

SN13 - Data Universe

Data Universe (SN13) est un sous-réseau ('subnet') au sein de l'écosystème Bittensor, spécialisé dans la collecte, le stockage et l'analyse de vastes quantités de données provenant de diverses sources, principalement des réseaux sociaux tels que X (anciennement Twitter), Reddit et les transcriptions YouTube. Son ambition est de créer le plus grand ensemble de données open-source de médias sociaux, optimisé pour être facilement accessible et utilisable par d'autres sous-réseaux et entreprises.

L'objectif fondamental de Data Universe est de résoudre le problème de fragmentation et d'accès aux données en temps réel nécessaires à l'entraînement et à la mise à jour continue des modèles d'intelligence artificielle, notamment les grands modèles linguistiques (LLM) et les algorithmes d'analyse de sentiment. En opérant de manière décentralisée, le projet distribue la charge de travail de collecte de données à travers un réseau de 'mineurs'. Ces derniers extraient et stockent les données, qui sont ensuite rendues interrogeables par les 'validators'. Cette architecture garantit non seulement la scalabilité, mais aussi la résilience des données collectées, potentiellement jusqu'à 50 pétaoctets de données réparties sur 200 mineurs.

Le token SN13 joue un rôle central dans cet écosystème. Bien que les détails précis de son utilité et de sa tokenomics ne soient pas exhaustivement décrits dans les sources, il est implicite qu'il sert de mécanisme d'incitation et de récompense pour les 'mineurs' qui collectent des données 'désirables' (basées sur la source, la catégorie, la fraîcheur et l'unicité des données). Il est également probable qu'il soit utilisé pour accéder aux services de collecte et d'analyse de données proposés par le réseau, potentiellement via une API. Des plateformes comme Subnet Tokens permettent l'échange de SN13, indiquant son rôle d'actif négociable au sein de l'écosystème.

Les cas d'usage de Data Universe sont multiples. Ils incluent la réduction des coûts de 'scraping' de données par rapport aux solutions centralisées, l'aide à la recherche de marché en collectant des informations sur les sujets d'actualité, l'amélioration de la précision des outils prédictifs grâce à des données fraîches, et la compréhension des sentiments en ligne mondiaux pour informer les stratégies d'entreprise. Des projets spécifiques au sein de Bittensor, comme le Subnet 44 (Score) pour les analyses sportives et le Subnet 64 (Chutes)

via Squad.ai pour la prévision de performance d'actifs, peuvent potentiellement utiliser les données fournies par Data Universe.

Le projet est développé et géré par Macrocosmos.ai, qui opère le subnet SN13 au sein de l'écosystème Bittensor. Macrocosmos s'efforce d'améliorer continuellement le subnet en ajoutant de nouvelles sources de données, en développant une API interrogeable et en créant des ensembles de données open-source. L'architecture est conçue pour être robuste, compétitive et hautement évolutive, illustrant le potentiel de conception décentralisée de Bittensor. Des outils comme "Gravity" sont mentionnés pour diriger les 'miners' vers des sujets spécifiques en temps réel, et "Nebula" pour la visualisation des données.

En résumé, Data Universe (SN13) se positionne comme une infrastructure de données décentralisée essentielle pour l'avancement de l'IA et de l'analyse de données. Il fournit la matière première (données) de manière efficace et évolutive, ouvrant la voie à des applications plus intelligentes et mieux informées au sein et en dehors de l'écosystème Bittensor.