GovTech Advisory and Guidance Tool (GTAG)

June 13th 2023

Asami Okahashi, Governance Specialist, Governance GP
Olushola Ibironne Joanne Martins, IT Officer, ITSTI
What is GovTech

GovTech is a whole-of-government approach to public sector modernization that promotes simple, efficient and transparent government with the citizen at the center of reforms.

The GovTech approach emphasizes three aspects of public sector modernization:

- A whole-of-government approach to digital transformation
- Universally accessible citizen-centric services
- Simple and efficient government systems

The GovTech agenda also encompasses:

- Effective use of disruptive technologies (e.g., AI/ML, cloud, IoT),
- Public data platforms (promoting the use of public data by individuals and firms),
- Local GovTech ecosystems (supporting local entrepreneurs and start-ups), and
- Greater use of public private partnerships to draw upon private sector skills, innovations and investments to address public sector challenges.
Assumption / Challenge
Our Journey: Approach GTAG Proof of Value

Discover the Problem

01 Prepare
Framing & Planning

02 Discovery
Exploration & Understanding

03 Interpretation
Research & Findings

Discover the Solution

04 Ideation
Concept Generation & Prototyping

05 Develop
Refinements & Experimentations
Questions:
1. What is your current process for researching / sponsoring / designing a project at country or regional level?
2. What data limitation do you have when preparing your projects / running your projects.
3. What data do you need and where do you get or access the data from?
4. Do you think this tool could be useful in supporting your day-to-day tasks?
5. If so, what sort of information/data do you want to see?
6. What would you like the data to tell you? (insights, validations, etc.)
Current Challenges, Desires, and Opportunities

What we heard from GovTech Experts, TTLs & Development Partners

Our clients would like to have access to information on similar project that have being done in other countries with close similarity to their needs. They need information on investment amount, procurement costs, project approach, lessons-learned.

Time consuming to review very lengthy documents to find the required information.

Current database is updated manually with new projects and extracted from ops portal and through communication with project owners.

There is no corporate mandate to keep this database alive. Hard to convince TTL to keep this database updated because it is time consuming.

Challenges keeping current database and sources up to date. It is a time-consuming manual activity for TTL and teams.

Information that informs our decision to invest through remote sensing of projects opportunities. Tracking Program & project level activities.

Highlight what projects have GovTech related components and who is investing in GovTech related projects.

Eliminate conflicting data from different sources. And ensuring better access to GovTech related data and components.

To support better communication between partner practices with common/similar goals/objectives.
Need of Surfacing Relevant Data & Info from Existing Data Assets

Digital Elements
- Project Cost & Procurement
  - Procurement packages on project
  - Most common expenditure for Low/Mid Income
- Project Info & Comparisons
  - Lessons Learned for a project
  - Comparisons on Lessons Learned on similar projects
- Project Results & Outcomes
  - Important performance indicators
  - Common types of digital components
- GTAG Knowledge Tree
  - Time and Resources
  - Cyber security
  - E-services
  - Data governance

Data Sources
- Operations Portal
- PADs
- ISR/ICR
- ASA
- GovTech Tools
- CPF/CPIA
- IEG
- GTMI Tool
Introduce an AI Enabled Tool to support GovTech Knowledge Advisory Services

GTAG Smart Repository/Dashboard

Searches info on relevant govtech data within WBG project documents (PADs, etc)

Excepted outputs
Govtech relevant/ hidden data within (PADs, ICRs, ISR,ASA) across following categories:
• Project information & comparisons
• Project Results & Outcomes
• Digital Elements

Google Search: provide a wider search based on Keywords or phrases

Guided Search based on frequently asked and answered questions across selected categories.

Data Snapshots such as SDG across projects

Data Sources

PAD/ICR/ISR, country related data, guidance on ASA & Diagnostics QER, Decision Meeting outputs.

Data Sources
(Govtech Database & Ops portal) GTMI Dashboard

User Groups Owner
WBG GovTech Experts

User Groups – Clients & Partners (internal), WBG Task team leads & Teams

Restricted access – out of scope for this prototype

User Groups – Donors & Partners (external)
Proposed Technical Building Blocks

**GTAG USERS**
- Gov Tech expert
- Task Team Leads & Task Teams
- Future Users - Donor & Development Partners (out of scope)

**Frontend**
- **GTAG Dashboard**
  - GTAG Search (AI)
  - AI4DG

**Data Sources**
- Agree data sources & Access levels.
  - PAD
  - ISR
  - ICR
  - IEG
  - ASA
  - GTMI
  - CPF
  - CPIA

**AI Algorithms**
- Development of Algorithms to support
- Data insights and queries.

**API integration**
- Internal and external data sources (operations Portal)
Surfacing Hidden Data
WBG Task Team Lead Scenario

Jane, a TTL has been working at the bank for just over a year. She is about to develop a new PAD for an ECA country, which will contain a GovTech related service. She is hoping that she could gather relevant data from past and ongoing project documentation. She is interested in the following data:

- Project that have GovTech components such as eservices across ECA region
- Information on Similar projects & indicators,
- Review any lesson learned and all associated costs.

How can AI help with surfacing Hidden data?

- **Semantic Search**: Using Azure Cognitive Search with Semantic capability. Embeddings are created based on the corpus (documents) and finding relevant context based on the query.

- **Prompt Engineering for GPT-4**: Using Semantic Search to select the most relevant section of a given document and gpt-4-32k model to generate responses based on user query.

- Hidden Data: Data that lies somewhere in documents, not easy to find

Proof of Value solution

- **Key Features Involved**
- Google wide search (current in POV)
- Guided Search for surfacing data within document (future)
- Data Insights around SDG Goals

List of Projects that contain E-services
This can be limited to selected filters such as regions, year etc.

Curated content from selected project and list of questions to guide the search.
Proposed Solution – Demo
Proposed Solution Architecture

**Front end**

**GTAG**

- Web App
  - GTag Intelligent Search
    - Google like Intelligent Search
  - GPT4DG
    - Guided Search (FAQ)
  - Data Snapshot-SDG
    - Data Snapshot & Correlation related to GovTech & subject interest

**About & Further Information**

- Introduction
- GTMI
- Data Source
- Power BI dashboard
- Others

**Back end**

**AI Model**

- Semantic Search
  - Semantic search denotes search with meaning, as distinguished from lexical search where the search engine looks for literal matches of the query words or variants of them, without understanding the overall meaning of the query

- **GPT-4**
  - Understand and generate natural language. GPT-4 is a large multimodal model (accepting image and text inputs, emitting text outputs) that, while less capable than humans in many real-world scenarios, exhibits human-level performance on various professional and academic benchmarks.

**Database / Cloud Develop**

- PAD
- ISR
- ICR
- IEG
- ASA
- GTMI
- CPF
- CPIA

**Azure Data Storage**