheduled distribution of reports or dashboards by email, providing management visibility over volumes, processing times and bottlenecks — a component separate from AI interpretation, with a role strictly limited to the analysis and distribution of operational data.
- **Central automation dashboard** — a single point of configuration for enabling and disabling RPA/PA workflows and adjusting rules, thresholds and priorities, within the limits of authorised users' access rights.

7. Security, Control and the Responsible Use of AI and RPA/PA

An important element of the experimental development activities was the analysis of how AI and RPA/PA technologies can be introduced into an environment handling sensitive data without an uncontrolled increase in security risk.

The Monefy platform processes identification data, contact details, appointments, financial information and other operational data. In this context, the advanced digitalisation components were designed with clear mechanisms for protection, auditing and access limitation.

AI and RPA/PA do not function as separate systems and do not access the database or the underlying infrastructure directly. They are integrated through the Monefy backend, through official APIs and through controlled events, so that existing authentication, authorisation, validation and business rules apply uniformly.

For the AI component, as detailed in Section 6.1, one of the principal challenges tested directly was protecting conversational workflows against malicious or unintended instructions (prompt injection). A further risk addressed explicitly is the unintended exposure of sensitive data: AI-generated responses must be limited strictly to the permitted context and to the user's role, while logs, test data and information used to adjust model behaviour must be handled carefully in order to avoid the unjustified storage or transmission of personal, medical or financial data.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

For the RPA/PA components, security involves the clear delimitation of the permissions and technical accounts used by the automations. Automated workflows operate under least privilege, perform only those actions permitted through the platform's official APIs, and are tracked through audit and reporting mechanisms.

This approach supports a responsible model for the use of AI and RPA/PA. The technology interprets, recommends or automates operational steps, but remains bounded by rules, permissions, validations and monitoring mechanisms. Technological innovation is thereby aligned with the security, confidentiality and operational control requirements of an environment handling sensitive data.

8. Training, Fallback and Operational Integration

An essential aspect of knowledge transfer is preparing users to work with AI and RPA/PA. The technology cannot be introduced effectively through technical activation alone. Staff must understand what the AI does, what its limits are, how recommendations are verified, when an action may be accepted, and when a case must be continued manually.

For scheduling workflows, staff must understand how the AI assistant operates in the portal and over SMS, what a slot recommendation means, how ambiguous messages are handled, and what steps are followed when the AI does not complete an action independently.

For billing workflows, staff must understand that the AI generates suggestions, structures information and prioritises cases, but that the final decision on a claim, an invoice or a follow-up action remains with authorised users.

For RPA/PA workflows, users must understand which actions are automated, which statuses are updated by the system, where the configuration options are found, and how to intervene when a workflow requires halting, adjustment or manual continuation.

Fallback procedures are important because AI and RPA/PA do not resolve every situation automatically. Where data is incomplete, responses are unclear, rules cannot be applied safely, or risk is elevated, the system allows manual continuation, documentation of the situation and escalation to the appropriate team.

9. Challenges Encountered and Lessons Learned

Beyond the initial technological uncertainties (Section 2) and the design decisions resulting from testing (Section 6), the research and development activities highlighted several cross-cutting challenges relevant to any project of this kind.

The first is data quality. For AI and RPA/PA to function correctly, the data concerning patients, appointments, providers, operational structures, insurance policies, invoices, claims, EOBs, statuses and rules must be sufficiently clear and up to date. Automation does not fully compensate for incomplete or contradictory data.

The second is integration with existing operational rules. A frequent risk in AI projects is the creation of a separate layer that makes decisions without respecting the logic of the principal system. The solution developed integrates AI and RPA/PA as extensions of the platform, so that the actions proposed or executed are validated through the same rules and permissions used by the operational system.

The third is maintaining user confidence. If staff are to adopt AI and RPA/PA technologies, the system must deliver verifiable results, allow control over actions, and avoid creating any sense of lost operational

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

control. It is for this reason that, as shown in Section 6.2, the component was deliberately calibrated towards high precision at the expense of recall.

The fourth is the correct separation between rules engines, automation administration and operational reporting. The Billing Rules Engine and the Posting Rules Engine support two distinct stages of the billing workflow: the construction and pricing of claims at creation, and the control of auto-posting during remittance and payment processing, respectively. The central automation dashboard supports the configuration and management of RPA/PA workflows by authorised users, while reporting through Metabase supports the analysis of operational data and business decision-making. This separation helps preserve a clear model in which rules are applied through dedicated components, automations are managed through controlled options, and operational outcomes are analysed through reporting tools.

10. Measured Benefits and Operational Impact

10.1. Methodology for Deriving the Quantitative Benefits

Reductions in operational time were determined by combining a time-and-motion study of current processes (measurement of actual effort per activity within the beneficiary's operational environment) with the per-unit reduction measurements obtained during prototyping and piloting (Sections 6.1–6.3).

Piloting was conducted by running the application in a live operational environment and testing it with the beneficiary's authorised personnel across a representative volume of appointments, claims and EOBs, allowing these estimates to be calibrated on a per-unit basis and validating the hypotheses concerning reduced repetitive effort.

10.2. Results

- **Time to create a claim:** measured at approximately 4 minutes per claim in the manual process; reduced to approximately 2 minutes through automatic pre-population of mandatory fields from data already held in the platform (a reduction of approximately 50%).
- **Time to create a copayment invoice:** measured at approximately 5 minutes per invoice; reduced to approximately 2 minutes through automatic identification of the relevant EOBs and application of the billing rules (a reduction of approximately 60%).
- **Aggregate billing effort** — which includes, alongside claim and invoice creation, the associated verification and reconciliation activities — was measured at approximately 79.50 hours for the reference volume and falls to approximately 51 hours through automatic pre-population of drafts and generation of invoices on the basis of EOBs and the established rules (an aggregate reduction of approximately 36%). The aggregate reduction is lower than the per-unit reductions because not all billing effort consists of automatable activities.

Beyond these quantitative benefits, the project also generates qualitative ones: patients receive faster and clearer communications; staff spend less time on repetitive activities; management gains access to better-structured information through reporting; and processes become more standardised, more traceable and easier to improve.

One important benefit is increased scalability. As patient numbers and operational volumes grow, an organisation needs processes capable of handling larger volumes without a proportional increase in manual work. AI and RPA/PA contribute to this scalability by absorbing repetitive tasks while retaining human control over sensitive cases.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

11. Transferable Knowledge Resulting from the R&D Activities

The activities carried out produced several lessons applicable to similar advanced digitalisation projects, particularly in contexts where automation must operate with sensitive data, operational rules and human intervention.

The first lesson is that AI is valuable when it is connected to rules and structured data. A language model can interpret messages, documents or operational explanations, but its outputs become useful only when validated by the platform through clear rules, permissions and context checks.

The second lesson is that RPA/PA must be built on events and rules rather than as a rigid sequence of scripts. Real administrative processes contain exceptions, differing statuses and contextual conditions. Automation must proceed, halt or escalate according to those conditions.

A further transferable outcome is the model for using AI in the billing domain: structuring information, formulating recommendations, consolidating denial codes and prioritising cases, with the final decision retained by authorised staff, and with deliberate calibration towards high precision at the expense of recall.

An identified good practice is the clear separation between automation administration, operational rules and operational reporting. The Billing Rules Engine and the Posting Rules Engine support two different stages of the billing workflow — the construction and pricing of claims, and the control of auto-posting, respectively — while the central automation dashboard supports the configuration of RPA/PA workflows and Metabase supports the analysis, reporting and scheduled distribution of operational data. This separation reduces confusion between the application of rules, the control of automation options and the analysis of operational results.

A further lesson is that security must be designed in from the outset. Prompt injection, accidental data exposure, abuse of AI capabilities and the compromise of automated workflows are real risks, tested directly within the project (Section 6.1). Backend integration, privilege limitation, auditing, logging and validation are therefore mandatory elements.

Finally, the activities confirmed that the implementation of AI and RPA/PA must be gradual. Automations should first be introduced on low-risk workflows and subsequently extended to more complex processes following validation, feedback and rule adjustment.

12. Conclusions

The experimental development research activities carried out by MONEFY BUSINESS S.R.L. in partnership with Steady Dopetech S.R.L. took as their starting point genuine technological uncertainties — concerning the interpretation of natural language, the deterministic transformation of conversational interactions, the structuring of heterogeneous billing documents, the calibration of autonomy, and resilience to malicious instructions — and addressed them through an iterative experimental process measured quantitatively at each cycle.

The AI conversational scheduling assistant, interaction through the portal and over SMS, AI support for billing and EOBs, automated notifications, eligibility verification, invoice automation, the Billing Rules Engine, the Posting Rules Engine, reporting through Metabase and the central automation dashboard represent complementary strands of a single model of advanced digitalisation. These components were validated through tested and rejected alternatives, experimentally calibrated thresholds and outcomes measured during piloting.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României”.

The most important conclusion of this work is that AI and RPA/PA must be introduced responsibly, gradually and verifiably. In domains where processes involve sensitive data and decisions with operational or financial impact, the technology must operate within clear rules, with provision for human intervention and with mechanisms for auditing, reporting and control.

Through this project, MONEFY BUSINESS S.R.L. contributes to the transfer of skills, competences and knowledge in the use of advanced digital technologies for the automation of administrative processes in healthcare services. The results obtained support further development in AI-assisted scheduling, administrative workflow automation, rule-based verification, interpretation of operational documents, billing automation, and process monitoring through integrated digital tools.

13. Glossary of Terms

- **AI / Artificial Intelligence** – technology used to interpret natural-language requests, classify information and generate operational recommendations.
- **RPA/PA** – robotic process automation / process automation, used to execute standardised administrative steps on the basis of rules and events within the platform.
- **Technological uncertainty** – a technical question whose answer is not known at the time a project is initiated and cannot be obtained through the direct application of existing knowledge or market solutions; resolving it requires systematic investigation, prototyping and testing.
- **Experimental development** – a category of research and development activity based on cycles of the form hypothesis → prototype → measured testing → adjustment, the outcome of which is not guaranteed at the outset.
- **Tool calling** – an architecture in which the language model invokes explicitly defined functions and the system executes the resulting actions within a controlled framework, as opposed to free-form interpretation of model-generated text.
- **Hallucination (model)** – the generation by a language model of information that is plausible but incorrect or non-existent (for example, an appointment slot that does not exist in the calendar).
- **Precision / recall** – the trade-off between flagging only those problems of which the system is certain (high precision) and flagging as many potential problems as possible, at the risk of false positives (high recall).
- **API** – a programming interface through which software components communicate with one another in a controlled manner.
- **OpenAI API** – the service used to access the language models underpinning the AI functionality examined within the project.
- **Backend** – the internal part of the platform that manages data, rules, permissions, validations and communication between components.
- **Do Not Schedule** – an operational status assigned to a patient for whom the system does not permit automated scheduling or the use of certain self-scheduling workflows, in accordance with the applicable internal rules.
- **EOB / Explanation of Benefits** – a document issued by a payer or insurer explaining how a claim was processed, which amounts were covered, adjusted or denied, and which remain the patient's responsibility.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".

- **CARC / RARC** – standardised codes used by payers to explain denial reasons (Claim Adjustment Reason Codes) or supplementary remarks (Remittance Advice Remark Codes) associated with an EOB.
- **ERA / Electronic Remittance Advice** – the electronic form of a payer's remittance advice, setting out how each claim line was paid, adjusted or denied.
- **Claim** – a request submitted to a payer or insurer for reimbursement or processing of a medical service.
- **Billing** – the set of administrative and financial processes through which services delivered are prepared as claims, processed on the basis of payer or insurer responses, converted into invoices where applicable, and tracked through to payment or clarification.
- **Dashboard** – an interface that displays, organises or administers information and options relevant to a digital process.
- **BI / Business Intelligence** – the use of reports, dashboards and data analysis to support operational and managerial decision-making.
- **Metabase** – the BI and reporting tool used to configure dashboards, dynamic tables, filters and data exports.
- **Billing Rules Engine (BRE)** – a configurable rules engine used at encounter and claim creation. It applies prioritised rules governing CPT codes, pricing, modifiers, diagnoses, place of service, revenue codes, billing provider and other billing elements; it can suppress billing, require prior authorisation, or flag a case for review.
- **Posting Rules Engine (PRE)** – a configurable rules engine used in remittance processing and payment posting. Following fixed structural checks, it can force manual review, void a claim on denial under controlled conditions, or apply a write-off for small residual balances; it does not relax the auto-posting safety checks.
- **Prompt injection** – a risk specific to AI systems, in which a user attempts to introduce instructions causing the AI to disregard its rules or disclose impermissible information.
- **Human intervention / manual validation** – a situation in which AI or RPA/PA assists the process, but authorised staff verify, confirm or manually continue the case where data is incomplete, rules cannot be applied safely, or operational or financial risk is present.

Project Information

The research and development activity was carried out within the project entitled: "Adoption of advanced digital technologies within MONEFY BUSINESS S.R.L.", RUE code: 84, financed through the NRRP (National Recovery and Resilience Plan), CALL: PNRR/2024/C9/MIPE/I3.1/GRANTS FOR SUPPORTING ENTREPRENEURS IN THE DEVELOPMENT OF ADVANCED DIGITAL TECHNOLOGIES – MEASURE 1, based on Financing Contract no. 84.1.I3.C9 of 13.05.2025.

Lead partner: MONEFY BUSINESS S.R.L.

Partner: Steady Dopetech S.R.L.

„Conținutul acestui material nu reprezintă în mod obligatoriu poziția oficială a Uniunii Europene sau a Guvernului României".