MEMOTEXT

Artificial Intelligence for Digital Engagement Playbook

PART II & III

July 2025

Part 2-3 Version 1.0. 07.04.2025

Why the AI for Digital Patient Engagement Playbook?

The introduction of AI in healthcare is far from recent. While today's advancements may feel novel and fast-moving, AI's history in healthcare dates to the 1960s, driving innovations in machine learning diagnostics, predictive analytics, and patient engagement. However, despite the overwhelming excitement, hype and attention about AI (in health) the widespread adoption of AI in healthcare remains in its early stages.

The widespread adoption and scaling of large language models (LLMs) has pushed AI and its clinical applications into mainstream conversations, accelerating curiosity, concern and some fear of being left behind. For the purposes of this *Playbook*, our focus is on AI in digital patient engagement and mobile health communication. In this context, patient engagement refers to the use of personalized, data-driven communications, mobile and digital user experiences with patients to enhance treatment adherence, improve care coordination, support self-management, and improve overall health and system level outcomes.

As AI and Machine Learning (ML) practitioners since 2016, enablers of conversational agent capabilities, predictive analytics, and secure LLM development, we understand both the excitement and risk aversion this emerging and rapidly evolving technology evokes. We at MEMOTEXT are sharing this *living* *Playbook* (1.0) document to provide some certainty and grounding in systems thinking as organizations grapple with the introduction of AI into care specifically in the realm of digital patient engagement communications. This playbook outlines an introductory guide to developing and integrating AI in ways that can maximize value, mitigate risk, and align the use of AI with patient safety and trust. We provide examples, resources and examples from MEMOTEXT's own case-studies to illustrate real world examples. From governance to deployment, we hope this resource serves as a foundation and a springboard for responsible and impactful AI-driven patient engagement.

We at MEMOTEXT are experiencing our partner's, client's and colleagues' curiosity as leaders and stakeholders in health systems of all sizes are seeking (potentially scrambling for) directionality, clarity and certainty on AI engagement and workflow integration.

We thank you for downloading the MEMOTEXT AI Engagement Playbook

To contact the authors or the MEMOTEXT team. Please contact memo@memotext.com AI Engagement Playbook Overview

Part I Summary

Part I of the AI for Digital Engagement Playbook laid the foundation for responsible, value-driven AI in healthcare engagement. The initial section focused on **defining high-impact use cases, building governance frameworks, and ensuring transparency, patient safety, and trust**. With frameworks like Design Thinking, SSM, and CRISP-DM, it emphasized aligning AI tools with real-world problems and ethical standards. Part I also introduced the distinction between prescriptive and augmentative AI solutions, reinforcing the need for human oversight and clear communication in patient-facing applications.

We are pleased to present parts II and III for your reading and sharing pleasure.

Part II

Implementation AI: The Nuts and Bolts

The development and deployment of AI for digital patient engagement is currently experiencing a dynamic phase of rapid experimentation, flux and growth. Organizations are faced with a variety of strategic choices such as adopting established AI models such as ChatGPT or Claude, leveraging open-source frameworks like Llama or Hugging Face, collaborating with specialized AI vendors, or developing entirely customized solutions.

Similarly, machine learning methodologies, analytics and data-science are now part of the everyday operational health experience. With numerous paths available, clarity in strategy and a robust understanding of each option's strengths and limitations become essential. Model selection, methodology and design/engineering will relate back to problem definition and objectives defined early in development.

Data Collection and Data Engineering
AI is only as good as the data you feed it. High-quality, data engineering methodologies for cleaning and cataloguing from various data repositories like EHRs, wearables, clinical data, claims, research, internal protocols / care pathways, patient-reported outcomes, and remote patient monitoring data are essential to the process of curating source data.

There are now several solutions and methodologies for optimization and data transformation to maximize efficiency, reliability and accuracy of data collection to reduce redundancies and convert messy data into AI consumable formats. Here are some guidelines to consider:

Ensure clean, ethical, structured, and interoperable data

AI models rely on clean data to work optimally. Setting up robust data validation pipelines will help to eliminate errors, duplicates, and inconsistencies before ingestion. Ensure patient data is both secure and ethically sourced and used throughout the AI development experience by ensuring data access, storage, encryption, and compliance (HIPAA, GDPR, PHIPA) oversight (see Data Governance)

Anonymized and deidentified

Data anonymization and de-identification methods should be employed where possible to enhance privacy protections. This process is meant to structure collected data it in a way that AI can maximize utility of the data. Often, the backbone of the AI strategy is a data collection and governance strategy.

Data governance strategy for activities such as inbound and outbound data driven patient engagement should ensure that patient data is collected, stored, and used securely and ethically, complying with regulations like HIPAA. This includes implementing clear policies on data access, encryption, and storage, defining who can access the data (clinicians, patients, researchers), and ensuring data quality through validation processes.

Data Tagging and Annotation Strategies

Beyond collection and cleaning, effective implementation of AI requires structured data tagging, also called annotation. This process assigns labels or metadata to raw datasets to categorize text, label images, or segmenting time-series signals (predictions) from actigraphy/wearables, for example to make them usable for machine learning. High-quality tagging is essential for supervised learning, personalization algorithms, and explainability.

There are three general approaches:

- **Manual Tagging:** Human annotators label data based on predefined schemas (e.g., tagging sentiment indicators in patient-reported text).
- **Semi-Automated Tagging:** Uses heuristics or weak AI to suggest tags, with human verification (e.g., flagging medication adherence patterns for review).
- **Fully Automated Tagging with Active Learning:** Models tag data based on trained classifiers, while uncertain or edge cases are escalated for human review.

Use of standard ontologies (e.g., SNOMED CT, LOINC) where applicable improves interoperability across health and information systems.

Model Training
It is critical to train your AI on a reliable and representative mass of real-world data. Data-scientists are trained to identify and account for bias and select/tune and adjust models to prevent unintentional reinforcement of bias. Thus, testing for bias is valuable at every stage. Clinical validation is an obvious requirement and ethics boards are beginning the pathway to more AI-related engagement. Regulatory bodies and IRBs are still defining best practices for AI ethics and validation in healthcare**.** Organizations need to prioritize bias mitigation, transparency, and human oversight while developing formalized AI governance frameworks.

Resource: https://www.nature.com/articles/s41746-024-01232-3

Best Practices for Maximizing AI in Engagement

The use of AI in digital patient engagement continues to evolve. The constraints and opportunities are strongly determined by the clinical/commercial environment within which patient engagement interventions are deployed. The interactivity, legal/regulatory environment, population and purpose of engagement interventions vary greatly across settings, but certain evidence-based strategies have emerged as best practices. MEMOTEXT's work across chronic disease, mental health, and high-acuity populations provides a practical foundation for applying these strategies in the real world.

AI can supercharge timing, but the logic must be clinically grounded.
Just-In-Time Adaptive Interventions (JITAIs) are designed to deliver outbound engagement at the right time, based on user context such as behavioural data, engagement patterns, and physiological or self-reported signals. In the REACH diabetes program, MEMOTEXT used JITAI logic to adjust message timing and content based on self-reported behaviour, wake-up times, and medication refill patterns. This approach, while not always AI-driven, yielded measurable improvements in medication adherence and self-efficacy among underserved populations. JITAIs create the framework that AI can optimize within, particularly for inbound and outbound personalization.

Micro-randomization is not AI, but it generates the data that makes AI smarter.
To enable ongoing optimization, micro-randomization techniques can test variations in tone, delivery time, or message type. For example, motivational versus directive messages were tested in MEMOTEXT mental health interventions for young adults, revealing which styles drove higher engagement. These granular insights can be used to continuously improve machine learning models and personalize AI messaging logic.

AI can manage multi-channel orchestration while maintaining personalization.
Effective engagement requires meeting patients on their preferred channels. AI can help orchestrate timing and sequencing across SMS, app notifications, email, and voice interfaces while preserving a consistent experience. In one deployment, MEMOTEXT combined SMS nudges with in-app surveys and post-visit email summaries to create a unified user journey. AI supports the coordination of timing, format, and content across channels to match patient preferences and contexts.

Conversational AI extends engagement beyond broadcast messaging.
Chatbots and conversational agents support two-way engagement, including outbound reminders and nudges and inbound support for navigation, literacy, and emotional check-ins. The KIT chatbot, developed with UHN and SickKids, provided tailored education and support to youth with diabetes transitioning to adult care. It used Retrieval-Augmented Generation (RAG) to surface clinically validated responses and supported mood tracking and care plan prompts. AI powered the natural language understanding and personalization while maintaining clinical control.

AI-powered conversational interfaces fall along a spectrum, from tightly controlled, rules-based chatbots to flexible, generative large language model (LLM) systems. Understanding the distinction helps teams choose the right level of complexity, safety, and flexibility for their use case.

- **Rules-Based / Intent Recognition Chatbots:** These rely on predefined scripts and decision trees. They are ideal for structured tasks like appointment scheduling, FAQs, or medication reminders. Their predictable behaviour makes them easier to validate and safer for regulated workflows. For example, MEMOTEXT's CAMH smoking cessation chatbot leveraged rules-based design to support quitting through timed reminders, structured coaching content, and logic-driven pathways validated through a randomized controlled trial.
- **LLM-Powered Conversational Agents:** These use generative AI to handle open-ended queries and provide nuanced responses. While they offer richer interactions and adaptability, they require additional layers of governance, prompt control, tuning, and human oversight to ensure safe deployment in healthcare contexts. In the *Keeping in Touch (KiT)* diabetes transition tool co-developed by SickKids, UHN and MEMOTEXT, a hybrid approach allows a conversational agent to manage navigation, while retrieving validated responses from a curated database.

These capabilities can also be hybridized whereby an LLM handles natural language understanding and routes requests to pre-validated responses, ensuring control without losing the user experience benefits of natural conversation.

Cultural Sensitivity and Accessibility
Programs like REACH-ES (Spanish-language adaptation) demonstrated the importance of culturally sensitive content. MEMOTEXT designed messages at a grade 5 reading level, localized terminology, and incorporated community-informed phrasing to ensure accessibility for Spanish-speaking adults with Type 2 diabetes leading to improvements in refill adherence and confidence in self-management.

Behavioural Segmentation and Persona-Based Engagement
Engagement can be greatly enhanced by tailoring not just to clinical profile but to behaviour. Segmenting users based on interaction frequency, risk profile, and prior responsiveness. For example, high-frequency responders can receive goal-oriented messages, while low-engagement users may receive re-engagement nudges. These approaches are often layered with persona design e.g., messages for "Confused First-Timers" vs. "Busy Chronic Users" to ensure the content speaks directly to each user's experience.

MEMOTEXT
Literature Summary 2025

For more detailed case studies, results, and implementation learnings, see the [MEMOTEXT Literature Review 2025] or visit https://mtxt.io/GetLit

Part III

Monitoring, KPIs, and Feedback Loops

Monitoring and measurement form the backbone of effective AI-driven engagement, enabling continuous improvement and real-world impact. Evaluation must be a native concept within an AI data-centric environment. It inherently cannot be external, episodic, or retrofitted but rather needs to be built in. Monitoring and evaluation form the basis of culture of safety and stewardship. Evaluation is embedded into the data strategy, not just as performance review but as a foundation for continuous learning and future decision-making.

Everything is measured not because we always know what we'll do with that information today, but because we recognize the value of being able to surface, revisit, and act on it when the moment comes. This kind of quantification isn't about surveillance or micromanagement it's about accountability, and the ability to evolve. By establishing clear metrics and feedback loops, healthcare organizations can ensure that AI systems are not only performing but evolving based on patient, clinician, and system needs.

Clinical Outcomes & Patient Safety

AI in digital engagement should ultimately demonstrate improved health not just clicks or responses. Effective measurement strategies align AI performance with clinical goals and safety benchmarks.

- **Health Impact Metrics**: Go beyond engagement. Monitor objective outcomes such as improved medication adherence (e.g., proportion of days covered, medication possession ratio), reduced HbA1c levels in diabetes management, or decreased emergency room utilization.

- **Patient Safety Surveillance**: Leverage AI to flag signs of deterioration, such as changes in symptom reporting or adherence behaviour, triggering timely human review or escalation.

- **Risk Mitigation Audits**: Routinely assess whether AI-generated communications inadvertently introduce clinical risk (e.g., conflicting advice, inappropriate timing). Validate models continuously—not just at launch.

Experience & Human Feedback Loops

Safety, personalization, and efficacy all improve when patient and clinician voices are integrated into iteration cycles.

- **Patient Feedback**: Collect qualitative input (surveys, in-message reactions, interviews) to fine-tune language, timing, tone, and emotional resonance. This ensures the AI doesn't just sound human but actually supports the human experience.

- **Clinician Insights**: Gather input from frontline staff to understand where AI communications help or hinder workflow. Use this feedback to redesign message logic or trigger escalation thresholds.

- **Continuous Learning Systems**: Build in mechanisms for ongoing improvement—whether through micro-randomized trials, A/B testing, or usage pattern analysis—to ensure the system adapts and stays relevant as patient populations and clinical practices evolve.

Engagement & Behaviour Analytics

Beyond vanity metrics understand how, when, and why users engage.

1. **Engagement Patterns**
    - **Basic Metrics:** Task completion, response rates, opt-ins.
    - **Content Metrics:** Click-through rates, goal completions, time-to-action.
    - **Temporal Trends:** Measuring engagement by day or time.
    - **Retention & Drop-Off:** When and why users disengage.
2. **Interaction Flow & Media**
    - **User Journeys:** Analyze sequences and journey mapping (e.g., survey → resource click → follow-up).
    - **Channel Comparison:** SMS vs. email vs. push notifications what drives better results? (Spoiler alert its SMS)
    - **Message Design:** Impact of content length, format, or tone on engagement.
3. **Frequency Optimization**
    - Compare response rates across message frequencies to avoid fatigue and optimize cadence.

Advanced Analytical Techniques

Use AI to continuously refine itself based on real-world behaviour.

- **NLP & Sentiment Analysis:** Detect emotion and trending topics in open-ended responses.
- **Clustering:** Segment users by engagement behaviour for tailored interventions.
- **Reinforcement Learning:** Dynamically optimize message timing using real-time data.
- **A/B Testing with ML Insights:** Test and validate personalized message variants.
- **Time-Series Analysis:** Predict optimal intervention windows using historical trends.

These interconnected monitoring strategies ensure AI-driven engagement is not static but smart, adaptive, and accountable. By grounding AI implementation in rigorous feedback and measurable outcomes, organizations build trust, demonstrate value, and ensure efficacy. Ultimately, success lies in creating systems that learn, adapt, and continue to improve.

LLM-as-a-Judge

Monitoring AI outputs for safety and consistency is no small task with large cost implications needed to resource the appropriate team to provide manual human review of AI engagement. This is where LLM-as-a-Judge can step in to help scale monitoring.

LLM-as-a-Judge refers to the process of using a large language model to evaluate, score or critique the output of another generative (or rules based) AI model instead of relying on manual moderation. This approach has multiple benefits:

- **Enables scale** across large volumes of outputs, which would be extremely time consuming for human moderators to manage
- **Support real-time monitoring and quality assurance** by flagging content for additional review
- **Mitigate common generative AI risks** such as hallucinations, bias or harmful language

One potential implementation example is through model evaluation. AI is a great source to offer resources and answer commonly asked questions. This was the foundation in which the KiT AI Chatbot at MEMOTEXT was developed. Due to the potential risk of hallucinations and harmful responses a LLM judge was added to score response and ensure all generated responses were safe and accurate.

Challenges and Mitigation

Effectively deploying AI in healthcare requires addressing several critical challenges to ensure systems operate responsibly, ethically, and securely. Organizations must proactively confront biases that could exacerbate disparities, maintain rigorous transparency to build trust with users, thoughtfully navigate resistance from patients and clinicians, and uphold stringent privacy and security standards to safeguard sensitive health data. Deploying AI responsibly requires addressing several critical challenges:

1. **Bias and Fairness:**
   Regularly audit AI systems for:
   - Training data biases that affect accuracy for underrepresented groups.
   - Labeling biases stemming from subjective judgments.
   - Statistical biases potentially amplifying unintended disparities.

2. **Transparency and Trust:**
   Clearly communicate AI system capabilities, ensuring transparency about limitations and providing clarity to maintain realistic expectations and facilitate human oversight.

3. **Patient Resistance:**
   Patient education about AI benefits, coupled with assurances of the complementary role of AI in healthcare, helps build acceptance without undermining human relationships. Adoption in a wider context requires thoughtful change management and direct line of sight from problem to solution.

4. **Privacy and Security:**
   AI involves handling large volumes of sensitive health data. Maintaining rigorous encryption standards, robust data storage solutions, and strict access controls are essential. Continuous monitoring across system interactions, data exchanges and vulnerabilities is key to ensuring safety. Large language models (LLMs) and conversational AI are vulnerable against prompt injection attacks, social engineering risks, and potentially harmful hallucinations.

Zero Trust is a security framework /concept based on the principle of *"never trust, always verify"*. It requires authentication and continuous validation at every level, from device identity to system interactions, across all transactions, whether it's multi-factor authentication, packet transfer protocols, or endpoint access policies. AI systems in healthcare must not assume trust at any level. Every access request, conversation, and data exchange in this scenario are continuously authenticated and validated to ensure safety and data integrity.

Some operational security examples include:

- ✓ **Identity & Access Controls:** Permission/Role-based access to different data within AI tools e.g., clinicians see only what they need, while administrative staff cannot access clinical outputs.
- ✓ **Real-Time Auditing:** Implementation of real-time logging and anomaly detection to monitor model outputs and flag irregularities or access violations.
- ✓ **Prompt Filtering & Output Controls:** Use of AI-specific security layers to intercept potentially unsafe prompts, prevent harmful output generation, and filter hallucinated or high-risk responses.
- ✓ **Stress Testing:** Regularly challenge the system with edge cases and adversarial inputs to detect vulnerabilities, particularly in large language models (LLMs) and conversational agents.

This is not a comprehensive list, but rather a prompt for application of rigorous security principles and thinking to LLMs, AI workflows, and patient-facing conversational agents, especially given the unique risks: prompt injection, social engineering, opaque decision paths, and hallucinations.

The Road Ahead

AI is only becoming more deeply embedded in healthcare, across EMRs, clinical documentation, diagnostics, imaging, care coordination, and patient engagement at a minimum. The possibilities are genuinely exciting. The accuracies and efficiencies offered by AI from decision support to automation and expert systems are accelerating at a pace that feels equal parts inspiring and disorienting. Agentic AI will unlock transformations we can't even yet predict. While we make no claims or predictions, we know that a science-fiction like future is materializing.

What won't change are the human realities of care. Clinicians will still carry the ethical weight of responsibility. Validation will remain non-negotiable. Health systems will continue to grow in complexity and scale, generating more data than ever. Social determinants like income, geography, and systemic inequity will continue to shape outcomes. And the fundamentals of self-management, behaviour change, adherence, and lifestyle will still drive long-term health.

AI may enhance workflows and unlock insights at an unprecedented scale, but it will never replace the fundamental need for human oversight, trust, and accountability in care. The challenge ahead is not just adopting AI but governing its integration responsibly ensuring it serves patients, supports clinicians, and strengthens healthcare systems without compromising safety, fairness, or transparency.

This playbook is just a primer for beginning an AI engagement development journey. The next step is **action** for healthcare leaders, clinicians, researchers, and AI developers that build, buy and adopt systems to prioritize trust, safety, and real-world impact. We hope to continue this exciting journey towards in our mission to make health data useful to improve outcomes across the healthcare spectrum.

Acknowledgement:

This document was developed through an iterative co-authoring process integrating humans (Amos Adler, Wenjia Zhou, Keerthana Mundlapudi) expertise with AI-assisted drafting. Additional human editing, iterations and contributions were provided by Dr. Darren Larsen. The collaboration involved structured ideation, content refinement, and real-time discussion. ChatGPT was used as an adaptive research assistant, thought partner, and editorial support tool.

Specifically, this process included:

- Structured Ideation: Humans provided initial drafts, conceptual frameworks, and the structured outline and iterative content development.
- Content Expansion & Refinement: The AI assisted in fleshing out certain concepts, research and reference provision, adding depth to sections. Through real-time dialogue, feedback loops, and clarification of intent, the document was refined to ensure coherence, domain specificity, and strategic alignment with the intended audience.