

AZURE AI SEARCH AND FOUNDRY IQ

GROUNDING ENTERPRISE AI
ON YOUR OWN DATA



JANNIK REINHARD

MICROSOFT MVP · [JANNIKREINHARD.COM](https://jannikreinhard.com)



jannikreinhard.com

About the Book

Every useful enterprise AI system has the same secret behind it, and it is not the model. It is retrieval — the ability to find the right few paragraphs, out of everything your organization knows, and put them in front of the model at the moment it answers. Get that right and a general model becomes an expert on *your* business. Get it wrong and you have a confident, fluent, well-branded way of being wrong at scale.

This book is about getting it right, on the Microsoft stack. The two tools at its center are **Azure AI Search** — the search and retrieval engine that used to be called Azure Cognitive Search — and **Foundry IQ**, the newer knowledge layer that sits on top of it and gives your agents *agentic retrieval*: the ability to plan a query, run several searches in parallel, rerank, and return a merged, cited answer. Azure AI Search is the muscle; Foundry IQ is the brain that decides how to use it.

I wrote this the way I work: hands-on, with real projects you can rebuild in your own subscription. We start from first principles — what retrieval is and why RAG beats fine-tuning for most enterprise problems — then build a real index end to end: ingest your data, chunk it, vectorize it with an embedding model, design the fields and scoring, and enrich it with skillsets. A dedicated, transparent Part III is reserved for the book's launch partner. Then we query the index properly in Part IV, with full-text, vector, and hybrid search, the semantic ranker, and Foundry IQ's agentic retrieval. Finally Part V takes it to production: grounding Foundry agents and Copilot Studio on your knowledge, securing it with RBAC and document-level permission trimming, and keeping it fast, fresh, and affordable.

You do not need to be a search specialist to start; basic Azure familiarity helps. Whether you are a developer building a RAG application, an architect designing a knowledge platform, or an admin who has been asked to “ground Copilot on our documents,” there is a path through this book for you. I checked the technical content against Microsoft's public documentation in

mid-2026. This area moves fast — verify current availability, prices, limits, and retirement dates before a production decision.

The code

Every code listing in this book — the REST calls, the Python, the index and skillset definitions, the queries — is on GitHub as a runnable file, organized by chapter, so you can clone it and run it against your own service instead of retyping from the page:

<https://github.com/JayRHa/ai-search-foundry-iq-code>



The files are keyless from the start — they authenticate with `DefaultAzureCredential`, so there is no admin key to paste anywhere. Each one names the environment variables it needs and adds the small bit of setup the book explains in prose. If something drifts after printing, that repository is where I keep it current.

About the Author



Jannik Reinhard is a Microsoft MVP for Security and Intune and for AI Platform and Foundry. Today he is Head of AI and Co-Founder/Lead at Epic Fusion Germany GmbH, focusing on Microsoft Foundry, AI agents, retrieval and knowledge platforms, Azure automation, security, and cloud architecture.

Before Epic Fusion, Jannik worked at BASF — one of the world’s largest chemical companies — as a Lead Architect and Technical Lead, where he led technical teams and built AI solutions for a global enterprise of 120,000 users. He was working on enterprise AI before the current hype: he started testing GPT-3 right after its release and was one of the early Azure OpenAI customers. Since then he has tested new models and platform features as they shipped — which is also how he learned to check what still works when a demo becomes a production service.

Jannik holds a stack of Microsoft certifications and is an active Microsoft Certified Trainer. He maintains many public repositories, contributes to open source, blogs at **jannikreinhard.com**, speaks at conferences around the world, and mentors young professionals. This is his third book, after *AI, Automation and Analytics: Transforming Device Management with Intune* and *Microsoft Foundry*. All three share the same idea: explain the topic practically enough that the reader can go and build it.

Copyright

Azure AI Search and Foundry IQ: Grounding Enterprise AI on Your Own Data

Copyright © 2026 by Jannik Reinhard

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the author, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law.

Published by: Jannik Reinhard, 76855 Annweiler, Germany.

First edition published 2026.

Product names, screenshots, prices, and preview/GA status reflect the state of the Microsoft cloud as it was verified in mid-2026. Cloud services change continuously; treat any dated fact as dated and recheck the linked source before you rely on it in production.

Trademarks and affiliation. Microsoft, Azure, Azure AI Search, Microsoft Foundry, Microsoft 365, Copilot, and related names and logos are trademarks of the Microsoft group of companies; all other product and company names mentioned are the trademarks of their respective owners. This is an independent publication and is not affiliated with, authorized, endorsed, or sponsored by Microsoft Corporation. Product screenshots are reproduced for identification and instructional purposes only.

Disclaimer of warranty and liability. The information and code in this book are provided “as is”, without warranty of any kind, express or implied. Every effort has been made to ensure accuracy, but the author makes no guarantee as to the completeness or correctness of the contents and assumes no responsibility for errors or omissions. The scripts, code samples, and configuration steps are provided for educational purposes; always test them in a non-

production environment before use. In no event shall the author be liable for any loss, damage, data loss, service disruption, or other liability arising from the use of, or reliance on, the contents of this book. Verify current details in the official documentation before making any production decision.

Contents

Part I: Foundations: Retrieval and Why It Matters

Chapter 1: Why Retrieval — Grounding, RAG, and the Confident-Wrong Answer.....	12
Chapter 2: Azure AI Search — The Platform.....	25
Chapter 3: Foundry IQ — The Knowledge Layer.....	45

Part II: Building the Index

Chapter 4: Data Sources and Ingestion.....	59
Chapter 5: Chunking and Integrated Vectorization.....	80
Chapter 6: Index Design — Fields, Analyzers, and Scoring.....	97
Chapter 7: Skillsets and Enrichment.....	123

Part III: Sponsor

Sponsor Chapter.....	143
----------------------	-----

Part IV: Querying and Ranking

Chapter 8: Full-Text, Vector, and Hybrid Search.....	145
Chapter 9: The Semantic Ranker and Reranking.....	162
Chapter 10: Agentic Retrieval and Foundry IQ Knowledge Bases.....	179
Chapter 11: Search Explorer, Debugging, and Evaluating Retrieval.....	196

Part V: Production: Grounding, Security, and Operations

Chapter 12: Grounding Foundry Agents and Copilot Studio.....	215
Chapter 13: Security — RBAC, Networking, and Permission Trimming.....	230
Chapter 14: Performance, Scale, and Cost.....	250
Chapter 15: Multimodal, Multilingual, and Advanced Retrieval.....	268
Chapter 16: Diagnosing and Fixing Bad Retrieval.....	285

Chapter 17: Operations, Monitoring, and Keeping the Index Fresh.....	304
--	-----