**E. DATABASE DESIGN, COGNITIVE SEARCH**

1) Azure Cognitive Search vs Traditional Search Engines

Azure Cognitive Search is a cloud-based search-as-a-service that skilfully includes AI improvements into its search pipeline, compared with traditional search engines like Lucene-based engines and web crawlers, which only index text based on keywords. Among its unique features are:

**AI Enrichment:** Cognitive Search uses searching skills to perform tasks like sentiment analysis, language detection, entity recognition (people, places), and OCR (text extraction from images). This makes it possible to search unstructured stuff, like scanned photos, papers, and audio transcripts. Traditional engines are unable to fully understand non-text formats.

**Vector and Semantic Search:** Cognitive Search uses embeddings to provide vector similarity and semantic ranking in addition to accurate match searches. Intentional searches (such as "email engagement" returning "open rates") may also combine keyword and vector searches to improve importance, which is not possible with typical search engines.

**Automatic Search:** Features like fuzzy matching, auto-complete, highlighted search results, geographical searches, and navigation maps can be customized to improve productivity in enterprise search and e-commerce.

**Enterprise Integration & Security:** It is ideal for private data environments because of its customized connectors for Azure Blob, SQL Database, Cosmos DB, and Azure Active Directory for role-based access control; public search engines are unable to store restricted or private data.

**Managed Scalability:** Developers' responsibility of supervising search clusters is reduced by improved availability, multi-region reproduction, and automated scalability.

LIMITATIONS AND MITIGATIONS

* **Field and Document Limits:** Each index can hold up to 1,000 fields (100 on the Basic tier). Prevention methods include combining rarely used fields or dividing large designs into several searches.
* **Paging Cap:** Skip + top searches are only able to get the first 100,000 documents. Use range or key-based paging on a special field designed to page through big result sets as a mitigation.
* **AI Transaction Costs:** An additional fee is applied each time a skill is performed. Mitigation strategies involve planning skill application, using researchers during off-peak hours, or storing enhanced outputs.
* **No Web Crawling:** Does not crawl other websites; requires you to supply all material. Mitigation: Incorporate data using unique pipelines or Azure Data Factory.

2) Importance of Database Normalization in Cloud-based Design

To remove redundancy, maintain data integrity, and stop updating, inserting, and deleting errors, normalization groups relational data into tables. Storage is saved and transactional updates are made easier with normalization in cloud-hosted relational databases (like Azure SQL Database and Azure SQL Managed Instance). These advantages include:

* **Reduced Storage & I/O Costs:** Cloud providers charge for both transactions and storage. Reducing duplicate data lowers expensive write operations and storage use.
* **Better Consistency & Maintainability:** Using a single source of truth makes it simpler to update and modify the design, which is crucial when using Azure's automated deployments and CI/CD pipelines.
* **Improved Transactional Performance:** OLTP techniques reduce performance under high write loads by requiring fewer columns to be locked or scanned in smaller tables.

A denormalized method is supported by next-generation cloud databases like Azure Cosmos DB, which group relevant data into a single JSON document. For instance, addresses and contact information may be embedded in a "Person" document, allowing for quicker, more responsive reading with fewer questions. This design is ideal for a strong workload. Since changing any portion of the document requires changing the entire thing, updates are more difficult. However, cloud-based systems, where customers greatly exceed writes, typically accept this cost.

To conclude in the cloud, normalization is still often important, to avoid redundancy, achieve consistency, and reduce storage costs.

To improve performance, cloud architects will carefully denormalize at the same time. Normalizing basic transactional data and using denormalized views or caches for high-read paths is a sensible strategy. Normalized tables in Azure SQL can be used for updates, whereas Azure Synapse or Cosmos can convert data to a denormalized structure for analytics. There can have an important effect on performance,denormalized structures can speed up searches (typically at the expense of write difficulty and storage), while normalized databases can have slower sophisticated searches through joins. By carefully weighing these trade-offs (and using Azure's indexing, caching, and elasticity features), architects may create systems that expand well without harming data integrity or unnecessarily increasing costs.

**Referencing**

Dev4Side (n.d.) *Azure Cognitive Search: features and use cases* [Online]. Available at: <https://dev4side.com>

(Accessed: 7 May 2025).

Smartbridge (n.d.) *Enhancing enterprise search with Azure Cognitive Search* [Online]. Available at: <https://smartbridge.com>

(Accessed: 7 May 2025).

Microsoft (2024) *What is Azure AI Search?* [Online]. Available at: <https://learn.microsoft.com/en-us/azure/search/search-what-is-azure-search> (Accessed: 8 May 2025).

BytePlus (n.d.) *Azure Cognitive Search pricing and scaling considerations* [Online]. Available at: <https://byteplus.com>

(Accessed: 10 May 2025).

Medium (n.d.) *Semantic and vector search in Azure Cognitive Search* [Online]. Available at: <https://medium.com>

(Accessed: 10 May 2025).

Stack Overflow (2019) ‘Azure search work around $skip limit’, *Stack Overflow* [Online]. Available at: <https://stackoverflow.com/questions/51067478/azure-search-work-around-skip-limit>

(Accessed: 10 May 2025).

Doc. Sitecore (n.d.) *Strategies for managing large Azure Cognitive Search indexes* [Online]. Available at: <https://doc.sitecore.com>

(Accessed: 10 May 2025).

TechTarget (n.d.) *Comparing traditional and AI-powered search engines* [Online]. Available at: <https://techtarget.com>

(Accessed: 13 May 2025).