Human judgment, made verifiable

Human judgment. Verifiable AI quality.

Voice AI QA, multilingual evaluation, red teaming, and consented speech data with real speakers across Southeast Asia and Myanmar—operated from our Bangkok hub in Japanese and English.

yamada@bku1.com · voice.bku1.com

01 / PROBLEM

Find broadly with machines. Confirm with people.

Automated tests generate failure hypotheses at scale. Accents, self-correction, noise, cultural assumptions, and safety boundaries often become visible only in conversations with real people.

AUTOMATION

Scale

Hypotheses

+

HUMANS

Reality

Evidence

02 / PORTFOLIO

One operation, from testing to data supply.

CORE

Voice AI Human QA

Real second-language speakers test conversation flows on real devices, networks, and environments.

BUNDLE 01

Multilingual AI Model Evaluation

Evaluate low-resource-language outputs for linguistic quality and cultural fit.

BUNDLE 02

Voice AI Red Teaming

Find safety and policy failures through adversarial personas, intent shifts, and conversational ambiguity.

BUNDLE 03

Consented Speech & Conversation Data

Run data collection that can explain who recorded what, under which consent.

03 / CORE

Voice AI Human QA

Real second-language speakers test conversation flows on real devices, networks, and environments.

What we do

01 Scenario and failure-condition design

02 Live calls with accents, noise, hesitation, and self-correction

03 Task completion, repeats, abandonment, and entity-error logging

04 Regression tests after model or prompt changes

How quality is controlled

01 Participant identity verification

02 Project-specific consent records

03 Scenario, device, and environment metadata

04 Independent QA review

04 / BUNDLE 01

Multilingual AI Model Evaluation

Evaluate low-resource-language outputs for linguistic quality and cultural fit.

What we do

01 Rubric and pass/fail design

02 Naturalness, meaning preservation, and cultural-fit scoring

03 Inter-rater agreement checks

04 Error taxonomy and improvement priorities

How quality is controlled

01 Evaluator qualification checks

02 Example-based evaluation guide

03 Double review and disagreement resolution

04 Deliverables with traceable rationale

05 / BUNDLE 02

Voice AI Red Teaming

Find safety and policy failures through adversarial personas, intent shifts, and conversational ambiguity.

What we do

01 Persona-based adversarial scenario design

02 Identity-bypass and unsafe-response tests

03 Conversations with ambiguity, contradiction, and emotion shifts

04 Regression comparison for non-deterministic responses

How quality is controlled

01 Written test authorization and boundaries

02 Reproduction steps and evidence

03 Severity and reproducibility ranking

04 Review summary for safety owners

06 / BUNDLE 03

Consented Speech & Conversation Data

Run data collection that can explain who recorded what, under which consent.

01 Speaker recruitment to specification

02 KYC and project-specific consent

03 Scripted, spontaneous, and conversational recording

04 Audio, transcripts, and speaker/environment metadata

07 / QUALITY PROCESS

Keep the who, why, and what traceable.

01	02	03	04	05
KYC	CONSENT	EXECUTE	QA REVIEW	DELIVER

Identity, consent, execution conditions, review, and deliverables stay linked at project level.

Run complex cross-border operations from start to finish.

Japan × Myanmar

Across two legal systems

10 years in SEA

Yuhei Yamada / since leaving DENSO

BKU runs recruitment, education, compliant procedures, travel, onboarding, and post-employment support across two countries, while expanding its Southeast Asian network from Bangkok.

09 / CONTACT

Tell us what must work, for whom, and under what conditions.

Share the target user, workflow, and cost of failure. We will design the smallest pilot that can answer the question.

Yuhei Yamada

yamada@bku1.com

voice.bku1.com

Languages: Languages across Southeast Asia, including Thai and Vietnamese, plus Burmese, English, and Japanese