E. Database Design, Cognitive Search

1. how Azure's Cognitive Search is different from traditional search engines and potential use cases where Cognitive Search would offer a clear advantage. What limitations does it possess, and how can they be mitigated.

Azure Cognitive Search is a managed, cloud-based search service that provides a good search experience over private, diverse content. It differs from traditional search engines (such as Google or Bing) in several ways. (Microsoft Azure Documentation)

Scope of Indexing:

Traditional search engines are designed to crawl and index the publicly accessible web, managing large quantities of unstructured and semi-structured data. In contrast, Azure Cognitive Search is optimized for indexing content specific to a particular application or organization. This content may include structured, semi-structured, or unstructured data and is typically housed within the Azure ecosystem or integrated sources such as Azure Blob Storage, Cosmos DB, or Azure SQL Database. (Microsoft Azure Documentation)

AI Enrichment:

A major strength of Azure Cognitive Search is in its AI enrichment capabilities. By integrating with Azure Cognitive Services, it can perform advanced content processing tasks such as optical character recognition (OCR), sentiment analysis, language detection, key phrase extraction, and custom entity recognition during indexing. These features allow better context-aware search functionality, especially for content types that are difficult to analyse using keyword-based methods. (Microsoft Azure Documentation)

Data and Indexing:

Azure Cognitive Search gives the user a lot of control over how data is organized and made searchable. The user can choose where the data comes from, decides how the search system should be structured, and even customize how the data gets processed. Eg,, you can set up special rules for how different fields are handled, use custom tools to break down text, and add smart AI features to improve search results. In comparison to most search engines where it is done automatically behind the scenes, so users don't have much say in how it works. (Microsoft Azure Documentation)

Integration within the Azure Ecosystem:

Azure Cognitive Search is deeply integrated with other Azure services, allowing organizations to build strong search experiences that agree smoothly with their existing cloud-based infrastructure. This means companies can easily add smart search features to their existing systems without starting from scratch. Since everything is connected within Azure, it's easier to use existing data and AI tools to build powerful and flexible search solutions that can grow as needed. (Microsoft Azure Documentation)

Security and Access Control:

Azure Cognitive Search has strong security features. It works with Azure Active Directory to manage who can access it. This means only the right people can see or search important data, which helps protect valuable information. In comparison, traditional search engines can only search public information and don't have built-in tools to control or protect private content. (Microsoft Azure Documentation)

Potential Use Cases Where Cognitive Search Offers a Clear Advantage:

1. Company Knowledge Bases:

It can organize and index things like internal documents, wikis, and help articles so employees can easily find what they need just by asking in normal language. AI also helps by giving smarter and more helpful results. . (Kleppmann, 2017)

2. Online Shopping Searches:

It makes it easier for customers to find products by letting them filter and sort results in helpful ways. It can even suggest items based on what people are looking for, using smart AI tools like semantic search and recommendations. (Kleppmann, 2017)

3. Searching Legal or Financial Documents:

If a company has tons of important documents, Cognitive Search can help find specific terms, rules, or legal information quickly. It can also use AI to highlight important details in those documents.(Kleppmann, 2017)

4. Searching Media Like Videos and Audio:

It can turn speech in videos into text, detect objects in videos, and even figure

out the mood or tone of the content. This makes it possible to search through videos and audio files based on what's said, shown, or the feeling behind it.(Kleppmann, 2017)

Limitations of Azure Cognitive Search and Mitigation Strategies:

1. Only Searches Indexed Data:
Azure Cognitive Search can only search the data you've specifically added to it. It can't search the entire internet like Google. (Erl et al., 2013)

mitigation:
Make sure you include all the important data you want searchable. If you need info from the web, you can connect it to regular search engines using APIs.

2. Can Be Complicated to Set Up:
Setting it up especially with AI features and custom tools can be tricky and may require some technical know-how. (Erl et al., 2013)

mitigation:
Use the helpful tools and guides in the Azure portal. You can also use Azure SDKs (software development kit) to automate things. Start simple, then improve step by step.

3. It Can Get Expensive:
You may have to pay for indexing, storage, and how often people search—especially with more advanced features. (Erl et al., 2013)

mitigation:
Choose the right plan for your needs and budget. Set smart schedules for updating your data and design searches to use fewer resources. Use features that update only what's changed to save costs.

4. Search Might Be Slow Sometimes(latency):
If your setup is complex or many users are searching at once, results may take a little longer. (Erl et al., 2013)

mitigation:
Improve how your data index is built, use a better service tier with more power, and add caching (storing common results temporarily) to speed things up.

2. Why is database normalisation important in cloud-based database design? Discuss the impact of both normalised and denormalised structures on performance and scalability in a cloud environment like Azure.

Importance of Database Normalization in Cloud-Based Design:

1.Data Integrity:

Just like in regular (on-premises) databases, normalization helps prevent problems when adding, updating, or deleting data. This means the data stays correct and consistent, which is important no matter where the database is hosted. (Silberschatz et al., 2020)

2.Reduced Redundancy:

Even though cloud storage is cheap, avoiding repeated data still helps save space. This can lower the overall storage costs in the cloud. (Silberschatz et al., 2020)

3.Improved Data Consistency:

When each piece of information is stored only once, any updates automatically show up wherever that data is used. This helps keep everything in sync. (Silberschatz et al., 2020)

4.Simplified Data Maintenance:

With a well-organized database, it's easier to make changes to the structure or update data without mistakes. (Silberschatz et al., 2020)

5.Scalability Benefits:

Cloud databases like Azure SQL Database often spread data across multiple servers. Smaller tables from a normalized database are easier to split up and managed more effectively. (Silberschatz et al., 2020)

Impact of Normalized and Deformalized Structures on Performance and Scalability in Azure:

1.Normalized Structures:

Performance:

Normalized databases usually need to use JOINs (is a way to combine data from two or more tables) to pull data from multiple tables. This can slow things down a bit in the cloud because of network delays or how the data is spread out. But Azure has smart tools and indexing to help reduce these slowdowns. (Kleppmann, 2017)

Scalability:

Normalized structures generally lend themselves well to write-heavy workloads and horizontal scaling Because the data is broken into smaller parts (tables), it's easier to spread across servers in the cloud. Updates are faster too since the user usually only need to change a small part of the data. (Kleppmann, 2017)

2.Denormalised Structures:

Performance:

Denormalization involves adding repetitive data to tables to reduce the need for JOINs, which makes reading data much faster. In a cloud environment, this can be beneficial for applications with high read traffic. This is especially helpful in tools like Azure Cosmos DB, which are built for fast reads. (Silberschatz et al., 2020)

Scalability:

Denormalized structures can sometimes complicate write operations and data consistency, mainly in distributed environments. Updates might require modifying repetitive data across multiple tables or documents. This makes it trickier to manage and scale, especially when dealing with lots of servers or big amounts of data. (Silberschatz et al., 2020)

Reference List

Erl, T., Puttini, R., & Mahmood, Z. (2013). *Cloud Computing: Concepts, Technology & Architecture*. Prentice Hall.

Kleppmann, M. (2017). *Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems*. O'Reilly Media.

Microsoft Azure Documentation. (n.d.). Retrieved from [Insert relevant Azure Documentation URL(s) here, e.g., the main Azure documentation page, specific pages on Cognitive Search and Azure SQL Database].

Silberschatz, A., Korth, H. F., & Sudarshan, S. (2020). *Database System Concepts* (7th ed.). McGraw Hill.