

PRODUCT STRATEGY · PRIORITISATION · RESPONSIBLE AI

BERT Product Roadmap & Prioritisation Framework

An evidence-informed plan to strengthen the individual value loop before expanding into collaborative and enterprise use cases.

Nanda Kishore Reddy Mangunuri

Early beta

30 September 2026

Owner: Nanda Kishore Reddy Mangunuri

Product stage: Early beta

Planning approach: Outcome-led, evidence-informed and capacity-aware

Last updated: 30 September 2026

1. Executive decision

BERT should not expand into team dashboards, enterprise administration or a broader behavioural platform until its individual value loop is consistently trusted and completed.

The roadmap therefore follows four outcome horizons:

1. **Now — earn trust and complete the core loop.** Remove authentication, onboarding, recovery and generation failure points.
2. **Next — make the Blueprint meaningfully better.** Improve specificity, contextual relevance, explainability and concision.
3. **Later — turn individual insight into collaborative value.** Strengthen sharing, feedback and small-team workflows.
4. **Explore — validate and scale responsibly.** Build the evidence, controls and platform capabilities required for B2B use.

This is a strategic sequence, not a promise of dates. Movement between horizons depends on release gates and evidence.

2. Product context

2.1 Current product

BERT is a cross-platform behavioural technology product. A user verifies their email, provides profile and birth information, completes an adaptive assessment and receives a personalised Blueprint designed to explain how they tend to work, decide, adapt and collaborate.

The current value loop is:

Discover BERT



Verify and onboard



Complete adaptive assessment



Generate Blueprint



Understand personal work patterns



Give feedback, return or share

2.2 Team and delivery constraints

The beta was delivered by a three-person contributor group: product/engineering, UI/UX and strategy. That makes focus a product requirement. Parallel feature programmes would increase delivery risk and reduce learning quality.

The roadmap assumes:

- one primary engineering owner;
- part-time design and strategic input;
- a shared FastAPI backend serving environment-specific deployments;
- Flutter clients for iOS and Android;
- Azure-hosted data, AI, communication and application services;
- a small beta population before wider acquisition;
- no claim of clinical, diagnostic or employment-selection validity.

3. Evidence informing the roadmap

The first beta produced 27 registered users, 20 completed Blueprints and 12 structured feedback responses.

Signal	Baseline	Product implication
Registered-user completion	74% (20/27)	The core loop works, but recovery and friction still matter
Journey rated easy or very easy	83%	Overall usability is promising
Journey completed within 15 minutes	92%	Time is broadly acceptable, but perceived repetition remains
OTP selected as difficult	42%	Authentication is the clearest journey friction
Birth-data explanation rated unclear	58%	Trust and purpose are not explained adequately
Blueprint moderately or very accurate	92%	The concept has a useful signal worth developing
Blueprint felt generic	67%	Specificity is the largest output-quality problem
Blueprint reported as repetitive	42%	Report synthesis and assessment wording require refinement
Blueprint at least moderately useful at work	100%	Workplace value is plausible but not yet proven longitudinally
Likely or very likely to share	67%	Collaboration is promising after quality improves
Recalled in-app feedback usage	17%	The learning loop is weak or poorly surfaced
Reported a technical issue	0%	Encouraging, but based on only 12 iOS respondents

Evidence limitations

- The sample is small and entirely iOS-based.
- Respondents are early adopters and may be unusually supportive.
- Survey accuracy is perceived accuracy, not scientific validation.
- No controlled comparison, longitudinal outcome study or workplace-impact study exists.
- Several funnel steps are inferable from database state but not yet captured as canonical product events.

The roadmap treats these results as directional signals, not proof of product-market fit.

4. Strategic product outcomes

Outcome A — Trusted activation

Users understand why information is requested, can authenticate reliably and can resume without losing progress.

Outcome B — Reliable value delivery

Every eligible completed assessment produces or retrieves a Blueprint, with a clear status when generation takes time or fails.

Outcome C — Distinctive insight

Blueprints feel specific, concise, context-aware and grounded in evidence rather than generic personality language.

Outcome D — Meaningful use

Users read, revisit, annotate through feedback or share their Blueprint for a real work-related purpose.

Outcome E — Responsible learning and scale

The team can improve the product using privacy-safe telemetry, versioned outputs, human review and explicit validation boundaries.

5. Prioritisation method

5.1 Two-lane model

Not every item should compete in one numerical ranking.

Lane 1 — mandatory gates

Security, privacy, legal compliance, account deletion, data integrity, critical reliability and store-policy requirements are release gates. They are addressed before discretionary growth work, even when their reach is low.

Lane 2 — product opportunities

Discretionary work is scored using a lightweight evidence-adjusted RICE model.

5.2 Evidence-adjusted RICE

Priority score = (Reach × Impact × Confidence) / Effort

Factor	Scale	Interpretation
Reach	1–5	Proportion of the active journey affected
Impact	1–5	Expected effect on core value or a release gate
Confidence	0.5, 0.7, 0.9	Hypothesis, directional evidence or strong direct evidence
Effort	Estimated person-weeks	Relative delivery cost including testing and rollout

Scores support discussion; they do not replace judgement. Dependencies, reversibility, risk and learning value can change the final order.

5.3 Decision principles

1. Fix failure before adding breadth.
2. Improve the core insight before monetising or scaling it.
3. Prefer changes that produce both user value and stronger evidence.
4. Keep deterministic behavioural evidence upstream of generative wording.
5. Do not use engagement mechanics that encourage unnecessary daily use.
6. Do not expose individual-level behavioural information to organisations without explicit consent and access controls.
7. Avoid false precision when sample sizes are small.

6. Prioritised opportunity inventory

Scores are planning estimates, not measured forecasts.

Initiative	Reach	Impact	Confidence	Effort	Score	Final priority
OTP reliability, autofill, paste and resend states	5	4	0.9	2	9.0	Now
Explain birth-data purpose, use and protection	5	4	0.9	2	9.0	Now
Assessment and generation resume recovery	4	5	0.7	3	4.7	Now
Generation status, retry and fallback reliability	4	5	0.7	3	4.7	Now
Canonical journey and report instrumentation	5	4	0.7	3	4.7	Now
Blueprint 2.0: specific, contextual, explainable output	5	5	0.9	8	2.8	Next — strategic bet
Assessment language, progress and repetition review	4	4	0.9	6	2.4	Next
Contextual feedback at evidence-card level	3	4	0.7	3	2.8	Next
Share flow and invited-reader experience	3	4	0.7	4	2.1	Next/Later
Scientific validation programme	3	5	0.7	10	1.1	Later — strategic gate
Team comparison and manager workflows	2	4	0.5	10	0.4	Explore
Enterprise administration and integrations	1	3	0.5	12	0.1	Explore

Blueprint 2.0 and validation score lower because they are substantial investments. They remain strategically necessary: simple scoring would otherwise over-reward small usability changes and underfund differentiation.

7. Now / Next / Later roadmap

NOW — Core-loop trust and reliability

Indicative horizon: 0–8 weeks

Outcome: More suitable users reach a trustworthy Blueprint without avoidable failure or ambiguity.

1. Authentication experience

- Improve OTP delivery messaging, timer and resend behaviour.
- Support full-code paste and platform autofill where available.
- Make error states distinguish invalid, expired, rate-limited and delivery-delayed codes.
- Track request, delivery proxy, first-attempt success, verification and abandonment without logging OTPs or email addresses.

Exit signal: OTP is no longer a leading reported difficulty, and verification conversion can be measured reliably.

2. Birth-data trust and usability

- Explain why birth date, time and place are requested before collection.
- Explain how the information influences the experience and how it is protected.
- Provide clear input alternatives and accessible controls.
- Preserve explicit consent and a path to the privacy and deletion policies.

Exit signal: Fewer than 25% of a sufficiently larger feedback sample rate the explanation unclear. This is a provisional target, not a statistical claim.

3. Interruption and recovery

- Restore the correct onboarding or assessment step after inactivity, logout or app termination.
- Reconcile local navigation with server-owned status.
- Make retries idempotent so repeated taps do not duplicate sessions, responses or reports.
- Provide a recovery path rather than “No assessment found” when resumable state exists.

Exit signal: At least 95% of observed resume attempts reach the next valid state without support intervention.

4. Blueprint generation reliability

- Persist explicit queued, processing, rendered and failed states.
- Show honest progress language rather than an indefinite loading state.
- Add retry/fallback rules and operational alerts.
- Capture generation duration, model/config version and failure category without report prose or PII in logs.

Exit signal: At least 98% generation success in a rolling beta window, with every failure diagnosable by category.

5. Measurement foundation

- Define canonical event names, schemas, owners and prohibited fields.
- Instrument onboarding, OTP, assessment resume, generation and report-section views.
- Version assessment, scoring and report outputs.
- Add the executive journey and reliability views to the admin console.

Exit signal: The team can reproduce the primary funnel and distinguish product drop-off from technical failure.

NEXT — Blueprint differentiation and learning loop

Indicative horizon: 8–16 weeks

Outcome: Users receive a more specific, contextual and explainable Blueprint and can improve it through feedback.

1. Blueprint 2.0

- Introduce a concise Blueprint summary with evidence-linked strength cards.
- Connect each statement to behavioural evidence, context and confidence language.
- Separate stable patterns from situational variation.
- Deduplicate overlapping content before AI rendering.
- Use constrained AI rewriting only after deterministic structure and automated checks.

Exit signals:

- genericness falls from 67% toward a provisional target below 40%;
- repetition falls from 42% toward below 25%;
- perceived accuracy does not regress;
- report generation reliability remains above the release gate.

2. Assessment experience refinement

- Audit repeated and low-information questions using response and item evidence.
- Explain that some related questions test consistency without revealing scoring mechanics.
- Add progress in a way that remains compatible with adaptive length.
- revise unclear or culturally narrow wording using UK English and concrete workplace scenarios;
- test whether fewer questions preserve evidence quality.

Exit signal: Completion and perceived clarity improve without lowering evidence coverage or increasing uncertainty.

3. Contextual feedback

- Ask “sounds like me” or “not quite” at the relevant claim, not only after the report.
- Capture optional context about when a claim does or does not fit.
- Explain how feedback improves future versions.

- Keep feedback versioned against the exact report, evidence model and content schema.

Exit signal: Feedback participation rises while report completion and trust remain stable.

4. Controlled share improvement

- Make the shared Blueprint concise and useful for a colleague or manager.
- Clarify owner consent, link expiry and revocation.
- Track link creation and aggregate consumption without invasive recipient tracking.
- Test whether sharing reflects value rather than social pressure.

Exit signal: Share creation and invited-reader engagement grow without privacy complaints or increased genericness.

LATER — Collaboration and evidence

Indicative horizon: 4–9 months

Outcome: Demonstrate repeatable workplace value beyond individual self-reflection.

1. Small-team pilot

- Recruit a small number of teams with explicit participant consent.
- Frame BERT as a conversation aid, not a selection or performance-ranking tool.
- Test team agreements, collaboration prompts and manager conversation guides.
- Measure whether teams report clearer working preferences and useful conversations.

2. Validation programme

- Define constructs and scoring hypotheses with a qualified psychometric advisor.
- Assess reliability, test–retest behaviour, convergent/divergent evidence and subgroup performance.
- Maintain a documented separation between perceived accuracy and scientific validity.
- Publish limitations and intended-use boundaries.

3. Durable value and retention

- Test later-day Blueprint return, updates after meaningful changes and work-scenario reflection.
- Avoid manufacturing habitual use where periodic use better fits the job.
- Measure Monthly Meaningful Blueprint Users rather than daily active users as the primary value signal.

4. Operational product administration

- Expand the admin console for version comparison, quality review, feedback analysis and support workflows.
- Apply role-based access, audit logging, least privilege and environment separation.
- Keep destructive user-data actions confirmation-based and traceable.

EXPLORE — Responsible B2B platform

Indicative horizon: 9–18 months, conditional on prior evidence

Outcome: Determine whether BERT can support organisations without compromising individual trust.

Potential discovery areas:

- team-level collaboration views with minimum cohort sizes;
- user-controlled invitations and permissioned organisational access;
- manager and facilitator workflows;
- HRIS or collaboration-tool integrations;
- paid pilots and packaging;

- multilingual and cross-cultural assessment research;
- behavioural representations for longitudinal research;
- privacy-preserving aggregate analytics.

These are discovery themes, not committed features.

8. Release structure and gates

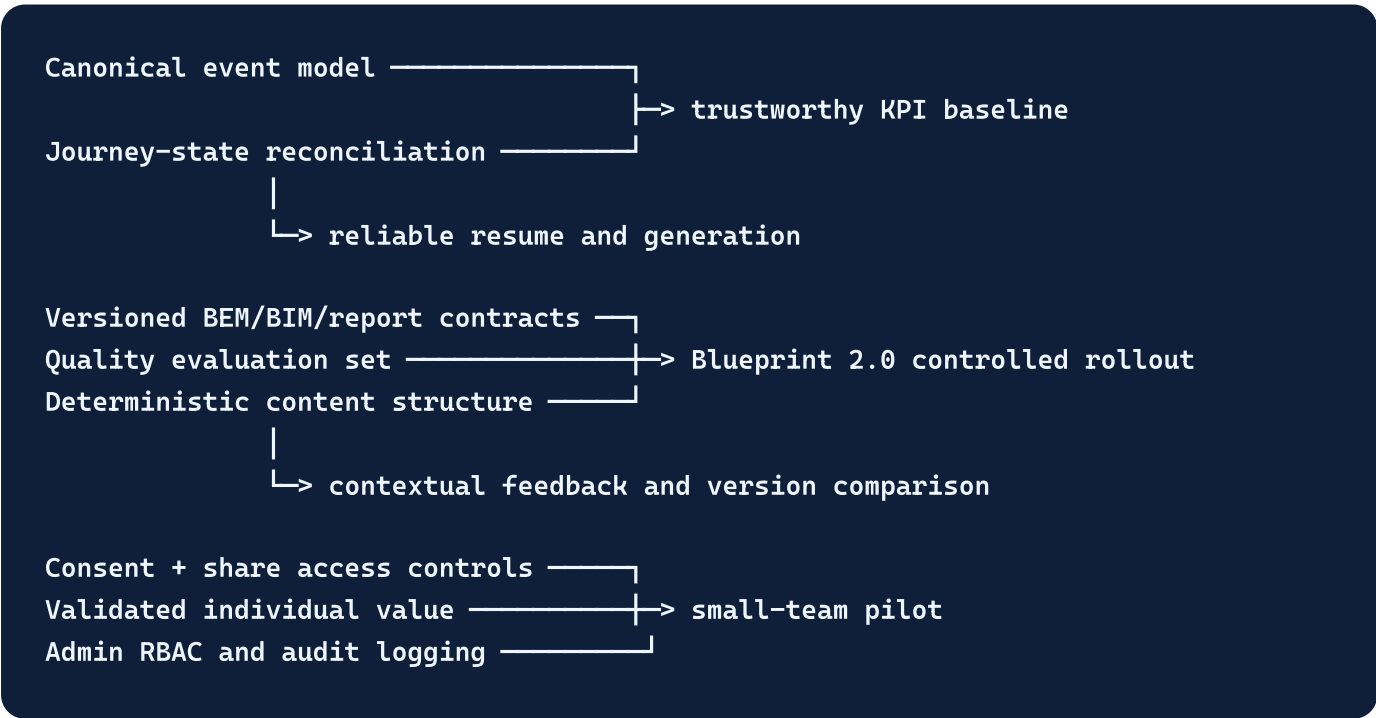
Release	Purpose	Must be true before release
R1 — Trusted journey	Stabilise authentication, onboarding, resume and generation	Critical flows pass UAT; privacy copy is correct; no known data-loss path; failure telemetry exists
R2 — Blueprint 2.0 controlled beta	Test more specific and explainable output	Offline quality set passes; schema and evidence checks pass; version rollback exists; human review completed
R3 — Meaningful-use loop	Improve feedback, return and sharing	Consent and revocation work; value events are measured; no regression in completion or reliability
R4 — Small-team pilot	Test workplace collaboration	Intended use and safeguards documented; role access verified; pilot participants consent; evaluation design agreed

Stop-the-line conditions

A release pauses when any of the following occurs:

- exposure of secrets, OTPs, raw birth information or report content in logs;
- unauthorised access to a Blueprint or user record;
- destructive account action without correct authentication or auditability;
- material scoring or report regression with no rollback path;
- AI output that violates schema, invents unsupported evidence or introduces unsafe claims;
- unexplained reliability degradation in authentication, assessment or generation.

9. Dependency and critical-path map



The critical path runs through instrumentation, reliable state, versioning and quality evaluation. Building team features before these foundations would create scale without trustworthy learning.

10. Capacity allocation

For a lean beta team, a reasonable planning envelope is:

Capacity lane	Share	Purpose
Core outcome work	55%	Roadmap initiatives tied to the active horizon
Reliability, security and privacy	25%	Defects, operational resilience and mandatory gates
Measurement and discovery	15%	Analytics, interviews, quality review and experiments
Unplanned support	5%	Tester support and urgent operational work

This allocation should be reviewed each cycle. A reliability incident temporarily overrides the split.

11. Outcome scorecard

Targets below are hypotheses for the next larger beta cohort and should be recalibrated when sample size and event quality improve.

Outcome	Current signal	Provisional target	Guardrail
Registered-user completion	74%	≥85%	No increase in support contacts
OTP reported as difficult	42%	<20%	Rate limiting and security remain intact
Birth explanation unclear	58%	<25%	Consent remains explicit, not coerced
Generation success	Not yet canonical	≥98%	No unsupported fallback content
Blueprint genericness	67%	<40%	Perceived accuracy does not decline
Blueprint repetition	42%	<25%	Required evidence coverage retained
Feedback participation	17% recalled usage	≥40% observed eligible use	Report completion does not fall
Meaningful Blueprint Use Rate	Not yet measurable	Establish baseline	Privacy-safe event collection only

12. Discovery and experiment plan

During NOW

- Observe five to eight users completing OTP and birth-data onboarding.
- Interview interrupted users about expected resume behaviour.
- Classify every generation failure and recovery route.
- Validate event completeness against database state.

During NEXT

- Run moderated comprehension tests on old versus Blueprint 2.0 examples.
- Blind-rate outputs for specificity, evidence support, repetition and actionable language.
- Use staged rollout or within-user preference testing rather than claiming significance from a tiny A/B test.
- Interview both Blueprint owners and invited readers.

During LATER

- Test a defined team job-to-be-done with a small pilot.
- Predefine success and failure criteria before onboarding organisations.
- Combine behavioural telemetry, structured feedback and qualitative interviews.
- Begin formal validation only with appropriate specialist input and sample planning.

13. What is deliberately not prioritised

Daily-engagement mechanics

BERT is currently an episodic insight product. Streaks, push-notification loops and daily-open goals could increase activity without increasing value.

Large enterprise feature breadth

SSO variants, deep HRIS integrations, complex billing and organisation-wide dashboards are premature before repeatable team value is demonstrated.

Automated employment decisions

BERT should not rank, reject or recommend candidates. The current evidence does not support high-stakes employment use, and such positioning would create substantial ethical, legal and trust risk.

Unconstrained generative personalisation

More AI-generated prose is not the objective. Personalisation must remain bounded by structured behavioural evidence, versioned prompts/schemas and quality checks.

Foundation-model claims

Behavioural representation learning is a long-term research direction, not a current product capability. It requires consented data, validity evidence, governance and a scale of observations the beta does not have.

14. Operating cadence

Weekly delivery review

- progress against the active outcome;
- release risks and blocked dependencies;
- reliability and privacy incidents;
- qualitative evidence from testers;
- decisions required from product, design or engineering.

Fortnightly product review

- funnel movement and data quality;
- report-quality sample review;
- completed discovery and changed assumptions;
- backlog re-ranking using new evidence;
- explicit continue, change, stop or investigate decisions.

Release review

- acceptance and release-gate evidence;

- rollback plan and environment promotion;
- monitoring ownership;
- user communication and support readiness;
- decision log updated after release.

Quarterly strategy review

- whether individual value is strong enough to expand;
- which segment and job-to-be-done show the strongest pull;
- whether validation evidence supports current claims;
- whether team or commercial discovery should accelerate, remain exploratory or stop.

15. Decision log template

Field	Required content
Decision	What was chosen
Problem	User or business problem being addressed
Evidence	Quantitative, qualitative and technical evidence
Alternatives	Options considered
Trade-off	What is delayed, reduced or accepted
Success signal	Observable outcome and window
Guardrail	What must not worsen
Revisit trigger	New evidence that would reopen the decision
Owner and date	Accountable decision-maker and timing

16. Portfolio interpretation

This roadmap demonstrates more than feature sequencing. It shows how an early product can be managed when evidence is promising but limited:

- problems are prioritised before solutions;
- mandatory trust work is separated from opportunity scoring;
- beta feedback is translated into measurable outcomes;
- technical foundations are treated as product dependencies;
- AI quality is governed through evidence, versioning and release gates;
- B2B expansion is conditional on individual value and responsible-use safeguards;
- targets are labelled as hypotheses rather than presented as achieved results.

The central strategic choice is deliberate: **make the individual Blueprint trustworthy, reliable and distinct before using it as the foundation for team or enterprise value.**