**IVR Presentation**

**1 What will your system do?**

* Our Project is a Multimodal IVR system that enables both voice and GUI oriented interactions
* Uses AI for natural language understanding (NLU), dialogue management (DM), and text-to-speech (TTS)
* Research goal is to find out if the implementation of multimodal features accelerates the process of completing tasks on IVR systems.

A piece of paper with writing on it

AI-generated content may be incorrect.

Figure Bank Statement Use Case

**2 Use Case** (*Make one for each customer journey*)

**Problems with the Current IVR Process**

1. **Redirection:**
   * Users are instructed to go **elsewhere for HMRC payments** instead of handling it within the system.
2. **Manual Input of Account Details:**
   * Users must **type in** their **sort code, account number, and customer ID**.
3. **Repeated Security Checks:**
   * Asking for **random digits from a security number** every time.
4. **Unnecessary Confirmation Prompts:**
   * Asking users whether they want **more info**, set up a **voice password**, etc.
5. **No Smart Call Handling:**
   * System provides **bank balance & overdrafts** even if the user didn’t ask.
6. **Rigid Interaction (Voice-Only):**
   * Users cannot **see their statement**—it is only read out loud.

**Simplify with IVR:**

**1. Smart Account Detection (Faster Logins)**

**Current:**

* User must manually **enter sort code, account number, and customer ID**.
* **Automatic Caller ID Detection**: The system detects the phone number and **links it to the bank account**.

**With Multimodal IVR:**

* If **calling from an unregistered number**, the IVR **displays the login screen on a GUI interface**.

🔹 **User Sees on Screen:**

* “We detected your account. Is this correct? (Yes / No)”
* **Option to manually enter details if needed.**

**2. Secure, Streamlined Authentication**

**Current:**

* Users **manually enter** digits from a security code.

**With Multimodal IVR:**

* **Voice Biometrics:** Users can authenticate using **voice recognition** instead of entering digits.

🔹 **User Sees on Screen:**

* “Authenticate using:
  + Voice Password
  + Enter security digits manually”

**3. Personalized IVR Menu & Faster Options**

**Current:**

* IVR asks why the user is calling and lists examples of questions.

**With Multimodal IVR:**

* **Visual Quick Options:** Instead of waiting for the IVR to read choices, the user **sees** a menu of the most common queries.

🔹 **User Sees on Screen:**

* "Do you need to:
  + **Check bank statement**
  + **Pay HMRC**
  + **Check recent transactions**
  + **Speak to an agent**"

🔹 **User Speaks or Clicks a Button** to proceed.

**4. Faster Bank Statement Retrieval**

**Current:**

* User asks for a bank statement, IVR confirms request, and then reads the statement aloud.

**With Multimodal IVR:**

* Instead of **forcing users to listen** to their statement, they **receive a visual copy** instantly.

🔹 **User Sees on Screen:**

* **“**Your Statement for February:
  + 🔹 Balance: £5,000
  + 🔹 Last 3 Transactions:
    - Coffee: -£5
    - Grocery: -£20
    - Salary Deposit: +£2,000
  + 🔹 [Download Full Statement]”

**3 Research Focus**

**Key Research Areas:**

* **Spoken Language Understanding (SLU)** → How well can IVR interpret user intent?
* **Dialogue Management (DM)** → Can IVR handle follow-up queries?
* **Natural Language Generation (NLG)** → Does IVR produce clear and natural responses?
* **Multimodal Interaction** → How effective is combining voice + GUI options?

**4. System Components & Tools Used**

| Component | Tool / Library |
| --- | --- |
| Speech-to-Text (ASR) | Azure ASR |
| Text-to-Speech (TTS) | Azure TTS |
| Dialogue Management (DM) | Llama AI (Ollama) |
| Multimodal UI (GUI) | Tkinter (desktop) |
| Data Processing | Python, Flask (Backend API) |

**5.** Evaluation Plan

How Will We Evaluate IVR Performance?

User Testing

* Collect real user feedback by observing them complete tasks using the IVR system.
* Track task success rates (e.g., "Did the IVR understand the request correctly?").
* Measure response time & accuracy.
* Will employ Latin Square.

Internal Testing

* Developers will test chats with the IVR model to see whether the system understands information provided to it and also assess its output (How will we be going about this?)

Metrics for Success

* Word Error Rate (WER) for ASR.
* Word Accuracy (generated word is measured against words in the excel file)
* Time taken for IVR system to respond (Quantitative data from observation).
* User Satisfaction Ratings (Qualitative data from questionnaire).

**Data Sources**

* No external data source is used. All data is internally generated and stored into an excel file.