**Research and Documentation of Blockchain**

***Objectives:***

1. ***Document the types of data available on each blockchain, such as transaction data, smart contract details, historical data etc.***
2. ***Create a summary of the pros and cons of each blockchain data source.***
3. ***Include information about any limitations or challenges associated with accessing specific data.------------***
4. ***Evaluate the technical requirements for connecting to each blockchain network.***
5. ***Estimate the potential costs and resource requirements for integrating with different data sources.***

***Understanding Blockchain Networks- A)Types of data available on each blockchain Network(Public,Private,Hybrid/Consortium):***

***Public:***

*Just to note, data which its stored-on Public can be used in another the other Blockchain. Essentially think of this as the base line. So all blockchain have these feature below and I will add the additional feature to each blockchain separately in their own category,*

1. *Transaction Data – Financial transactions that may include details about a sender, recipient, amount and timestamp.*
2. *Smart Contract – Contracts that self-execute when the initial terms/agreements are met. These can automate and enforce the execution of predefined rules and actions when certain conditions are met.*
3. *Cryptocurrencies*
4. *Tokenized Assets – Assets like real estate, art, or other physical and digital assets can be tokenized and represented on a blockchain. This allows ownership tracking and transparent/trustful transfer.*
5. *Immutable Records – Once records are added into a public blockchain it cannot be altered or deleted as it would create the whole chain false. Therefore,* ***historical data*** *is preserved all the way.*
6. *Identity and Authentication Data – P.B can store identity related information. Users can prove ownership of their digital identities without relying on a central authority. (This is often done through cryptographic methods). Since added data is Immutable identities cannot be tampered with.*
7. *Supply Chain Info – This could be tracking the production, shipment, and delivery of products in a supply chain.*

***Private: (In addition to Public features, it has these below)***

1. *Sensitive Business Data – Since this is a private you can store sensitive business data, such as financial records, property records (recording and storing property rights), supply chain information, or proprietary information relevant to the participants in the network.*
2. *Regulatory Compliance Data – Records may be store here to prove a company is meeting legal requirements or industry standards.*
3. *Internal Processes and Workflows – Imagine a private blockchain as a special digital communication space for a company. Users can use it to manage projects, share documents, and communicate. It's like having a secure and organized online space for getting work done within the company.*
4. *Operational Data – Details about how much inventory (products or supplies) is available, how well different parts of the company are performing, or any other data.*

***Hybrid and Consortium have a combination of all the above.***

***Understanding Blockchain Networks- B)Advantages and Disadvantages of Blockchain Network types:***

***Public:***

Decentralization: Public blockchains are decentralized, which means that no central authority controls the network. This makes the network more resilient to attacks and censorship.

Transparency: Public blockchains are transparent, and anyone can view all the transactions on the network. This makes it easier to verify the authenticity of transactions and reduces the risk of fraud.

No permission required: Anyone can join a public blockchain network without needing permission from a central authority. This makes it more inclusive and accessible to everyone.

Immutability: Once a transaction is recorded on a public blockchain, it cannot be altered or deleted. This ensures the integrity of the network and reduces the risk of fraud.

Scalability: Public blockchains have limited scalability, which means that they can only handle a limited number of transactions per second. This can lead to slow transaction processing times and high fees during peak usage.

Lack of privacy: Public blockchains are transparent, which means that anyone can view all the transactions on the network. This can be a concern for users who want to keep their transactions private.

Energy consumption: Public blockchains require a lot of computational power to maintain their security and integrity. This leads to high energy consumption, which can be a concern for environmentalists.

***Private:***

*Privacy: Private blockchains offer greater privacy compared to public blockchains. Only authorized users can access the network, and transactions can be kept private.*

*Scalability: Private blockchains can be more scalable than public blockchains, as the number of users is limited. This can lead to faster transaction processing times and lower fees.*

*Control: Private blockchains offer greater control over the network, as the network can be customized to suit the needs of the organization.*

*Centralization: Private blockchains are centralized, as they are controlled by a select group of users. This can lead to concerns about the integrity and security of the network.*

*Limited inclusivity: Private blockchains are restricted to a select group of users, which can limit their inclusivity and accessibility.*

*Cost: Private blockchains can be more expensive to set up and maintain compared to public blockchains.*

*Forked Chains: Since private blockchain users need to trust each other for smooth functioning, which might not always be possible due to conflicting interests among members of the network. This could result in disputes and disagreements leading to issues like forked chains (two separate cryptocurrencies (old and new)).*

***Consortium (Private Blockchain but owned by more than 1):***

*Collaboration: Consortium blockchains enable collaboration among a group of organizations. This collaboration can streamline business processes, reduce friction, and enable faster decision-making.*

*Transparency: High level of transparency compared to traditional centralized systems (non-blockchain). Since changes to a consortium blockchain requires all parties agree on the validity of transactions.*

*Security: Consortium blockchains benefit from the security features of public blockchains while maintaining a level of control among the participating organizations.*

*Governance: Consortium blockchains allow for collective decision-making among the participating organizations. No single organization can exert control on the decision-making process.*

*Complexity: Consortium blockchains are more complex to set up and manage compared to private blockchains. This is because of the need for collaboration between multiple different entities.*

*Centralisation: Consortium blockchains suffer from a higher degree of centralisation than public blockchains. This may result in operational difficulties if rules and regulations are not clearly defined. (Rules and regulations are still being defined by SEC as nodes which the blockchain operates on can be in many different jurisdictions causing issues what is legal or not legal)*

*Limited Access: Consortium blockchains have restricted access to a specific group of organizations. No one outside these organizations is allowed to operate the blockchain.*

*Trust Issues: Consortium blockchains rely on trust among the participating organizations. This can be problematic in practice, as many competitors may not want to operate transparently.*

***Hybrid:***

*Flexibility: Hybrid blockchains provide flexibility by allowing both public and private transactions. It offers the best of both worlds and allows for customisation based on the needs of the application.*

*Security: Public transactions provide immutability and transparency, while private transactions offer privacy and confidentiality. This combination of security features enhances the overall security of hybrid blockchains.*

*Scalability: Public transactions can be used for high-volume, low-security transactions. Meanwhile, private transactions can handle sensitive or high-security transactions.*

*Interoperability: Hybrid blockchains offer interoperability by allowing interaction with other blockchains or external networks. This interoperability enables seamless data sharing and collaboration across different blockchain networks or systems.*

*Complexity: Hybrid blockchains can be more complex to design, implement, and manage compared to single-type blockchains. The integration of public and private components requires careful planning and coordination.*

*Governance Issues: Hybrid blockchains face governance challenges similar to those of public blockchains. Disagreements regarding protocol upgrades or changes can lead to conflicts within the hybrid blockchain network.*

*Interoperability Challenges: Interoperability remains a significant challenge for hybrid blockchains. Ensuring seamless data transfer and communication between different blockchains or external networks requires standardized protocols and interfaces.*

*Integration Challenges: Organizations need to ensure compatibility and smooth data flow between different blockchain networks and traditional IT systems. Overcoming integration challenges and establishing robust interfaces are crucial for successful implementation.*

***Understanding Blockchain Networks- C)What are the technical/resource requirements needed to connect to each Blockchain***

***Technical Requirements Blockchain Network:***

*Node Setup: Node Setup will be needed to be able to connect to a blockchain network as this computer contains a copy of the blockchain ledger.*

*Software Instillation: Instillation of the necessary software will be needed to access/use the blockchain, and this should be outlined by the blockchain provider.*

*Synchronisation: Synchronize your node (computer) with the blockchain network. (This involves installing the entire blockchain to your node, time required will depend on the size of the blockchain and internet connection strength)*

*API Keys: Collect an API key if the blockchain requires it.*

*Programming Packages/Strong Understanding in a language: Choose the correct programming languages/ SDKs (software development kits, all the tools you’d need in one place) to help you interact or develop in the blockchain network.*

*Security Evaluations: Check the security structure of the network (Public or Private) then implement security practises to protect your sensitive data or private keys.*

***Resource Requirements for connecting and using Blockchains:***

*Resource requirements can vary depending on the use case of the block chain, however majority include:*

*Device(Laptop/Computer): To help you access and interact with the blockchain.*

*Internet Connection: A reliable connection is needed to communicate with the blockchain.*

*Server Hardware:*

*Processors: Mining (computational work that network nodes undertake to validate the*  *information contained in blocks) involves solving computational puzzles, so a powerful CPU is*  *beneficial. Some blockchains may favour CPUs with specific characteristics.*

*Graphics Processing Units: In many proof-of-work blockchains, especially those based on hashing algorithms like Bitcoin's SHA-256, GPUs are commonly used for mining due to their parallel processing capabilities.*

*Application-Specific Integrated Circuit (ASIC): In some cases, particularly for widely adopted blockchains, miners may use specialized hardware called ASICs, designed specifically for mining and offering significantly higher efficiency.*

*Cooling Systems: Mining operations can generate a significant amount of heat. Adequate cooling systems are necessary to prevent overheating and ensure the stable operation of the hardware.*

*Power Supplies: Mining hardware consumes a considerable amount of electricity. Ensure a stable power supply and, if possible, consider backup power solutions to prevent disruptions.*

***How can we use blockchain to store data- A) Fundamentals:***

Firstly there are two fundamental ways to store data in a blockchain which is **Off-Chain** or **On-Chain:**

**On-Chain storage:** This is the costly method of storing the data in the blockchain as the data is stored inside each block on the chain. If an attack happens then data can be restored and used.

-Storing small and critical pieces of data, such as token balances or smart contract state variables.

-High data integrity as everything is stored on each block

-Increased security against tampering or loss

**Off-Chain Storage:** In this type of storage only the metadata (just info which describes about your data) is stored in the chain. The entirety is not stored in the chain so if any attack happens then it might not be possible to restore the data. This is a cost-efficient method of data storage.

-Storing large files, media content, or extensive databases where the full content is not needed on the blockchain itself.

-Reduced blockchain bloat allows for easier scalability.

***How can we use blockchain to store data- B) What are the different types of Blockchain Platforms:***

There are many different types of blockchain technologies which can use and here are some of the most common examples:

**Bitcoin:** In bitcoin, the blockchain is a public decentralised network and the data within the structure includes the entire history of all the bitcoin transactions. If one node has an error then it uses thousands of other nodes as a reference point to correct itself. Bitcoin uses infrastructure consists of many chained-Blocks: Transactions are grouped into blocks, and each block contains a reference to the previous block, creating a chain of blocks. This structure ensures the chronological order and immutability of transactions.

To add new blocks to the blockchain is called mining. This is done through the Consensus Mechanism where miners solve complex mathematical problems to validate and add transactions to the blockchain. This process secures the network and incentivizes participants. However, this process uses a substantial amount of energy (not eco friendly) and very time consuming as it takes on average 10min for a transaction to be validated.

**Ethereum:** Ethereum is also completely decentralized and boasts complete support for smart contracts; also ethereum uses a trie data structure to store data. It separates the temporary data from the mined transaction data. The data is added to the transaction trie only when the transaction is confirmed. There are three types of trie in Ethereum for data storage:

State Trie: This global state trie is constantly updated and it contains the key-value pair for every account which exists on the Ethereum Network.

Storage Trie: This stores the contract data. Each Ethereum account has a storage trie and the 256-bit hash value of the storage trie’s root node is stored in the global state trie. Simple words Storage trie is a data structure within Ethereum that organizes and stores contract- specific data for each Ethereum account

Transaction Trie: Each Ethereum block has its own separate transaction tire. A block contains many transactions

**Corda:** This is the open-source blockchain platform that doesn’t use the global broadcast. Corda is built to enable direct transactions between parties involved in a specific business agreement or transaction. It can cut the record-keeping cost while at the same time streamlining the business operations. Here, the communication between the peers can be verified without the need to download the whole data due to the use of graphs and persistent queues.

‘**Global Broadcast’ =** Is used for consensus. This means that every participant in the network receives and validates every transaction. The entire network maintains a copy of the same ledger, providing transparency and ensuring that all participants have a synchronized view of the blockchain's state.

**XRPL (XRP Ledger):** This is a decentralised public blockchain and uses Global broadcast within their operations. Unlike most other blockchains, the XRPL requires no mining and uses negligible energy, key to long-term growth and stability. Furthermore, XRPL is among the fastest transaction speeds as XRP transactions are validated every three to five seconds.

**IBM Blockchain:** Commercial clients who are less risk-averse have found the fastest blockchain platform with IBM Blockchain private, decentralized blockchain network. The developer believes that its ability to connect to the business cloud and legacy systems (old systems) more easily than other decentralized networks presents the greatest prospects. This is crucial for businesses that have invested heavily in older technologies and need to modernize without disrupting their current operations. Also, this can be particularly appealing to enterprises that are less willing to take risks associated with a complete overhaul of their existing technology infrastructure.

**EOSIO:** Regarded as the "top open-source platform" for companies and programmers using blockchain, according to Eosio. It is quick and adaptable, supports scaling, is highly configurable, and gives developers various tools and resources. Due to Eosio's foundation in C++ (general-purpose programming and coding language), both private and public networks can benefit from its "very flexible" environment.

***How can we use blockchain to store data- C) what are the techniques used to store data in a blockchain platform:***

**Hashing:** This is a cost-efficient way of storing the data in the blockchain. In this method, only the hash value of the data is stored in the blockchain. The raw data can be stored in the file system and the hash id of the blockchain will be attached to the raw data.

**TiesDB:** This is an Ethereum-based decentralized application (dApps) to store non-financial data and search through their documents. This allows advanced search and document modifications.

**BigChainDB:** This database allows developers and enterprises to deploy blockchain proofs-of-concept, platforms, and applications with a blockchain database. This offers immutable data storage, built-in-asset support, low latency, powerful query functionality, and high throughput, thus this is a database with blockchain characteristics.

**Distributed database:** Distribute databases like MongoDB, Apache, and Rethink DB can be used to store data. They are quick and versatile, but they are not Byzantine verified. This means any hacker can corrupt the entire information base as all the hubs of the information completely trust one another.

**Decentralized cloud storage:** Decentralized cloud storages allow for the storage of static data where data is not stored on the company server but instead on the devices of the renters. This storage can be used online thus making them fast and efficient but they are costly too. (Renting a DB)

**Interplanetary file system:** This is a blockchain technology that breaks up data into shards (a small part of a whole) and stores them in multiple instances. It is a peer-to-peer solution where the files get downloaded only if the person needs them. Thus, this is the address-dependent storage solution.

***How can we use blockchain to store data- D) Potential Costs for connecting to Blockchains Platforms:***

Factors to consider:

1. **Development and Implementation Costs:** Developing and implementing blockchain solutions often require special knowledge/expertise. Therefore it would be recommended to hire a blockchain developer/consultant to handle this process.
2. **Infrastructure Costs:** Since Blockchains system typically require a decentralised network of nodes to maintain the distributed ledger. Depending on the blockchain platform setting up and maintaining this infrastructure can involve in hardware costs. E.g Servers, Storage, Cooling systems.
3. **Training and Education:** Integrating blockchain technology may require training existing staff or hiring new personnel with blockchain expertise. Providing training programs or enlisting external training services can incur additional costs to ensure the team has the necessary knowledge to manage and utilize the technology effectively.
4. **Security and Auditing:** Blockchain technology can enhance security through its decentralized and immutable nature. However, ensuring the security of the overall system, including data encryption, smart contract audits, and protection against hacking attempts, is crucial. Implementing robust security measures and conducting regular audits can lead to additional costs.
5. **Scalability Considerations:** Scaling blockchain networks can be challenging and may involve additional costs. As the number of transactions increases, the network's capacity and performance need to be maintained. Depending on the chosen blockchain platform, scaling solutions such as sharding or layer-two protocols may require additional investments.
6. **Regulatory and Compliance Costs:** Depending on the industry and jurisdiction, integrating blockchain may introduce additional compliance requirements and regulatory obligations. Ensuring adherence to relevant regulations, such as data privacy laws or anti-money laundering regulations, may involve additional costs for legal consultations, compliance audits, and implementing necessary measures.

***Thoughts:***

The blockchain is a revolutionary system for securely recording transactions, but it still suffers from one major flaw secrecy. The entire point of using a blockchain is distributed trust and the ability to verify transactions in a trustless environment. While this sounds compelling, the reality is that it can only be used to encrypt data with public access. If a user wishes to have complete control over his/her data, he/she will need to keep it local or use an external device.