

White paper: decentralization and AI – breaking the silos

Abstract:

What could the future of the internet look like? Imagine an AI assistant acting as an operating system generating a custom interface for navigating through the internet's data, all stored on users' personal decentralized cloud service.

Three major developments have made this possible: the rise in open-source AI LLMs (Large Language Models), the stabilisation of blockchain technology, and the investment in the metaverse.

These three technologies combined give you a hint as regards what the future holds for the internet, becoming humanity's collective mind space, where humans will be able to synchronize their brains, exchange ideas and thoughts, and coordinate their actions, in an unprecedented pace, birthing a collective hive mind for humanity's super-organism.

Introduction

The internet, often celebrated as a decentralized haven of global connectivity, has undergone a significant transformation over the decades. Today's digital landscape tells a contrasting tale, marked by centralization and dominated by a few tech behemoths. Let's delve deeper into the roots and ramifications of this shift.

The Centralization Creep

When the World Wide Web was in its infancy, it promised a level playing field—a space where anyone could stake their claim, voice their thoughts, and start a revolution from their garage. Over the years, however, this vast, decentralized space of boundless opportunity began to change. The internet's once-diverse ecosystem has increasingly become monopolized by a handful of giants. Facebook, TikTok, YouTube, Twitter, and the like have started to resemble digital kingdoms, each with its realm of influence, leaving little room for smaller entities to thrive. This dominance is felt not just in the social media space, but also in areas like web hosting and data centers. The centralization has birthed an environment where innovation risks being stifled, and the digital world's democratized essence is fading.

The 'Lock-in' Phenomenon

It's a classic case of everyone being somewhere because everyone else is there too. Platforms, especially the likes of Facebook, wield significant power owing to their massive user bases. This results in a peculiar "lock-in effect," a digital inertia, if you will. Users find themselves anchored to a platform not necessarily because of its superior features or privacy policies, but because it's where their friends, family, and connections are. It's akin to attending a party not because you love the venue, but simply because everyone you know is there. This has critical implications for competition. New platforms find it tough to lure users, not necessarily because they lack innovative features, but because breaking the gravitational pull of the dominant platforms is a Herculean task.

Data – The Golden Chains That Bind Us

But there's a deeper, more insidious aspect to this centralization. Our data—every like, share, click, and comment—has become the new oil for these platforms. Users often find themselves shackled to a platform due to the sheer volume of personal data that's been amassed. The very thought of starting anew on a different platform feels daunting. Imagine spending years curating your music playlists, photo albums, and friend lists only to start from scratch elsewhere. But it's not just the

aggregation of data that's the problem. It's its exploitation. These platforms deftly utilize our data, often covertly, to shape our online experiences. Whether it's tailoring advertisements to eerily match our recent conversations or manipulating our feeds to ensure we stay glued for just a few more minutes, our personal data is incessantly used to drive platform revenues and often in ways we might not wholeheartedly approve of.

Centralization has its advantages, of course, with economies of scale and streamlined user experiences. However, it's essential to question at what cost?

Breaking the silos: encouraging signs

In today's interconnected world, the notion of a free, open, and decentralized internet seems almost utopian. However, several parallel developments are hinting at a brighter, more open digital landscape. Let's dive into three of the most promising shifts that might redefine the future of the web.

Regulation: Empowering Users Through Data Portability

One of the most prominent steps towards a decentralized internet has been taken in the form of regulation. The General Data Protection Regulation (GDPR) introduced the concept of data portability, an empowering right that allows users to download their data in a machine-readable format. At first glance, this seems like a substantial leap towards digital freedom. However, the reality is slightly more complex. While the idea of retrieving all your data sounds fantastic, its execution leaves much to be desired. The downloaded data is often unorganized and not directly usable, necessitating manual re-upload into new services. The inherent incompatibility in data structure across platforms makes migration an uphill battle. For data portability to realize its true potential, there's an urgent need for standardized formats and smoother transitions between platforms.

Decentralized Hosting: Charting a New Path Forward

Decentralized hosting isn't a new idea. However, recent technological advances have pushed it from the realm of idealism into a tangible reality. Projects like the InterPlanetary File System (IPFS) offer a fascinating glimpse into decentralized data sharing and storage. But the real game-changer lies in the marriage of blockchain technology with decentralized storage solutions. Enter projects like Sia or Filecoin. By leveraging the immutable and distributed nature of blockchain, these projects are pioneering a new paradigm for hosting, ensuring data security, accessibility, and longevity.

Artificial Intelligence: The Universal Digital Interface

Imagine an AI that not only understands human language but also computer code. Enter the Large Language Models like ChatGPT. These AI systems can effortlessly convert one stream of data into another, bridging the gap between diverse platforms and formats. This capability unlocks a myriad of possibilities, from translating user preferences across platforms to potentially birthing an AI operating system. Such a system could plug into all user accounts, filtering through data and presenting it via a custom-built, real-time user interface. The implications? Users might never need to navigate a platform's website or its native interface. Instead, an intelligent AI could extract relevant content from all services, molding it into a customized experience. Moreover, an AI operating system could broadcast a user's input across multiple platforms, breaking down silos like never before.

Future vision:

The internet has always been a fluid entity, ever-shifting, and continuously evolving. But have you ever envisioned a future where every byte of data is under the strict dominion of its user? Welcome to the future of the internet, a user-centric paradigm shift that's redefining our interaction with the digital realm.

The Age of Personal Data Sovereignty

Picture an internet where you hold the reins of all your data, be it on a decentralized cloud powered by blockchain or a home-based server. No more data resting on third-party platforms. Instead, cryptographic assets pay for storage, ensuring it remains encrypted and utterly private. Whenever you upload, a blockchain entry grants specific services access, based on your privacy settings. The data is always yours; platforms merely access it based on permissions you grant.

Rise of the Personal AI Assistant

In this digital utopia, the spotlight is not on the platforms but on the individual user. Picture an AI, a dynamic blend of an operating system and a personal aide, which scrapes and reorganizes data from myriad sources, creating a real-time user interface tailor-made for you. Whether sharing a memory or voicing an opinion, it's your AI assistant that takes the lead. It streamlines data dissemination across platforms as per your directives. No more juggling apps or remembering passwords. Your AI is the bridge between you and the digital world.

The Decline of Traditional Platforms

In such an evolved ecosystem, traditional platforms become obsolete. Your connections, relationships, and even casual interactions get etched on the blockchain. Why toggle between platforms when your AI assistant can generate a unique interface for each social circle? These AI aides, leveraging their deep learning capacities, filter and present information that aligns with your interests. From news articles to event suggestions, everything is curated just for you. The term 'internet surfing' becomes archaic. Instead, your AI sifts through the vast web, delivering precisely what you seek.

An AI-Constructed Digital Ecosystem

The innovation doesn't stop there. Personal AI assistants evolve to play matchmaker, aligning their users based on mutual interests or needs. Imagine signaling your AI about a job opportunity, and it networks with a newly designed AI-constructed internet. This network optimally coordinates and identifies synergies between users' desires. The internet no longer remains a chaotic web. AIs meticulously structure it, creating intricate virtual pathways linking data across individual clouds.

The future of the internet is not just about faster speeds or broader coverage. It's about restructuring the very fabric of online interactions, making it user-centric, private, and tailored. The lines between man and machine blur, but in this dance, it's the user that leads, and the AI that follows. The dawn of the personalized internet age is upon us, and it promises an era of unparalleled customization and control.

Path to the future:

How do we get from a predominantly centralized internet to a decentralized AI powered one? Here is a detailed roadmap of the various phases of this transition and the necessary prerequisites.

Phase 0: Prototyping

- Develop a simple AI assistant prototype which is capable of scraping data from a website and feed it into a custom built user interface.

Phase 1: Pre-Transition & Data Control

1. GDPR Integration:

- Implement data retrieval protocols to exercise GDPR's 'right to data portability'.
- Ensure all user data, including metadata, is downloadable in a comprehensive package.

2. Open-Source AI Development:

- Begin the design of an AI assistant model.
- Integrate conversion modules for transforming the downloaded user data into SQL or JSON formats.
- Adopt open standards to ensure compatibility with future decentralized frameworks.

Phase 2: Data Migration & AI Training

3. Decentralized Cloud Setup:

- Design and set up personal decentralized clouds with encryption and redundancy measures.
- Implement secure protocols for data uploads.

4. AI Training Mechanism:

- Construct a robust AI training pipeline.
- Feed the AI with articles, personal interests, and other user-specific data.
- Design a retraining mechanism that prioritizes new data, ensuring AI remains up-to-date.

Phase 3: Data Aggregation & AI Integration

5. Aggregation Module Design:

- Develop modules for AI to access and aggregate data from diverse online sources like emails, forums, and video platforms.
- Ensure secure protocols for data scraping, ensuring privacy and data integrity.

6. AI Device Integration:

- Develop a light-weight client app for Android and other major platforms.
- Construct customizable interface components, such as tabs and notification systems.
- Implement a data sync mechanism for real-time updates.

Phase 4: Distribution and Social Connectivity

7. Data Deployment Mechanism:

- Design a content distribution module for the AI.
- Implement multi-platform deployment with the option for simultaneous uploads to platforms like Vimeo, YouTube, and Instagram.

8. Blockchain Integration for Social Bonds:

- Design a blockchain system to map and store user's social connections.
- Integrate AI to interpret and use blockchain data for custom interface generation.

Phase 5: Native Online Space Creation & AI Augmentation

9. Design of Native Online Spaces:

- Develop a framework for users to establish their personalized online spaces using AI and blockchain.
- Facilitate user-to-user interactions without the need for centralized platforms.

10. AI Collaboration and Plugin Integration:

- Construct an open-source API for AI plugins.
- Facilitate community-driven enhancement projects, such as improved sorting algorithms and custom interfaces.

Final Phase: Continuous Improvement, Future Integration and Monetization

11. Feedback Loop:

- Establish a system for users to provide feedback.
- Use AI to analyze feedback and suggest areas of improvement.

12. Future-Proofing:

- Continually monitor emerging technologies and integrate beneficial ones into the system.
- Ensure the decentralized internet model remains scalable, secure, and user-friendly.

13. Monetization Through Microtransactions:

- Integrate public blockchains to facilitate secure and transparent microtransactions.
- Develop a module within the AI to handle instant payments using crypto-assets for services or content access.
- Design a flexible tariff system where users can set or negotiate micropayment rates for the content or services they offer.
- Ensure compatibility with a range of crypto-assets and continually update based on market trends and user preferences.

This roadmap is a strategic blueprint to transition from a centralized to a decentralized internet, focusing on technicalities at every phase. Regular evaluations at each phase will ensure the transition is smooth and achieves its goal of a user-centric digital realm.

Prototype development:

1. Setting Up the Development Environment:

- We will use Python as a development environment and the pip (Python's package installer) to install necessary libraries: beautifulsoup4, requests, and flask.

2. Selecting a News Website:

- For this prototype, we will select a straightforward news website, e.g., BBC News, while respecting the website's robots.txt file, which defines scraping permissions.

3. Web Scraping:

- We will create a Python script to scrape the news headlines from the selected website.

4. Integrate Summarization AI:

- We will create a function that will utilize the transformers library from Hugging Face to summarize text.

5. Displaying the Contents:

- We will modify the Flask app to display both the article content and its summary.
- We will update the index.html file to display the article and its summary.

6. Run the Updated Prototype:

- We will execute the Flask script and navigate to <http://127.0.0.1:5000/> to see the scraped article and its AI-generated summary.

Additional Considerations:

Model Performance: The BART model for summarization, especially the facebook/bart-large-cnn variant, is designed for news articles and is one of the best open-source options available. However, it's still not perfect. The summary quality might vary based on the article's complexity.

Loading Time: Summarization models can be sizable, leading to increased loading times when the server starts. We will consider loading the model once during the server's initialization rather than every time we summarize to reduce waiting times.

Optimization: For a production environment, we will need to optimize further, such as by running the AI model on a dedicated server or using a GPU for faster computations.

Conclusion:

Given the meteoric rise of AI and its use cases, it is clear that an AI intermediated internet is the future. However, this can take many forms. Existing silos and centralized services can become even more centralized, trying to lock in users and their precious data via rolling out proprietary opaque AI assistants, under the control of existing mega-corporations. Under such a scenario, it is dubious whose interests these AI assistants will be designed to serve: the interests of the user, or the interests of the shareholders of those giant companies. Only by investing in open source AI assistants can we secure a rich, diversified and open AI intermediated internet, serving the interests of users first and foremost.