

HR Helper Chatbot

Design choices, key features and development process

Platform: Microsoft Copilot Studio | **Scope:** Internal HR policy assistance

Design Choices

The HR Helper was designed as a focused, internal assistant for Nexara Studios rather than a general-purpose chatbot. Its scope was limited to employee handbook and social media policy questions so that answers would remain relevant and auditable. Claude Sonnet 4.6 was selected for instruction-following, controlled refusals and clear writing. Generative orchestration was retained for flexible language understanding, while web search and ungrounded responses were disabled. This combination allowed the model to reason about a request but prevented it from answering without approved evidence. Microsoft authentication was chosen to keep access within the organisation.

Key Features

The chatbot uses retrieval-augmented generation to search the Nexara Employee Handbook and Social Media Policy, combine relevant information and cite the supporting source. Its instructions require accurate, policy-based answers, explicit uncertainty and referral to HR when information is unavailable. A custom Conflict Report topic provides deterministic dialog management: questions involving conflict, harassment, theft or misconduct receive a fixed confidential escalation message instead of an improvised response. The agent is published through Microsoft Teams and Microsoft 365 Copilot for convenient employee access.

Development Process

I first created a blank agent, defined its name, description and model, and tested the default behaviour to expose unsupported answers. I then disabled web search and ungrounded responses and confirmed that the agent refused questions without evidence. Next, I uploaded the two policy documents, added detailed knowledge-source descriptions and waited for indexing to finish. After inserting role, citation, refusal and escalation instructions, I tested single-document and cross-document questions and inspected the retrieved sources. I built and refined the Conflict Report topic using obvious and disguised misconduct scenarios. Finally, I enabled Microsoft authentication, published the agent, added the Teams and Microsoft 365 Copilot channel, and completed end-to-end and adversarial tests before release.

Project basis: 06. Copilot Studios - HR Helper.docx