FUTURE HEALTH | A GLOBAL INITIATIVE BY ABU DHABI | H2S HACK2SKILL

Future Health Digital Innovation Challenge

From signals to solutions.

Co-creator Microsoft

Co-creator inception42

Co-creator Mohamed bin Zayed University of Artificial Intelligence

Team Name: **ZenAIDigital (product: ZenAI Behaviour Twin)**
Team Size: **1**
Problem Statement: **Wellbeing · Challenge 09, Sustaining Healthy Behaviors**

Start with the person: Noura, 41

Additional slide. Noura is demo persona P03: synthetic, invented for illustration

THE PERSON

Noura, 41, Khalifa City

Emirati mother of four. Chose a 15-minute family walk after the evening meal, 5 days a week. Her phone counts her steps.

THE PROFESSIONAL

Partner clinician (DoH-licensed)

Sees people after the habit has gone. Gets a shared report and FHIR export only with Noura's consent.

THE PROGRAMME LEAD

Provider wellness programme

Runs campaigns but cannot see who has lapsed or why. Sees groups of 20+ only. Owns the outcome.

THE MOMENT THAT MATTERS **A heatwave and busy Ramadan evenings: Noura misses 3 planned days in a row. Most apps break the streak, keep the same goal and go quiet. That week decides whether she stops for good.**

Why it stops: five whys

1 She stopped walking. Why?
2 Three missed days broke her streak. Why did that matter?
3 The app treated it as failure and kept the same goal. Why?
4 Plans are set once, at peak motivation. Why?
5 Nothing plans for the lapse.

Root cause: **support is built for starting, not for recovering.**

How might we

help an adult in Abu Dhabi who has missed a week restart a routine she chose, within 14 days, without guilt?

The rules act, not just advise: the plan gets lighter with a 24-hour undo; one restart plan after a lapse; increases are only proposed and she decides.

Measured and owned: back on plan within 14 days and activity against her own baseline, owned by the partner's programme lead.

Problem statement: Challenge 09, Sustaining Healthy Behaviors

The challenge, the problem and the objective from the brief, and how ZenAI Behaviour Twin answers each

Brief	What it asks	How we answer it (release v8.6.3, movement)	Status
The challenge	Help people start, maintain and recover healthier behaviours, with support that adapts to their goals, circumstances, motivation and setbacks over time	Start: safety screen, then one action the person chooses (what, when, where, minutes, days). Maintain: daily check-in, at most one reminder a day, increases only with consent. Recover: a lapse is spotted early, the plan gets smaller once, a welcome back follows and progress is never reset	**Built**
The problem	Information alone rarely sustains change; motivation dips; routines break; one lapse can end engagement; language, culture, health literacy, work patterns, disability and access differ	Every message carries a small action; a motivation barrier gets two tiny options; Ramadan and iftar timing, heat (indoor options), shift work and rest days; Arabic and English; homemakers, students and people seeking work; People of Determination count minutes with adapted actions; plain language and an optional guide	**Built** (Arabic and accessibility review pending)
The gap	Assets exist (screening, patient apps, Health Age, wearables, campaigns, workplaces, gamification) but no adaptive loop turns them into small actions, learns what works, supports recovery and measures sustained change	The loop is built for movement and learns what works for each person: if an action keeps failing, it suggests one that fits the person's main barrier, and they decide. FHIR export to patient apps; optional wearable data; workplace view for groups of 20+. Health Age and campaign links are designed only	**Built** (links designed)
The objective	Help people start, keep up and get back to healthier behaviours, with support that adapts over time	Working end to end on synthetic people; whether it changes real behaviour is the pilot question	**Built; benefit untested**

Problem statement: the five desired outcomes

What the brief asks for, what release v8.6.3 delivers, and what is still to do

Desired outcome	What we deliver	Status	Still to do
Adaptive personal plans	Goals, preferences and health context become one small chosen action; smaller after missed days, bigger only with consent; steps or minutes	**Built**	A second behaviour (sleep or diet) deliberately not added yet
Continuous learning and re-engagement	Spots a lapse early, captures barriers, suggests a better-fitting action, offers a different reminder time or frequency after five ignored, no blame, progress kept; the guide explains why the plan changed	**Built**	Set thresholds on real data; ML and personal twin built on synthetic data only (A8)
Evidence-based behaviour support	Technique codes and sources shown with limits; family and companion actions; workplace group view; opt-in risk cards always paired with an action	**Partial**	No incentives or game screens; no campaign flow
User autonomy and inclusion	Person chooses and edits the action, approves increases, consents per purpose and can withdraw, export or close; Arabic and English; larger text, high contrast; chosen guide or none; employers and insurers see nothing individual	**Built** (accessibility partial)	Arabic review; accessibility testing with users
Sustained outcome measurement	Activity against own baseline, persistence, recovery within 14 days, retention, equity by group with small groups hidden; a pre-registered comparison with a fixed goal	**Partial**	No biomarker yet; real data only in the pilot

Built = working and tested in release v8.6.3 on synthetic data · **Partial** = part built, rest designed · demo deployed on Azure Container Apps; nothing externally validated yet

Ecosystem fit: we make existing initiatives last

Additional slide. Existing assets from the brief, and how the recovery loop connects to each (release v8.6.3, synthetic)

Existing asset	Status	How it connects
Wearable data	**Built (sample)**	Import of daily totals from Apple Health or Health Connect tested on synthetic exports; phone app shell designed; manual entry always works
Patient app journeys	**Built**	Consented FHIR R4 report and export; Malaffi and UAE Pass later through the partner
Workplace initiatives	**Built**	Workplace view for groups of 20 or more; employers and insurers never see individuals
Segmentation and impact measurement	**Built**	KPIs and equity breakdowns, programme simulator, insights check; ML segments built on synthetic data
Health Age scoring	**Designed**	Integration point through the partner, Stage B
Community campaigns	**Designed**	Campaign sign-ups flow into the recovery loop
Gamification and incentives	**Not built by choice**	Streaks punish lapses; we reward a return (best of 54 arms, Milkman 2021); incentives as a pilot option
Screening programmes	**Not linked**	A "next step after screening" referral is a later link

The loop between them: a small chosen action → adapt → recover after a lapse → measure lasting change (KPI-13, KPI-02, KPI-04, KPI-07). We don't replace these assets; we make them last.

Why people and programmes choose it

Additional slide. Our hypotheses for the pilot to test, not results

Participant

"Helps me get back on track, not makes me feel bad for falling off."

- No broken streak or guilt: an easier plan and a restart
- My plan, my choice; increases only proposed
- Fits life here: after iftar, indoor in the heat, Arabic or English
- My employer can't see me; free through my programme

Programme or provider

"Your campaign gets them started; we keep them going and show it lasted."

- Brings back people who lapse
- Measures lasting change, not downloads (KPI-13)
- Groups of 20+ only; consented clinician link
- Auditable rules; UAE-only, vendor-neutral hosting

	Campaigns and challenges	Phone trackers	Patient apps	ZenAI Behaviour Twin
Keeps people going after a lapse	Weak	Weak	Weak	**Core strength**
Personal, chosen action	Rarely	Generic goals	Rarely	**Yes**
Local context (Ramadan, heat, Arabic)	Sometimes	No	Partly	**Yes**
Measures lasting change	Rarely	No	Rarely	**Yes**
Privacy-safe group view (20+)	No	No	No	**Yes**

Team members*
[Name, country, role, organization, age confirmation]

S.No	Name	Country	Role	Organization	Age confirmation
1	Sunil Mysore Kempegowda	Australia	Product and architecture lead	ZenAIDigital Pty Ltd	Confirmed, 18+
2					
3					
4					
5					

Team Capability Summary*

Maximum 1,000 characters

Dr. Sunil Mysore Kempegowda is a senior Enterprise Architect and digital transformation leader with extensive cross-sector experience across the Asia-Pacific. His practice bridges healthcare, education, law enforcement, financial services, media, and environmental management. Awarded his doctorate from the University of Technology Sydney in 2018, his research pioneered an Enterprise Architecture-driven approach to digital transformation and established the extended Comprehensive Capability Maturity Model (CCMM) assessment framework.

He is the author of Enterprise Architecture in the Era of Disruptive Digital Technologies and co-author, with Dr. Zenon Chaczko, of Governing a Moving Target.

Through ZenAIDigital Pty Ltd (Sydney) he leads ZenAI Behaviour Twin, applying 25+ years of health, government and AI architecture to safe, measurable behaviour support.

25+ years
enterprise and AI architecture

PhD
enterprise architecture, UTS

110/110 · 82/82
rule scenarios · browser tests, v8.6.3

AR · EN
right-to-left, inclusive design

866 of 1,000 characters

Architecture or workflow diagram*

Daily loop: rules decide, the person chooses

Every check-in runs the same versioned DMN tables in the app, the simulator and the tests

1. **Check in**
   Done / partly / not today; steps or active minutes; optional barrier and note
2. **Safety first (DEC-02)**
   Red flag or pain: coaching stops; 998 and see your doctor
3. **Classify the day (DEC-01)**
   On track, at risk (3 planned days missed in a row), lapse (7 days)
4. **Plan decision (DEC-05)**
   Smaller plan with 24-hour undo; one restart plan per lapse; increases only proposed
5. **Message and reminder (DEC-04, DEC-06)**
   Encouragement by default; opt-in risk cards always paired with an action
6. **Person decides; guide explains**
   Accept, undo or change the action; optional avatar guide explains why (template speech, no AI model)
7. **Trace and measure**
   CloudEvent with rule ID, version and traceparent; KPIs with groups below 20 suppressed

Architecture: what is built and what is designed

Release v8.6 · demo live on Azure Container Apps, UAE North · synthetic data only

Experience — Arabic and English, right-to-left

Participant app and avatar guide **Built**	Professional report and FHIR export **Built**	Programme and workplace views (n ≥ 20) **Built**	Staff queue and appointments **Designed (Stage B)**

Decision core — pure code, no cloud SDK

DMN 1.4 tables DEC-01 to DEC-06 **Built**	Rules: goals, recovery, consent, retention **Built**	Evidence-tagged messages and cards **Built**	KPIs and equity strata **Built**

Analysis — propose only

Programme policy simulator **Built**	Insights check (pre-set rule) **Built**	Agent functions, template mode **Built**	Azure OpenAI ranks approved actions **Live**

Minimum deployment, UAE North — Azure

Container Apps: app and rules API **Live**	PostgreSQL, UAE-only, 25-year retention **Designed (pilot)**	Sign-in and consent records **Designed (pilot)**	Traces: Application Insights, Log Analytics **Live**

Built = tested: 109/109 rule scenarios, 82/82 browser tests · **Live** = running on Azure now · **Designed** = specified, not built

Live version: behaviourtwin.zenaidigital.com, tab "Architecture & evidence" (Azure services confirmed by /api/health); details in annex A9 and A12

Target users and intended outcome*

Initial focus (hypothesis to test)

Adults in one provider-led wellness programme in Abu Dhabi who drop off after a setback such as heat, Ramadan, shift changes, travel or family duties.

Participants (8 synthetic demo personas, more occupations)

- Shift workers, office staff, students, homemakers, people seeking work, retirees
- Emirati women organising family life; men 45–64 with cardiometabolic risk; seniors 65+
- People of Determination: wheelchair users counting active minutes; people with low vision

Professional users

- Clinicians see a report only when the participant shares it; programme leads see groups of 20 or more only

Intended outcome

People restart after a setback and keep up a movement routine they chose, with more of them at or above their own starting activity at 12 weeks.

Challenge outcomes

- **Adaptive personal plans:** chosen action, time and place; smaller plan after missed days; increases only with consent
- **Continuous re-engagement:** welcome-back, one restart plan per lapse, optional guide
- **Sustained measurement:** activity against own baseline, recovery, retention, equity

66.1% of UAE adults were insufficiently active in 2022, the highest of 197 countries (modelled; Strain 2024)

How will success be measured?*

Measure	Definition	Prototype (synthetic, v8.6.3)	Pilot
Primary: above own baseline	Share with mean activity in weeks 9–12 at or above own baseline week (steps; minutes in the pilot)	Computed (KPI-13)	Primary outcome vs equal-attention fixed goal
Persistence	Share completing at least half of planned days, weeks 9–12	Computed (KPI-01)	Secondary
Recovery after a lapse	Back on plan within 14 days; days to recover	Computed (KPI-02, KPI-03)	Secondary
Retention	Still checking in at week 12; 6-month follow-up	Computed (KPI-04)	Secondary
Change in activity	Median change from baseline	Computed (KPI-05)	Secondary
Equity	All measures by age, gender, language, occupation, disability; groups below 20 suppressed	Computed (KPI-07)	Secondary
Safety and burden	Reminders per week; safety pauses and adverse events	Reminders computed (KPI-06)	All monitored monthly
Engagement features	Avatar guide on vs off; opt-in risk cards: understanding, distress, blame	Not measurable	Embedded secondary questions

Go / no-go to Stage B: primary outcome no worse than the fixed goal beyond a margin agreed in advance, no rise in safety events, comparable retention, partner agrees. Software: 110/110 rule scenarios and 82/82 browser tests pass on the published build.

Evidence supporting the approach*

Citations, studies, benchmarks, or prior testing.

Published evidence

- **Small but real effects:** activity trackers raise activity by about 1,235 steps/day, low to moderate certainty, 121 RCTs (Larsen 2022); effects roughly halve beyond 6 months (Mönninghoff 2021), so sustaining is the gap
- **Adaptive goals** beat fixed goals over 12 months (Adams 2022); self-chosen goals sustained gains (ENGAGE, Patel 2021)
- **Rewarding a return** after a missed workout was the best of 54 arms in a 61,293-person megastudy (Milkman 2021)
- **Chatbots** modestly raise daily steps, SMD 0.28, mostly low-quality trials (Singh 2023): the basis for testing the guide
- **UAE need:** 66.1% insufficient activity, the highest of 197 countries (Strain 2024)

Prior testing (software and simulation only)

- **110/110** rule scenarios and **82/82** browser tests on the exact published build, including safety, consent, suppression and Arabic
- **Insights check:** found a planted pattern in 19/20 held-out datasets, 0/20 false alarms; false alarms still occur
- **Machine learning:** early lapse signal reaches held-out AUC 0.72 on synthetic people; most flags are false alarms (27% precision), so flags only offer easier support
- **Honest simulation result:** adaptive restart keeps more people on plan, but a fixed goal leaves more above their own baseline. A pre-registered regrowth policy narrowed the gap (+6.7 points) and still lost to a fixed goal (−13.7). The pilot must settle this with real people
- **Not yet done:** user walkthroughs, expert review, real-world outcomes

Data sources and legal right to use them*

List source, type, owner, sensitivity, permissions, and whether personal data is used.

Source	Type	Owner	Sensitivity	Permission	Personal data
Synthetic participants and 12-week daily records	Generated by the simulator	ZenAIDigital	None: no real people	Own	No
Synthetic personas (8 participants, 4 staff roles in the demo)	Written test cases	ZenAIDigital	None	Own	No
Published statistics (SCAD 2024 population mix, UAE National Health Survey, WHO, IDF)	Aggregate figures, cited	Publishers	None	Citation	No
Evidence library (journal articles)	Citations and DOIs only; no text copied	Publishers	None	Citation	No
LOINC and UCUM codes in the FHIR export	Terminology	Regenstrief Institute	None	Free licences; attribution to add before pilot	No
Optional calibration: NHANES minute steps; Open-Meteo Abu Dhabi weather	Public datasets (not used in v8.6.3)	US CDC (NCHS); Open-Meteo	Low	US federal public domain; CC BY 4.0	No

No personal data is used. The demo rejects common real identifiers (Emirates ID, email, diagnosis text). A pilot would collect data only with consent per purpose, stored in the UAE and kept at least 25 years (Federal Law No. 2 of 2019; legal review pending).

Risks and safeguards*

Privacy; security; bias; explainability; misuse; human oversight; failure modes.

Area	Main risk	Safeguard	Status
Privacy	Employer or insurer sees individuals	Organisations see groups of 20+ only; sharing is per purpose and revocable	**Built**
Security	Health data leaves the UAE or is exposed	UAE-only hosting, identity and audit; no personal data in model prompts	**Designed**
Bias	Works only for the already active or advantaged	Equity strata on every measure; adapted actions; minutes as well as steps	**Built**
Explainability	People do not know why the plan changed	Versioned rules with IDs on every decision; the guide explains in plain words	**Built**
Misuse	Shaming, fear or weight stigma	Encouragement by default; opt-in risk cards with an action; banned-wording tests	**Built**
Human oversight	Rules or AI change behaviour unseen	Agent functions only propose (built); rule changes need a new version, full re-test and approval (process)	**Built**
Failure: safety	A symptom is missed or coached through	Safety screen first; red flags stop coaching and show 998; guide stays silent	**Built**
Failure: guide	A friendly avatar is mistaken for a clinician	Labelled AI helper; template speech only; never a plan input	**Built**
Failure: evidence	Simulation is read as proof of benefit	Labelled synthetic; results shown even when unfavourable	**Built**
Failure: AI or ML	A model gives advice, errs or treats a group unfairly	AI and ML only propose; never-tier blocks; wording check with fallback; fairness by group; shadow before real use (A8)	**Built (synthetic)**

Hardware requirements

Required if wearable, sensor, connected device, or other hardware is there. Include sourcing and safety.

No hardware required

Works on any phone or computer browser. The demo is one web page.

Optional: devices people already own

Phone step counts or a wearable, chosen by the participant, read as daily totals from Apple Health or Health Connect. Import tested on synthetic exports; the phone app shell is designed. Manual entry always works.

Wellness data only

Consumer devices give wellness data, never medical measurements. No device-driven alerts or thresholds.

Sourcing and safety

Nothing is procured or supplied. Safety screening and red-flag rules apply whatever the data source. The demo uses simulated self-reported entries; a synthetic device-format sample is in the data package.

If the application requires login authentication, please provide valid access credentials
or enable a "Continue as Guest" option with all necessary features and functionalities accessible to guest users.

No login needed

The competition demo runs on Azure Container Apps in UAE North: no accounts, no passwords, synthetic data. The future pilot adds sign-in, consent and personal data controls.

Link: https://behaviourtwin.zenaidigital.com

Backup: BehaviourTwin_Prototype_Demo_v8.6.3.html opens in any modern browser (offline with fallback fonts)

Data: synthetic only; nothing is stored after the page closes

How judges can test it

1. Open the link and press "Start a movement-support demonstration" (10 guided steps, about 3 minutes)
2. Watch the baseline, check-ins and a lapse; the plan shrinks without blame; keep or undo it
3. See recovery time, then the aggregate Executive view and the trace for each decision
4. Switch views: Participant, Professional, Programme, Workplace lead, Executive, Agents, Rules & proof
5. Auditors: "For auditors: checklist (25 tests)" gives each check, its expected result and a Go button; "Architecture & evidence" shows Azure services and claims

Path to pilot and scale*

Stakeholders, integration, regulatory needs, timeline assumptions, and sustainability.

Now
Prototype v8.6.3, synthetic data, tests on the published build

Nov 2026
Register by 10 Nov; submit 20 Nov; partner talks

Stage A
12-week pilot plus 6-month follow-up; coaching only

Gate
Go / no-go on the primary outcome, safety and retention

Stage B
Staff workflow; AI and ML sub-study; then scale

Stakeholders
DoH Abu Dhabi (organiser and regulator)

One DoH-licensed provider wellness programme (partner, to confirm)

Participants, partner clinicians, programme leads

Integration
Shared report and FHIR R4 export to the partner

Clinical care stays in the partner's own systems in Stage A

Pilot app: Apple Health or Health Connect daily totals; later Malaffi and UAE Pass

Regulatory
DoH Tele-Medicine Standard for any clinical follow-up

Federal Law No. 2 of 2019: UAE storage, 25-year retention

ADHICS via the partner; ethics approval for the pilot

Sustainability
Licence per active participant plus set-up fee (price to test)

Embeddings only (cents a month); vendor-neutral containers in a UAE region

Staff time owned by buyer or facility; no revenue claimed

Our ask: a partner cohort, accountable clinical and behaviour review, and evaluation support. **Outcome owner:** the partner's programme lead.

Pre-existing IP and third-party materials*

Identify ownership, licenses, open-source components, and restrictions.

Component	Owner	Licence	Restrictions
Code, decision tables, specifications, synthetic data, avatar illustrations and message text	ZenAIDigital Pty Ltd	Proprietary	Pre-existing and created for this entry; no third-party code in the app bundle
esbuild; fast-xml-parser (build tools)	esbuild authors; NaturalIntelligence	MIT	Build time only
@cucumber/cucumber (tests); Playwright (browser tests)	Cucumber; Microsoft	MIT; Apache-2.0	Test time only; not shipped
Fonts: IBM Plex Mono, IBM Plex Sans Arabic, Readex Pro (Google Fonts)	IBM; font authors	SIL Open Font License 1.1	Loaded from Google Fonts
Standards: DMN, FHIR R4, LOINC, UCUM, CloudEvents, Open mHealth	OMG, HL7, Regenstrief, CNCF, Open mHealth	Open specifications; LOINC and UCUM free with attribution	LOINC and UCUM notice to add before pilot
Microsoft Azure (planned hosting)	Microsoft	Commercial terms	Not yet used
Cited research	Publishers	Citation only	No text or figures reproduced
Patents	None	None claimed	No exclusive method claimed

Conflict of interest declaration*

Disclose relationships with organizers, partners, or judges.

Declaration

The entrant declares no relationships with the organisers (Department of Health Abu Dhabi), the co-creators (Microsoft, MBZUAI, Inception), challenge partners or judges.

[To confirm by the entrant before submission. List any employment, contract, partnership, funding or family relationship here.]

Note: the planned hosting uses Microsoft Azure under standard commercial terms. Using a co-creator's commercial service is not a relationship with the organisers, but it is stated here for transparency.

A lapse is not a failure: Noura's illustrative journey

Invented for illustration, not a result. No outcome numbers are shown.

Baseline week

Noura picks a 15-minute family walk after the evening meal, 5 days

Weeks 1–3

Steady; guide encourages; no plan changes she did not choose

Heat barrier

Misses 3 planned days in a row; plan gets lighter: a shorter indoor walk, 24-hour undo

Lapse

A week away; one restart plan from her recent days

Welcome back

Done on 4 of 7 days; a small step up is offered, she decides

Inclusion: a wheelchair user counts active minutes of wheelchair pushes; every message, chart and the FHIR export use minutes, and suggestions stay within adapted actions.

Programme view: the programme lead sees recovery for groups of at least 20 only, never the individual, with the honest simulation result alongside.

Additional Slide

Annex: technical specification

Replaces the separate technical document. Release v8.6.3, 21 November 2026

In one paragraph

A rules-first coaching service that helps adults in Abu Dhabi restart a movement routine they choose after a setback. Versioned decision tables decide safety, day status, plan changes, messages, reminders and who may see what; the person approves every increase. The same tables run the app, a programme policy simulator and 192 automated tests; machine learning and a personal twin run on synthetic data and only propose. The demo runs live on Azure Container Apps in UAE North on synthetic data; nothing is externally validated yet.

Contents

A1 Release and how to verify it
A2 Architecture components
A3 Decision tables
A4 Data and simulator assumptions
A5 Safety, privacy and governance
A6 Testing, results and assumptions
A7 Dependencies, limitations and references
A8 AI, ML and cloud: built vs designed
A9 Architecture and evidence
A10 Auditor checklist: the journey (A1-A10)
A11 Auditor checklist: safety, privacy, Azure (B1-B15)
A12 Azure service mapping: live and designed
A13 Cost at scale and beyond movement

A1 Release and how to verify it

Everything in this deck is bound to these files and hashes

Item	Value
Demo	BehaviourTwin_Prototype_Demo_v8.6.3.html: one self-contained page, no login, synthetic data only
File SHA-256	8C581E4522F3C8A31BBD72DC57C3842728D0C6E48AB8FF7F4A2086EB9A63D365
Bundle SHA-256	e8f6555fa5592456359cbac5ea66e8c586196d503ab9df6dfebedec4ceda1504
Tests on that bundle	110/110 rule scenarios (Gherkin) and 82/82 browser tests (Playwright); results embedded in the page and in release-manifest_v8.6.3.json
Code	BehaviourTwin_Prototype_Code_v8.6.3.zip. Run npm install, then npm test (rules) and npm run build:release (test, build, browser tests, rebuild, manifest)
Claims	BehaviourTwin_Claims_Manifest_v11: every claim with status, file, test, limitation and owner. Scale: built locally, demonstrated, designed, deployment verified, externally validated
Status today	Built and demonstrated; public demo and decision API deployment verified on Azure Container Apps (UAE North) with Application Insights. Nothing externally validated

A2 Architecture components

Pure decision core, Azure behind ports; the first deployment is deliberately small

Decision core (built)

DMN 1.4 evaluator (hit policy FIRST); goals, recovery, consent and retention rules; evidence-tagged messages and cards; reports and KPIs. No cloud SDK.

Experience (built)

Participant app in Arabic and English, right to left; professional report with FHIR R4 export; guided 10-step judge demo with a trace per decision and a safety panel; programme, workplace and executive views (groups of 20+, persistence at weeks 4/8/12, recovery, equity flags, CSV); optional avatar guide.

Analysis (built)

Programme policy simulator runs synthetic people through the same rules for 12 weeks; insights check with a rule fixed in advance; four agent functions that only propose; ML-01, ML-02, ML-04 and a personal twin on synthetic data, proposals only. Live: Azure OpenAI embeddings rank approved restart actions after the rules; generative wording not used by design.

Deployment (demo live; pilot designed)

Live: Azure Container Apps in UAE North running the page and decision API, Application Insights and Log Analytics, release package in Storage. Designed for the pilot: PostgreSQL, identity, consent, audit, backups in UAE Central. Portable containers.

A3 Decision tables

Versioned DMN 1.4 tables shared by the app, the simulator and the tests

Table	Version	Decides	Possible results
DEC-02 Safety boundary	1.1.0	Runs first. Red-flag phrases or symptoms stop coaching; "yes" or "not sure" on activity limits, pregnancy or a long-term condition gives a gentle plan	Stop and show help (998); gentle starter; standard
DEC-01 Day classification	1.0.0	Day status from check-ins and planned days	Safety stop, paused, lapse (7 days), at risk (3 planned days missed in a row), disengaged, reminder, unknown, on track
DEC-05 Goal adjustment	1.0.0	Plan changes: decreases apply with a 24-hour undo; increases are only proposed	Restart plan (once per lapse), smaller plan, two tiny actions, hold, propose +10%
DEC-04 Message mode	1.1.0	Encouragement by default; opt-in risk information always with an action; never on the gentle plan	Supportive only; encourage; inform with action
DEC-06 Nudge (reminder)	1.0.0	At most one reminder a day, in the person's window, after iftar in Ramadan; after five ignored, offer a change	No reminder; defer; shift time; offer change; send
DEC-03 Report visibility	1.0.0	Who sees what	Own data; shared report with consent and valid link; aggregate 20+; suppressed; deny

A4 Data and simulator assumptions

No personal data is used; every person-like record is synthetic and marked as such

Simulator parameter	Value	Basis
Gender and age mix	67% men; 22 / 45 / 27 / 6% by age band	SCAD 2024; shaped assumption
Baseline steps by age band	6,800 / 6,300 / 5,800 / 4,500, SD 25%	Assumption; calibration path to NHANES (public domain)
Lapse shocks; dropout	About 3% a day after day 10; 0.4% a day, ×3 during a lapse	Shaped on published lapse and attrition rates
Accept increase proposals	70%	Assumption
Smaller plan easier in recovery	×1.15 completion (×1.0 in the pre-registered test)	Assumption that favours adaptation; tested both ways
Behaviour when a plan is done	Exactly the planned amount	Simplification: any lower goal means lower activity by construction

Formats: JSON; CloudEvents 1.0; Open mHealth step-count sample; FHIR R4 Bundle (steps LOINC 41950-7, minutes LOINC 55411-3 in UCUM min); Frictionless Data Package.

For a pilot: consent per purpose, UAE-only storage, records kept at least 25 years and archived on closure (Federal Law No. 2 of 2019; legal review pending).

A5 Safety, privacy and governance

What protects participants in the demo today, and what the pilot adds

Built and tested in v8.6.3

- Safety screen before any plan; pain on any day pauses coaching until reassessed; red flags show 998
- Organisations see groups of 20 or more only; employers and insurers are denied; sharing with a professional is per purpose and revocable
- Wording checks on every template, card and guide line in English and Arabic: no guilt, weight, prediction or medical advice
- Agent functions only propose; a rule change ships only as a new table version that passes the full test suite
- Avatar guide: labelled AI helper, template speech only, silent during a safety pause, never a plan input (same day tested with and without a guide)

Care model (designed)

- Stage A pilot: coaching only; clinical care stays in a DoH-licensed partner facility's own systems
- Stage B: staff queue, appointments and role- and licence-based access only after they are built and approved
- Clinical roles need a DoH licence; wellness coaches are non-clinical
- Follows the DoH Standard on Tele-Medicine (2020) for any clinical follow-up

A6 Testing, results and assumptions

Software and simulation evidence only; no real-world outcomes yet

Test	Result	Meaning
Rule contracts (110 scenarios)	110/110 pass	Decision tables, safety, consent, retention, reports, agent guardrails, units, guide
Browser tests (70)	82/82 pass on the published bundle	Journeys in both languages, accessibility controls, reports, views
ML-02 early lapse signal, held-out seeds	AUC 0.72; at threshold 45% sensitivity, 27% precision; fairness gap 0.04	Proposals only; most flags are false alarms; real-world performance unknown
Insights check, held-out data	Planted pattern found 19/20 (95% interval 0.76–0.99); 0/20 false alarms (0–0.16)	False alarms still occur outside this set
Pre-registered regrowth test (10 new seeds × 300, no assumed boost)	Above own baseline: fixed 75.4%, adaptive 54.9%, restart plus regrowth 61.7%; regrowth vs fixed −13.7 points (−16.9 to −10.6)	Regrowth narrows the gap but a fixed goal still wins in the simulator; the pilot tests real behaviour

Assumptions

- Thresholds (3 planned days missed in a row, 7-day lapse, +10% increases, groups of 20) await clinical and behaviour-change review; simulator parameters are not calibrated to Abu Dhabi data
- A DoH-licensed partner provides any clinical follow-up in Stage A; participants have a phone or browser; Azure UAE North can host the minimum deployment

A7 Dependencies, limitations and references

What the solution relies on and what it cannot yet show

Dependencies

- Build and test only: Node.js, esbuild and fast-xml-parser (MIT), Cucumber (MIT), Playwright (Apache-2.0); no third-party code in the app
- Fonts: IBM Plex Mono, IBM Plex Sans Arabic, Readex Pro (SIL Open Font License 1.1)
- Standards: DMN 1.4, FHIR R4, LOINC, UCUM, CloudEvents, W3C Trace Context, Open mHealth
- Services (designed): OCI containers on any UAE provider (reference Azure UAE North); PostgreSQL, OpenID Connect, OpenTelemetry, OpenTofu
- People: clinical and behaviour-change advisor, Arabic and cultural reviewer, DoH-licensed partner, pilot statistician

Limitations

- Synthetic data shows how the rules behave, not whether people change
- No user walkthroughs, expert review or real outcomes yet; pilot stack not deployed
- Staff workflow designed only; no FHIR validator run or accessibility audit; LOINC and UCUM notice to add
- Arabic text and avatars await review; ML and personal twin tested on synthetic data only; AI limited to ranking approved actions

Key references

Larsen 2022 BMJ 10.1136/bmj-2021-068047 · Mönninghoff 2021 JMIR 10.2196/26699 · Adams 2022 AJPM 10.1016/j.amepre.2021.09.014 · Patel 2021 JAMA Cardiol 10.1001/jamacardio.2021.3176 · Milkman 2021 Nature 10.1038/s41586-021-04128-4 · Singh 2023 npj Digit Med 10.1038/s41746-023-00856-1 · Strain 2024 Lancet Glob Health 10.1016/S2214-109X(24)00150-5 · DoH Tele-Medicine Standard 2020 · Federal Law No. 2 of 2019

A8 AI, machine learning and cloud: built vs designed

Release v8.6.3: built and tested on synthetic data; demo live on Azure Container Apps; nothing validated on real people

AI personal assistant

- Built: tool tiers. Acts within rules (explain, log a check-in, move a reminder); proposes (step up, sharing); never (medical advice, safety rules)
- Built: refused during a safety pause; rule-based replies
- Live: Azure OpenAI embeddings (UAE North) rank an approved library of restart actions after the rules; wording stays reviewed templates by design

Machine learning and personal twin

- Built: ML-01 routine; ML-02 early lapse signal (held-out AUC 0.72; never shown); ML-04 segments (20+)
- Built: personal twin prototype (14+ check-in days; proposes only)
- Designed: ML-03 what works for whom; ML-05 candidate patterns
- Models only rank or time options rules allow

Vendor-neutral cloud and devices

- Built: device import from Apple Health, Health Connect, Open mHealth; heart rate refused
- Live: demo container on Azure Container Apps, UAE North, with Application Insights; portable to any UAE provider
- Designed: Azure UAE North reference, backups in UAE Central
- Designed: phone app shell (iOS, Android)

Staging: Stage A no ML in decisions (consented data, offline analysis) → Stage B shadow → randomised sub-study with ethics approval → adopt only if gates pass: uplift in KPI-02 or KPI-13, no group worse off (KPI-07), no rise in reminder burden (KPI-06). Nothing predictive is ever shown to participants.

A9 Architecture and evidence

The same content is on the live site: behaviourtwin.zenaidigital.com, tab "Architecture & evidence"

Participant
check-in, own words

→

Rules first
DMN 1.4, DEC-01 to DEC-06; red flags stop, pain pauses

→

Azure OpenAI
embeddings rank 9 approved actions

→

Person chooses
keep or undo within 24 h

→

New plan
versioned; progress never reset

Trace per decision: rule ID and version, CloudEvents, W3C traceparent → Application Insights and Log Analytics

Programme and executive views: groups of 20+ only; persistence, recovery, equity

How the AI is bounded

- Rules run first; red flags show 998; pain pauses suggestions; no AI call
- Identifiers removed; text capped at 200 characters; not stored
- AI ranks approved actions with sources; it never writes advice
- Labelled keyword fallback if Azure is unavailable

Standards used

- DMN 1.4 (six versioned tables)
- HL7 FHIR R4 with LOINC; Open mHealth
- CloudEvents; W3C trace context; OpenTelemetry
- ArchiMate 3.1; BCT Taxonomy v1

Built · designed · not yet

- Built and tested: safety stop, smaller restart plan, live Azure rules API and AI ranking, outcome measures, Arabic and English
- Designed: sign-in, consent store, backups in UAE Central, staff workflow, phone apps
- Not claimed: health benefit, expert review, user or buyer feedback

A10 Auditor checklist: the journey (A1-A10)

Built into the live site: "For auditors: checklist". Each test has a Go button to the exact state, the expected result, and Pass / Fail / Note

ID	What is tested	Expected result
A1	Persona and goal chosen by the person	Noura, 41 (synthetic); family walk, 15 minutes, after the evening meal
A2	Seven-day baseline	Seven daily step values; average 3,784 steps
A3	One small action, plan v1	Plan v1 with action, minutes, timing and the reason
A4	Daily check-ins	Active-day count rises by 5 after "Check in for 5 days"
A5	Lapse detection	Status changes to paused or lapse; no streak broken
A6	Plan shrinks without blame	Plan v3: 5 minutes; "progress is never reset"; no guilt words
A7	Keep or undo the change	Keep gives plan v4; undo restores the previous plan
A8	Recovery time for her	"Back on plan 4 day(s) after the setback started"
A9	Aggregate outcome dashboard	Persistence 87/57/46%; median recovery 5 days; 76% within 14 days
A10	Trace of one decision	Rules and versions, output, plan version, KPI events, traceparent

A11 Auditor checklist: safety, privacy, measurement and Azure (B1-B15)

All figures come from a fixed random seed, so they are the same on every run (synthetic data)

ID	What is tested	Expected result
B1	Safety screen before any plan	Warning symptom: coaching stops, 998 shown, no plan
B2	Consent per purpose	Separate choices; coaching works without sharing
B3	Assistant refuses a medical question	"Refused", pointing to a clinician
B4	Assistant acts within rules	"Acted within rules"; check-in logged
B5	Arabic, right to left	Participant view in Arabic, right to left
B6	Wheelchair user, active minutes	Plan and check-ins in minutes, not steps
B7	Aggregates only	Groups of 20+; CSV without names or IDs
B8	Same rules run in Azure	Browser = Azure; fingerprints 20d73377 / ad1bf607
B9	KPIs computed in Azure	53% above baseline; 76% recovered within 14 days
B10	AI cannot bypass rules	Approved message with technique and source
B11	Live Azure service	/api/health: Container Apps, UAE North, current release
B12	Works on a phone	Fits the screen; tour works
B13	AI ranks restart actions; rules decide	Azure OpenAI embeddings rank the approved library
B14	Rules stop the AI on pain or red flags	Stopped by the rules before any AI call
B15	Architecture & evidence tab	Live Azure rows match /api/health

A12 Azure service mapping: live and designed

Resource group in UAE North; vendor-neutral containers, so the same build runs at another UAE provider

Azure service	What it does here	Status
Azure Container Apps (UAE North)	Public demo and decision API; same rules as the page, same fingerprint	**Live**
Azure OpenAI embeddings (UAE North)	Live semantic matching of approved restart actions after the rules; never writes advice	**Live**
Generative Azure OpenAI model	Not configured in this release; reviewed templates by design (health check says so)	**Not configured**
Application Insights	OpenTelemetry traces (rules.decide, assistant.match, KPIs) with rule versions	**Live**
Log Analytics	Container logs and trace store	**Live**
Azure Storage	Tested release package	**Live**
Custom domain and certificate	behaviourtwin.zenaidigital.com → CNAME to ca-zenai-behaviourtwin.kindsmoke-dbb61d85.uaenorth.azurecontainerapps.io; both serve the app over HTTPS (free managed certificate)	**Live**
PostgreSQL flexible server	Participant records, consent (FHIR), decision events; 25-year retention	**Designed for pilot**
Microsoft Entra ID / External ID	Sign-in for participants and staff; UAE Pass later via the partner	**Designed for pilot**
Key Vault, Private Endpoints	Secrets and private networking; managed identities	**Designed for pilot**
Azure Policy	Allowed locations: UAE North and UAE Central only	**Designed for pilot**
Backups in UAE Central (Abu Dhabi)	Geo-redundant backup; quarterly restore test (access must be requested)	**Designed for pilot**
Azure Machine Learning, AI Search	Model registry and monitoring; curated evidence index	**Designed for pilot**

A13 Cost at scale and beyond movement

Measured on 21 November 2026; prices checked the same day; extension is designed, not built

Cost at scale

- **Rules, not a chat model:** the six DMN tables evaluate one person-day in about 0.03 ms on one CPU core (27,000 synthetic person-days in 0.85 s). One million daily check-ins ≈ one minute of compute a day.
- **Only optional AI call:** embeddings for a person's own words, about US$0.02 per million tokens (text-embedding-3-small).
- **Containers scale with demand:** Azure Container Apps; background jobs can scale to zero (the demo keeps one warm replica).
- **Small storage:** daily totals only, never raw sensor streams.
- **Real costs:** database, sign-in, messaging, clinical review and support. Pilot: about US$50–100 a month of Azure for 300 people (estimate).

Beyond movement (designed, not built)

- **Reused unchanged:** day status (DEC-01), who sees what (DEC-03), message tone (DEC-04), reminders (DEC-06).
- **New per behaviour:** safety screen (DEC-02 variant), smaller-step ladder (DEC-05 variant), approved actions with evidence cards, measures, Arabic, clinical review, tests.
- **Sleep:** regular wake time the person chooses; red flags for sleep apnoea, drowsy driving, low mood.
- **Nutrition:** one positive swap a day; no calories or weight targets; red flags for eating-disorder language and low-blood-sugar medicines.
- **Smoking cessation:** only with the partner's cessation service (medication).
- **Effort:** about 4–6 weeks of content and rules per behaviour, clinical review in parallel; the platform is unchanged.

FUTURE HEALTH | A GLOBAL INITIATIVE BY ABU DHABI | H2S HACK2SKILL

Future Health Digital Innovation Challenge

From signals to solutions.

Thank you

Co-creator

Microsoft

Co-creator

inception42

Co-creator

Mohamed bin Zayed University of Artificial Intelligence