Extra Labs FAQ

Last update 02.25.2024

Please keep in mind that the ideas and conceptions indicated in this document are likely to evolve in the future, and may not be entirely up to date with Extra Labs' current opinions.
The best way to clarify a point is to contact Charlie Durand, CEO of Extra Labs (charlie@extralabs.xyz, +33625710869, Telegram @CharlieExtra).

Market references and specific knowledge

1. Introduction

1.1. What do we do?

We are building a decentralized protocol called CIRCUM, that incentivizes owners of geospatial sensors to pool their data. The goal is to create a large-scale, near-real-time data source for 3D maps.

The owners of geospatial sensors, which can include autonomous vehicle operators, telecom companies or surveying companies already collect this data. They could pool it if they had a financial incentive to do so. The protocol provides this incentive through a token that distributes value between data consumers and resource providers.

The collected data consists of point clouds, obtained using LiDAR sensors or photogrammetric processing. This data can be reconstructed and rendered in any 3D engine. The resulting 3D map has various applications, including urban planification and smart city, construction, telecoms, autonomous mobility, simulation, gaming and other consumer use cases.

1.2. What is the problem we are solving?

Local governments and companies struggle with outdated and incomplete 3D map data because the acquisition process is very expensive.

Local governments fund large-scale aerial acquisition programs for tens of millions of dollars and the data gets obsolete very fast. For example, the french government has launched a 5-years long 3D modeling program of the whole country for 60M€ (LiDAR HD program). By the time the acquisition is complete, part of the 3D model will already be 5 years old. It's too old to be efficiently used for urban planification and management of land use cases (monitoring of soil artificialization, deforestation, flooding, project insertion, heat island…).

A lot of companies use 3D maps/models for building games, improving user experience of their app, planning an event or construction work, simulating the effect of real world phenomenon (digital twinning) or supporting autonomous vehicles. These companies often have to collect their own data or rely on solutions like Google Maps or Open Street Map. Unfortunately the latter run into the same issues of obsolescence. Google Maps also uses low-flying aircraft for its data acquisition which is very costly. Their data is routinely +8 years old outside of the main urban centers. It is also not very detailed and looks blocky, which is not suitable for designing immersive interfaces (see Zenly experience with this issue in the part Market reference).

Our CEO encountered these issues while working at her first company and trying to implement a 3D map into a mobile application. We believe that in 2024, this implementation should give much better results and that is the origin of Extra Labs.

1.3. Why is now the best time to solve this problem?

In the early 2000s, Google established a monopoly in the field of 2D maps, as they were the only entity capable of making massive investments to collect large-scale data and build processing capabilities. This monopoly, combined with their substantial investments, inevitably resulted in increased prices, such as the 1600% price hike of Google Maps API in 2018.

Today, the race for 3D maps has begun, but the landscape has changed. Sensors are now widely distributed worldwide, and a vast amount of data is already collected. This means that there is no longer a need for massive data collection investments. Additionally, the emergence of distributed ledger technology like blockchain enables the aggregation of latent capacities of these sensors through tokenization, facilitating their collaboration. These factors create a situation where 3D maps can be produced and maintained at a low cost. Furthermore, this approach allows for significant improvements in level of detail and recency by directly connecting to on-the-field sensors and leveraging the “constantly expanding” feature of permissionless systems.

1.4. Who are the co-founders?

Our founding team is made up of 3 people, here are our detailed biographies:

Charlie Durand, Chief Executive Officer, 26, French, Paris-based.

Charlie holds a degree in cultural history from the Sorbonne and another degree in business management from EMLyon Business School. In 2019, she founded her first company, Karacal, and assumed the position of CEO. Karacal is a collection of apps that utilize positioning data to deliver content about points of interest to users. Karacal didn’t raise funds, shifted to a service business model and achieved profitability through clients such as Renault and Radio France.

The transition to a service business model prompted Charlie to leave Karacal in the summer of 2022, with the support of her co-founders. Her departure was also motivated by the idea of establishing a new company focused on 3D maps. Indeed the Karacal team had attempted to implement a 3D map on their mobile app, but struggled to find a solution that offered sufficient detail and immersion.

Charlie excels in designing functional systems, such as protocols, tokenomics, community programs, and business models. She has gained experience in running a business and has led sales activities during her time at Karacal. She is experienced in working with and selling to public institutions. Since 2021, she has been actively involved in the blockchain ecosystem.

Laurent Caraffa, Chief Research Officer, 38, French, Paris-based

Laurent Caraffa completed a Computer Science degree at the University of Nice Sophia Antipolis, followed by a PhD in 3D surface reconstruction at the University Pierre and Marie Curie. Since 2014, he has been a researcher at the French Institute of Geographic and Forestry Information (IGN), specializing in photogrammetry, geometric algorithms and deep learning applied to 3D surface reconstruction challenges. Laurent supervises numerous theses, teaches, and publishes articles on these subjects.

In addition to his research, Laurent has been active in the blockchain ecosystem for several years. As a result, he was present at the meeting where Charlie presented the concept of Extra to the directors of IGN in the fall of 2022. Enthusiastic, he joined the team the same day. The co-founders decided that Laurent would serve as the Chief Research Officer, leveraging his position at IGN to access the institute's vast resources.

Subsequently, the company Extra S.A.S was founded in January 2023.

Dhruv Malik, Chief Technical Officer, 29, Indian, Paris-based

Born and educated in India, Dhruv Malik moved to Paris in 2018 to pursue a master's degree in cybersecurity and cryptography at the prestigious Polytechnic school. After completing his education, he gained experience working for fast-growing companies in the crypto space. In the spring of 2023, he joined the newly formed Extra team, initially as an employee, before eventually becoming a cofounder.

Dhruv is actively involved in the open-source community and frequently contributes to projects. He also participates in hackathons and was a key member of the Extra team that won one of the Protocol Labs prizes at the 2023 EthGlobal Hackathon.

Currently, the company's capital is distributed as follows, with shareholder agreements in place. It is anticipated that 15% of the company will be allocated as compensation for future employees over the next 5 years.

Charlie Durand	52%
Laurent Caraffa	23%
Dhruv Malik	20%
Hanna Feyler	5%

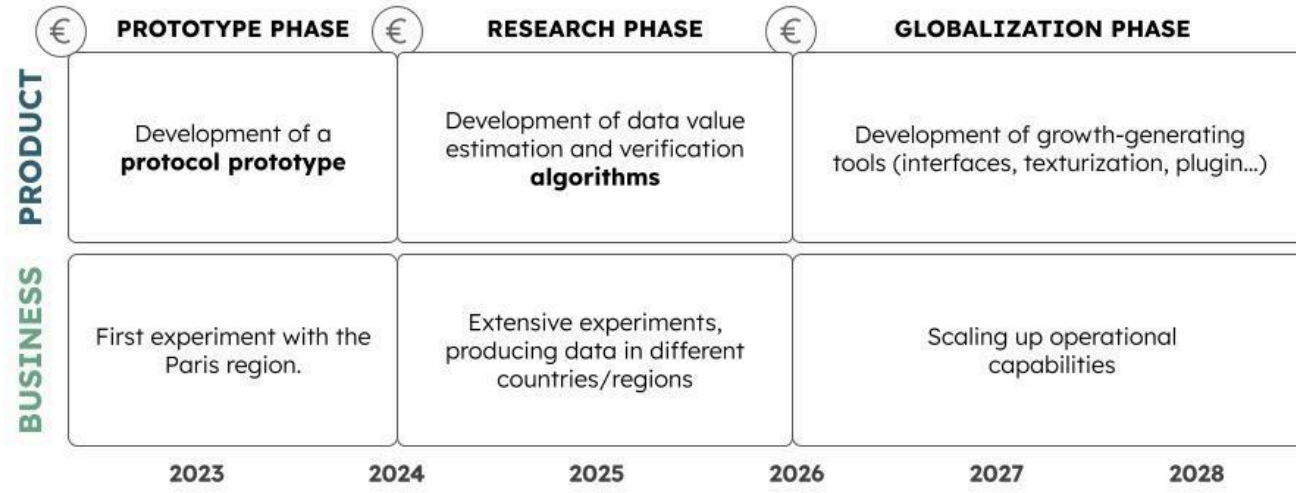
Hanna Feyler is co-founder of Charlie Durand's first company Karacal, and received shares in compensation for Karacal's contribution in designing Extra.

1.5. Where are we right now?

We are about to complete the development of a first prototype of our protocol, which will be tested in the coming months in a paying experiment with the Paris region. More precisely we:

- Have deployed the compute-over-data infrastructure that will support the protocol by enabling storage and processing of the data. This first phase of development was successfully tested through a first contract for a drone acquisition company (future contributor of the protocol). We also started deploying geospatial processing pipelines on this infrastructure.
- Have produced smart contracts to link contribution of the data (stored on IPFS) with emission of a token. We have also deployed an account abstraction mechanism to make the contribution of the data easier.

Here is our roadmap



Following the deployment of our first prototype, we plan to spend the next two years running profitable experiments with this prototype at different scales in different countries and regions, while developing the algorithms needed to scale the protocol globally (see 2.8.).

2. Product

Our protocol CIRCUM serves as a catalyst for pooling different resources such as geospatial data, computing power, storage capacity, and 3D surface reconstruction algorithms. These resources are essential for the protocol to fulfill its mission of generating detailed and up-to-date data for 3D maps.

The protocol consists of various mechanisms that attract resources, validate them, process them, and enable easy access and consumption of data for the end-user.

To facilitate the explanation of CIRCUM, we have divided it into four layers based on the four roles mentioned above: value exchange layer, trust layer, surface reconstruction layer, complexity abstraction layer.

2.1. Why are we using blockchain?

Using blockchain to coordinate resource providers and data consumers enables us to ensure a smooth service while avoiding high CapEx. The token incentivizes owners of geospatial sensors to share their data, even when there are no immediate data consumers. This helps CIRCUM build a dense and coherent capacity to fulfill requests on-demand. By linking the token value to the actual usage of the protocol, we avoid the need to invest in purchasing data without certainty about their resale value.

The token also helps solve the chicken and egg problem. Data providers can join the market to acquire tokens, understanding that they may not immediately cash out on them. In exchange, they expect the token price to increase over time. Additionally, the proof-of-stake mechanism described in the "Trust layer" section provides liquidity to early token holders.

2.2. What blockchain are we using?

We currently use EVM-based L2s (Polygon) because of the advantages they offer:

- Taking into consideration the trilemma issues (Stable performance from genesis, Scalability in terms of high TPS while keeping txn costs low and security with strong proof guarantees on mainnet with concepts like Validium).
- Strong support of all major components of the web3 infrastructure (Oracles for on-chain data acquisition of, smart account wallets for abstracting the web3 workflows etc).
- With their further ascent towards building an aggregated layer of modular blockchains (as explained in their thesis), there will be significant added value in building an application-specific chain to fit the requirements of our protocol's use case (i.e. having a higher DA of point cloud data processed, a unique and customized staking mechanism, etc.).

2.3. What format of data are we accepting in the protocol? From which sensors?

For the time being, our protocol will only accept data in the form of point clouds (.las or .laz). This format can be produced using both LiDAR and photogrammetric techniques, which account for the majority of 3D mapping efforts. We will accept point clouds containing RGB data or classification, but we will not include this information in the protocol output for now.

We have chosen to focus on the point cloud because it protects privacy, is widely used by geospatial players and will enable our data comparison research (see 2.4 and 2.5) to focus on a single format.

LiDAR requires a dedicated sensor, which can be costly. The cheapest 3D LiDAR equipment for our use case starts at around $1000, but the cost can be very high for high-performance equipment. However a lot of companies already use these sensors for data acquisition and autonomous mobility.
Photogrammetry is a technique performed on videos or photos, and is therefore less respectful of privacy and less precise. It is, however, an interesting alternative for reducing costs.

We won't be building our own equipment, but we will be able to recommend certain sensors based on the best value for money for contributors.

2.4. How is the value distributed between resource providers and data consumers (aka value distribution layer)?

This layer of the protocol ensures the fair distribution of the value paid by the data consumers to the resource providers (data, storage, compute and algorithms).

Each resource provider is rewarded with tokens in proportion to their work. These tokens serve as keys that grant access to the data produced by CIRCUM. They are fungible, meaning that each token has the same value as any other, regardless of the actions that led to its mining.

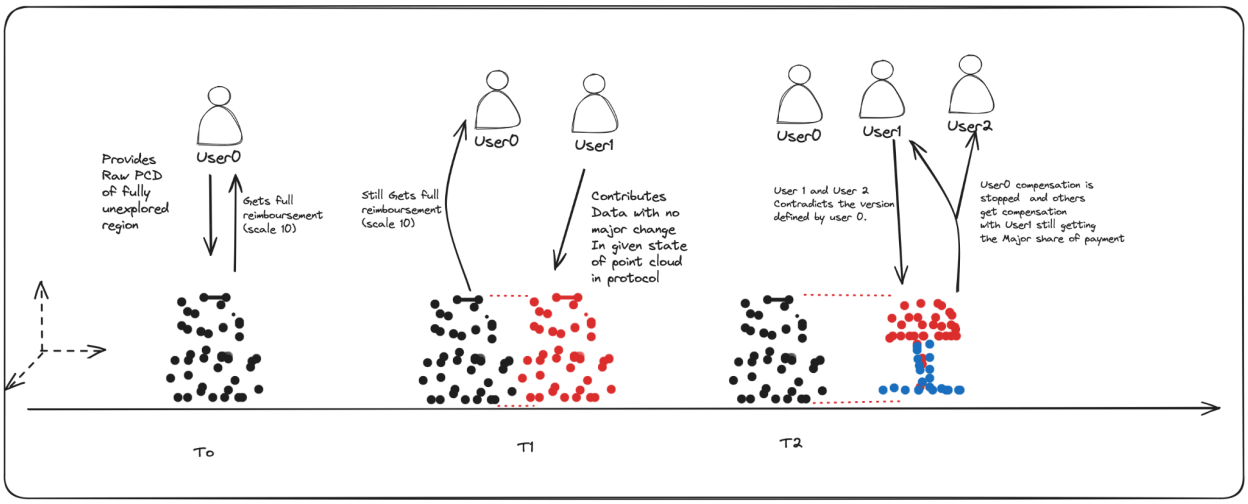
Here is the process for data provider:

1. A data provider contributes a data set.

2. The data set is compared to the existing data state for a given geometric area in the protocol. This comparison provides an relevance grade of the new dataset by finding data freshness. For example, a dataset in a location never mapped previously would be graded a 10 and a dataset contributed twice would be graded 0 the second time.

3. The protocol follows an epoch cycle, the duration of which has not yet been fixed. The map will be updated at each epoch to determine the “current state of the map”. The data that makes it into each new current state of the map is rewarded by tokens, based on its relevance grade and its volume.

For example as shown in the diagram:



1. A contributor provides a point cloud dataset of a street from a terrestrial view, at epoch 0.

2. He gets a relevance grade of 10 because he provided the first data there and thus he gets the full amount of token for its original contribution.

3. During the epoch between T0 and T1, someone proposed a dataset of the same location. The location and nature of contribution hasn't changed, the first contributor keeps earning the full amount of token for its original contribution.

4. During the epoch between T1 and T2, something changed in the neighborhood. Someone else provides data for this change. The first contributor gets a lower relevance grade and gets less token at the end of the epoch.

This system is designed to incentivize contributors to be honest in order to keep collecting tokens for a contribution for a long time. It also incentivizes contributors to look for less well mapped locations, ensuring a better completion of the map.

By default, the amount of token rewarding a contribution is the same, regardless of the location of this contribution. However, we could implement a bounty system to speed up the mapping of certain areas or the completion of the map.

The process is simpler for others resource providers:

1. A resource provider performs a “task”, in providing storage, compute or processing through algorithms. We’ll define what constitutes a task for each type of support resource provider.

2. This task is rewarded by a fixed amount of token.

To access the data, data consumers need to purchase tokens from a resource provider. Once tokens have granted access to the data, they are burned, ensuring a sustainable circulating supply.

This layer will be addressed mainly in our phase of development 1 and 2.

2.5. How do we ensure that the data is accurate (aka trust layer)?

The second layer is also the most theoretical one atm. As a permissionless system, CIRCUM cannot rely on the honesty of its geospatial data providers to maintain an accurate map. These providers may share false, outdated, or low-quality data for various reasons. To address this issue, we aim to align the economic interests of all participants in ensuring the trustworthiness of the data.

To achieve this, we propose using a Proof-of-Stake mechanism combined with a Reputation Score. Contributors are required to stake tokens before they can start submitting data. If their data is frequently contradicted by other contributors or shows others signs of low quality, their Reputation Score will decrease. Once the score falls below a certain threshold, their stake will be slashed, and they will be unable to submit new data or receive any tokens until it is renewed.

This "negative" mechanism is complemented by the "positive" one described above, where the contributor is rewarded as long as their data remains the best version available for the current state of the map.

This layer is especially challenging. We’ll focus on this topic in the phase 2 of the development process. We have applied for an innovation program grant from the French Ministry of Research on the topic: "Comparison and estimation of the respective value of 3D representations". In any case, this subject will require long-term research and thus in the meantime, trust can be ensured through a partially manual validation process.

The algorithms developed by Extra Labs for this comparison and value estimation mission will form part of the token remuneration system. Each task completed to ensure data reliability will be rewarded in tokens.

2.6. How do we process the data (aka surface reconstruction layer)?

The surface reconstruction layer is responsible for storing the raw geospatial data provided by contributors and processing it. The algorithms, compute power, and storage in this layer are rewarded by tokens in the same way as geospatial data. Extra Labs should be the primary provider of algorithms in this layer and earn tokens for each task completed. The processing tasks in this layer include cropping,

cleaning, and indexing datasets, computing normals, and running reconstruction through geometric algorithms or neural networks.

During our phase of development 1, we focused on this layer. In the long run, Extra Labs' activities and research should prioritize the surface reconstruction topic, as it will be our source of revenue.

2.7. How do we allow the data consumers to access the data (aka complexity abstraction layer)?

This final layer is the closest to the data consumers as it manages their interaction with CIRCUM. The goal is to provide the developers with an experience similar to integrating the familiar Google Maps API/SDK. To achieve this, we will use account abstraction, allowing data consumers to access the data without directly interacting with the blockchain. However, the main challenge lies in providing the consumer with a stable data price, while the token price is not stable.

To address this, we will link the number of requests a consumer can make to extract or visualize the data, to the price paid in fiat for a token, rather than to the token itself. For example, if 1000 maploads cost $10, a token purchased for $5 will grant 500 maploads. If the price increases later on, a token purchased for $15 would grant 1500 maploads.

Thanks to account abstraction this approach is fully transparent for the data consumer, who only needs to complete a KYC process and provide their credit card information. By ensuring the token price can rise and fall following usage, we protect the protocol's resiliency. We express this solution through the equation: 1000 maploads = $10 ≠ 1 token.

This layer will be addressed mainly during the phase 2 of the development process.

2.8. What will our data processing research involve?

The value determination, comparison and reputation estimation described on 2.4. and 2.5. will be the main subject of our work for the next years. It will be based mainly on deep learning techniques.

One of the major challenges of our protocol is to ensure data consistency in a context where data is collected in a decentralized and uncoordinated way. This means that the same information about the shape and precise location of an object such as a building can be shared separately by different contributors, either partially or totally, and at the same or different times. This is what enables the protocol to compose complete models, for example by combining aerial data that mainly captures the upper parts of buildings with data collected from ground vehicles. The overlapping part of the data can then be realigned to produce a coherent 3D model. But this phenomenon is also a threat to the economic balance of the protocol, due to its principle of remunerating contributions via its token. Indeed, this remuneration is made on the assumption that the data has a relative value for the protocol. In the case of the contribution of information already known to the protocol, the relative value of the data for the protocol is approximately zero, but the protocol will still remunerate it. This produces a surplus of tokens in relation to the value actually present in the protocol, and therefore a reduction in the price of the token and the value distributed to each contributor. A player identifying this opportunity could even take advantage of it by

repeatedly submitting the same data, extracting a significant proportion of the protocol's value and leading to its collapse in the short term.

It is therefore crucial that we are able to identify when a specific piece of information is new to the protocol, so that we can remunerate contributions based solely on the number of new pieces of information, and not on the total volume of information.

The next logical step in this research will be to apply our data comparison capabilities to the risk of contributing false information, due to error or malicious intent.Indeed, the ability to detect that a piece of information is new - true or false - and to compare it with information contributed on this site before and especially after the information studied will be crucial in determining it as false.

2.9. Which part of the protocol will be on-chain/off-chain?

In short, the following is the breakdown of operations segregated for web3 operations in Circum:

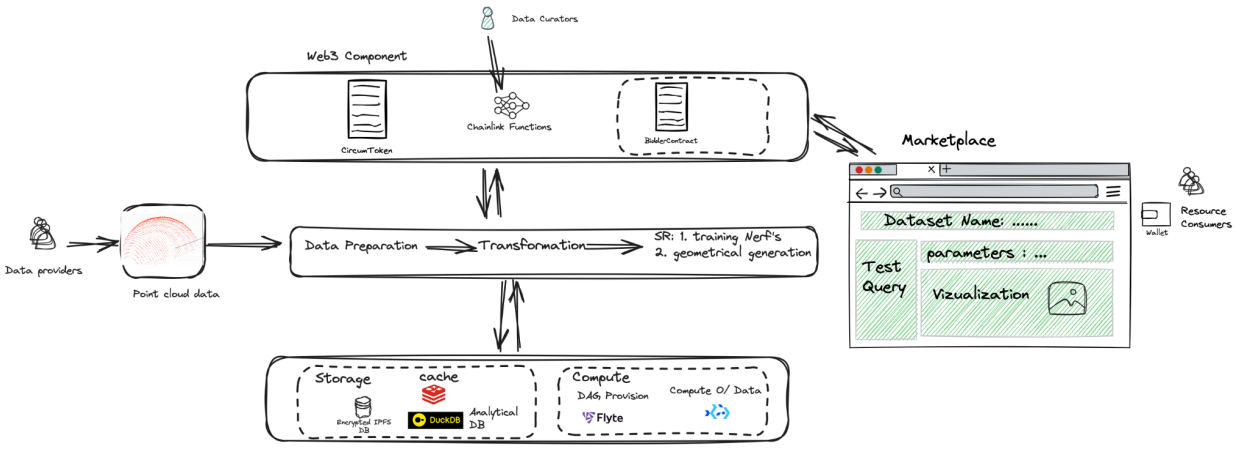
Offchain: all the operations related to the following categories:

- Data providers: Giving the raw point cloud and other measurements and their corresponding metadata.
- Point Cloud Data processing (microservices running the job on compute over data instances), data storage, calculation of the texture for reconstruction and data engineering operations (cropping, scaling and merging the pointcloud contributions). Also calculation of the contribution of each data provider to get back the results.
- The consumer's personal details.

Onchain: It will consist of the smart contracts that represent the value distribution part of the Circum protocol, like:

- Token contract that will be assigned as a unit of contribution to the network to the data providers. Simultaneously, they can consume the data from the hosted datasets based on the cost defined by the cost incurred for reconstruction + selling price of the data provider.
- Staking contract for data providers and other participants to rate the various datasets. Unlike the typical jargon of DeFI, here we don't give interest to the users who stake in a given dataset. Instead, they will maintain the reputation score of the dataset and user. Based on the utilization of the tokens for data access, the underlying value of the collateral will always be kept improving.
- Information from off-chain computers that will be used for attributing token rewards, providing proof of veracity of results, and the resulting proof data (by using oracles).
- Eventually, contracts that will allow the users to sell and buy access to the curated Point Cloud datasets using CircumToken.

The diagram of how both parts of the onchain protocol interact is as follows:



1. Upon receiving raw point cloud data, the geospatial-pipelines service automatically applies relevant transformation algorithms. Flyte orchestrates workflows, executing each transformation on the Bacalhau compute framework. Final formatted data and user/dataset metadata are then stored in the database.

2. Once the given compute operation is finished, comes the step of compensating the data providers via the CircumToken contract. This is done by first providing the value of the net contribution parameter using the ChainLink functions to the token contract (which is defined currently by the curator managed by Extralabs + with the amount of tokens that are staked by the users for the given dataset using the staking contract).

3. On the side of data consumers, they can access the data by buying the API calls to the given dataset given their value is fixed via the Circum Token.

2.10. What is the protocol tokenomic?

The protocol's tokenomic is designed to guarantee actual data consumption and to enable the protocol to continue to reward resource providers fairly over the long term.

- The supply of the token is not capped.
- Minted tokens are intended to be staked (see 2.5.) or to be burnt during data consumption (see 2.4.).
- The total supply of the token is allocated to reward the contribution of resources. No distribution to investors or team.

2.11. What does our product roadmap look like?

To prevent disconnection from the market we use an iterative approach in building this protocol. Each phase of development is concluded by a contract signed with a paying customer. See more information about our current customers in 4. Go-to-Market.

Phase 1 - Prototype phase - May 2023 - July 2024

Our initial development focused on achieving the capability to run 3D surface reconstruction algorithms on a decentralized computing infrastructure.
We established the foundation for our ability to store and process data using the compute-over-data paradigm.

Stack: IPFS, Filecoin, Bacalhau

This phase then aimed to deploy the capability to record the parameters of geospatial data contribution on-chain, while storing the actual data on a decentralized service. Recording the contribution triggers the mining of a token as a reward. This token grants access to the decentralized database and enables the holder to consume the data.

This phase produces a prototype of the protocol, showcasing its data-as-a-collateral principle and making it possible to produce data at medium scale (region/small country) with some task being monitored/realized manually. During this phase, we will build and run a controlled, medium-scale version of the protocol, focusing on a region or a country. This version will be used to deploy and test numerous features from the value distribution and complexity abstraction layers. Additionally, we will experiment and fine-tune the internal economy balance of the protocol.

Phase 2 - Research phase - July 2024 - July 2026

This phase will see heavy research efforts in order to implement the capability to better estimate and reward the value of each data contributed (see 2.8.). We'll then build on this capability to implement an automatic trust building system using proof-of-stake, reputation score, and periodic rewards instead of the manual verification system used with the prototype.

During this phase, we'll also use the prototype first to test numerous behaviors of the protocol's participants as well as features from the value distribution and complexity abstraction layers. Additionally, we will experiment and fine-tune the internal economy balance of the protocol.

Phase 3 - Globalization phase - July 2026 - …

This final phase serves as the general access phase of the protocol. While nationwide versions of the protocol may still exist alongside it, our primary focus will shift to the global version. The protocol will transition from the testnet to being independent, with a functional trust layer and no geographical constraints.

At this stage, our team will start developing tools facilitating the usage of the protocol for more and more participants: interfaces, improving modeling, texturization, plugin…

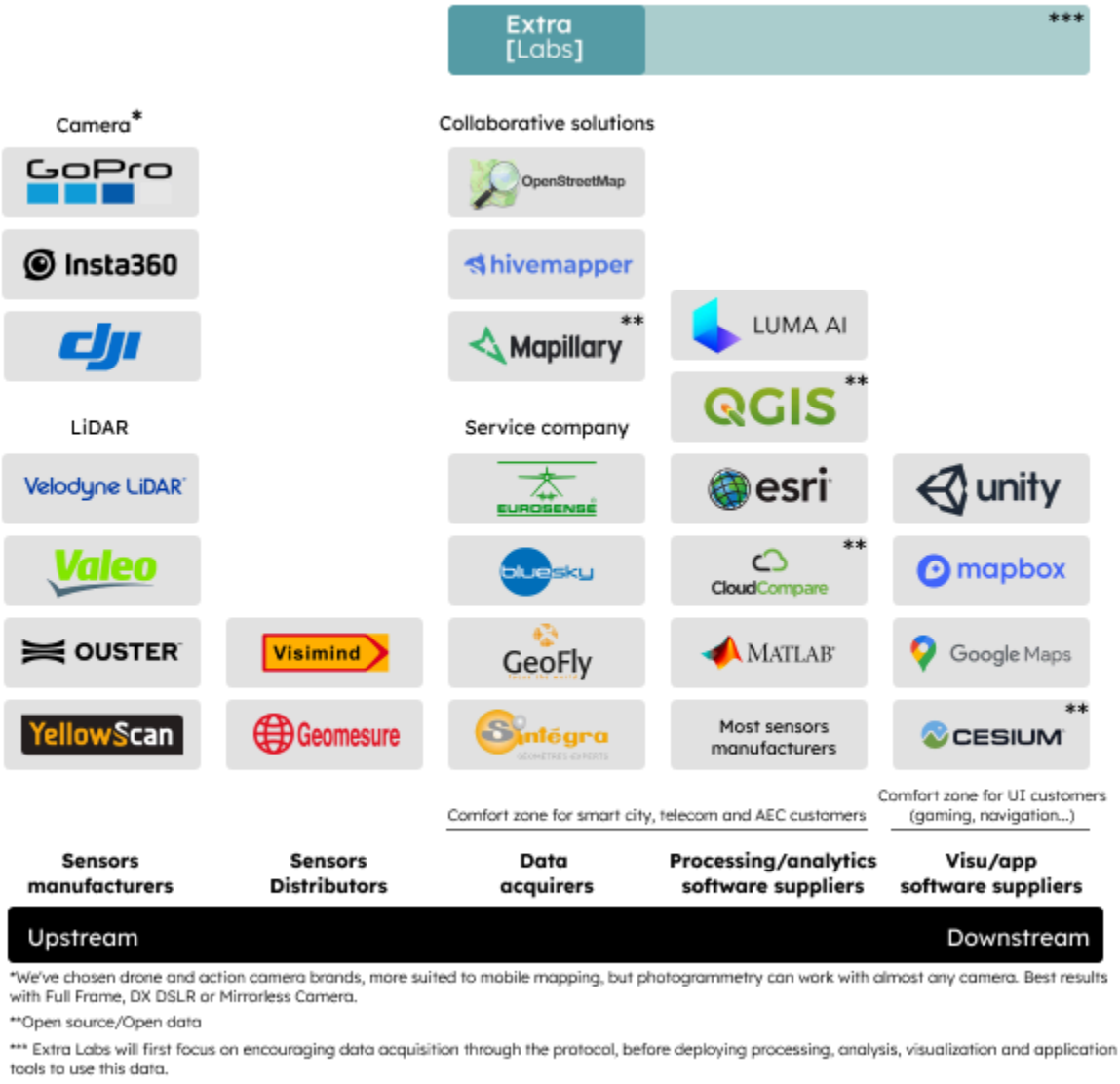
2.12. Where do we stand from a regulatory perspective?

Our activities will fall under the MiCA european reglementation. We plan to undergo the registration process in 2024 to be recognized as a digital asset service provider and allowed to perform custody of digital assets on behalf of third parties and purchase-sale of digital assets with fiat currency.

We are working with the law firm D&A partners, specialized in blockchain and tech entrepreneurship. The entire auditing and registration process is expected to cost between €100,000 and €130,000.

3. Industry

Edit 18/03/2024: Adding this map of the market to help better understand our position and relationship with other players on the market.



*We've chosen drone and action camera brands, more suited to mobile mapping, but photogrammetry can work with almost any camera. Best results with Full Frame, DX DSLR or Mirrorless Camera.

**Open source/Open data

*** Extra Labs will first focus on encouraging data acquisition through the protocol, before deploying processing, analysis, visualization and application tools to use this data.

3.1. What is the market size and dynamics?

The 3D maps market is experiencing rapid evolution due to advancements in technology, including methods, sensors, software, and processing algorithms. The market includes some large, long-established companies, a few emerging players and a wide variety of SMEs, consulting and service companies and startups. As such it's a pretty fragmented market due to its wide range of needs and use cases. Company buyouts are frequent in this market, with a certain diversity in the profile of buyers, as positioning and navigation are crucial issues for many players, whether in the geospatial sector (ESRI, Trimble), mobility (Uber, TomTom, HERE), software (Google, Apple, Microsoft, Samsung) or telecommunication (Qualcomm, Intel).

Firstly, a distinction is made between "upstream" solutions, corresponding to the data collection, extraction, cleaning and preparation phases, and "downstream" solutions, corresponding to analysis, modeling, visualization and other data processing.
The market is divided into several segments: urban planning and smart cities, autonomous mobility, telecommunications, defense, emergency, video games/AR/VR, education, agriculture, tourism, architecture/construction and climate/environment. All these segments are growing, especially autonomous mobility and urban planning. To enter the market, we will be concentrating our efforts on urban planning, where the existing players are less innovative, the competition less significant and the solutions less effective. We estimate the overall size of this segment at $430 million[1], and the size of the 3D map market at $8-12 billion[2].

In the urban planification segment, data is generally collected and pre-processed by specialized local companies. These SMEs have small twin-engine aircraft and drones equipped with sensors, including LiDAR sensors. They carry out inspection and mapping missions for local authorities and companies. For example, in France, the LiDAR HD program is run by a consortium comprising Geofly, Sintegra, PixAir, BlueSky, and Eurosense. On the downstream side, a handful of large companies offer Geographic Information Systems (GIS) solutions, often with products geared towards digital twins or urban planning needs. ESRI is the flagship company in the GIS market, with its ArcGIS software. Trimble provides software solutions and specialized hardware for the construction, agriculture and mobility sectors, as well as for data acquisition. Bentley Systems and Dassault Systèmes are particularly active in the fields of digital twins and urban planning. In 2010, Dassault worked on Rennes Métropole's digital twin, one of the first projects of its kind in France.

3.2. Who are our competitors?

We plan to focus our initial efforts on the urban planning and smart city segment of the 3D map market, more specifically on the upstream part, i.e. data collection and preprocessing.

Ground and aerial survey companies

There are relatively few alternatives to the collection programs commissioned by local governments from ground and aerial survey companies. However, this solution is costly and the data quickly becomes obsolete. In addition, the time lag between the expression of need by community departments and the delivery of pre-processed 3D models is usually several years. We see these companies as potential contributors to the protocol, opening up a new business model for their activities.

Google Earth and Maps

The giants of 3D maps are Google Earth and Google Maps. Google Earth is the original 3D visualization platform, mainly used for educational purposes, while Google Maps is primarily intended for navigation.

---

[1] Estimation basée sur les chiffres de lUN World Urban Prospects et de la European Commission Européenne: 4.3 milliards de personnes vivent dans 10 000 villes à travers le monde. Nous supposons qu'une municipalité est prête à dépenser en moyenne 100 dollars par an pour une carte actualisée de l'espace de vie de 1 000 personnes.
[2]https://wgicouncil.org/world-geospatial-industry-council-releases-a-report-on-spatial-digital-twins-trends-and-opportunities/

However, both platforms include a combination of hyper-realistic textured buildings and 3D models in metropolitan areas. As with other centralized maps, Google faces problems of data obsolescence, particularly for 3D, which requires significant investment. Data for areas outside major Western cities is often more than 5 years old. New 3D features such as Immersive View are mainly available in a few major cities, and while we can anticipate progress in this area, it's unlikely that the majority of the map will ever benefit from this level of service. Above all, Google does not allow developers to extract data for use in other software. Google Earth proposes an engine that allows developers to run a simulation or create a custom model in Google Earth for demonstration purposes, but they cannot use Google's data for most smart city use cases that require specialized software (and more up-to-date data).

Luma.ai and NeRF companies

A few companies are emerging on the 3D building modeling market thanks to advances in artificial intelligence in the sector. Luma.ai, for example, relies on Neural Radiance Field (NeRF) technology to reconstruct an implicit 3D model from a video that can be captured using a simple smartphone. We have also mastered this technology and tested it with good results on our decentralized computing infrastructure. However, this method requires a relatively long capture time with multiple shots for each element to be reconstructed. Therefore, it is suitable for the reconstruction of buildings but not entire cities. However, we are closely following the progress of this technology and in the future, we could potentially offer a pipeline to allow NeRF users to become contributors to the protocol.

Hivemapper

The company closest to Extra's model and objectives is Hivemapper. We see them as our main competitor among innovative solutions, even though their focus isn’t currently on 3D. Hivemapper manufactures dashcams that contributors must purchase (two models, usually around $300 for the original model and around $650 for a version with better cooling and video features) and install in their cars to start providing data. The collected data consists of street images and Hivemapper also uses contributors to manually train the dataset in order to extract information about traffic signs, road obstacles, or roadworks. Hivemapper also compensates its contributors through a token that is used to share value among network participants. Hivemapper raised a total of $23M, with the last round being a Series A of $18M in 2022.

We believe that their dashcam-based model and supply-led go-to-market isn’t the right solution for the use cases we are trying to serve.Indeed, their decision to engineer their own dashcam, instead of utilizing data from various devices, stems from their experiences with heterogenous datasets that didn't adequately meet their customers' needs or align with their processing technology. This may have been right when Hivemapper started developing their dashcam a few years back, but the AI breakthroughs have significantly changed the situation in the meantime. Hivemapper hasn’t communicated on any of their own progress on AI processing. Being a hardware provider comes with its own challenges, at the time of writing the delay between order and receipt of a Hivemapper dashcam in the US is about 6 months. It’s also unlikely that Hivemapper makes any profit on their dashcam. Moreover, the choice of using a dashcam instead of a roof-mounted system exposes the data to reflection on the windshield. Hivemapper also encountered issues with the volume of data to transfer. Drivers are meant to connect their phone to the dashcam at the end of their drive to upload the datasets. However, both the dashcams are 4k and the volume of data was such that drivers were complaining about the uploading time. Hivemapper chose to reduce the quality of the datasets that were collected in order to reduce the uploading time. The result is that the data quality is lower than that of Street View. This poor quality and the limited angles coming from a road-bound camera make it impossible to produce good quality 3D models.

The token of the network provides liquidity for the suppliers through speculation, making it easier for Hivemapper to attract drivers. They made the choice to focus on supply first and at the time of writing claim to have mapped about 10% of the world's roads, with mappers in more than 90 countries. They make abundant use of token drops and award drivers a significant amount of token each week. This amount is quite disconnected from what Hivemapper (shyly) announces as an estimated price for data consumers. To date, we haven't been able to find a single confirmed consumer of Hivemapper's data. This introduces uncertainty into the value of the collected data, a significant issue in a data-as-collateral model.

OVER the reality

OVER focuses on the metaverse and augmented reality market. Contributors use a mobile application to take photos and videos of the world, organized in the form of ground-drawn hexagons. OVER operates in a closed ecosystem, and the tokens that are mined to reward contributors are mainly used for purchasing Metaverse experiences. This weakness of their internal economy has led to a continuous decline in token price since mid-2021, and its usage seems to have significantly decreased since then. Neither Hivemapper nor Over have made significant progress on the digital twin segment and the needs of local governments in terms of 3D mapping.

Therefore, our positioning is based on the practicality and cost of 3D data. We rely on the same technology as aerial and terrestrial surveying companies, namely LiDAR and photogrammetry. However, whereas these companies charge a very high cost to a single client for flying planes solely for data collection, our model allows us to motivate actors to collect and contribute this data alongside another activity or need. The compensation price for these data providers is also supported by a large number of consumers. This allows us to offer a low price and almost immediate availability for up-to-date data.

4. Go-to-Market

4.1. What is our business model?

During our experimentation phase (roughly 2024-2026) we'll get direct revenu from our experimentation partners, which are local governments and companies involved in the acquisition and processing of 3D geospatial data. The work we carry out together focuses on crucial issues for these players, enabling us to validate the suitability of the solutions we develop for market needs, and to begin forging links with future contributors and consumers of the protocol.

Once the protocol will be up and running, we'll transition to a business model based on the volume of value circulating on the protocol. Indeed our company aims to develop a "lab" business model, popular in the DeFi ecosystem with companies like Morpho Labs or Mangrove DAO. It involves generating revenue by contributing to the technical development and maintenance of a public good. CIRCUM is designed to support this business model by utilizing a token that effectively distributes a portion of the protocol's revenue to resource providers. The company EXTRA will be one of these resource providers, earning tokens by completing tasks that require surface reconstruction or data comparison algorithms. The tokens collected will then be sold on the market. We estimate that EXTRA will be able to capture around 20% of the value passing through the protocol.

Our first offering to data consumers will be based on the ability to extract 3D data for use in other software. This capability is essential for users in the urban planning and smart cities sector, which is the one we'll be targeting first. The price at which consumers will be able to extract the data - via the purchase of tokens - is still uncertain, but the models we are running for the first experiments on the subject show a price of around €200 for the reconstruction of a 1 km² tile. Various ETL pipelines (Extract, Transform, Load) will be made available to enable transfer and exploitation of this data.

In the longer term, we will be rolling out a second offering enabling developers to integrate protocol data for visualization only. This will take the form of SDKs and plugins for 3D engines such as Unity or UnReal Engine. Its price will be close to that of Google Maps, $10-15/1000 maploads. The aim of this offering will be to target users in the gaming and navigation sectors, as well as enabling the development of much more advanced tools than those possible via our ETL offering. To complete the source of data here, we would develop texturing services and 3D models completion through stable diffusion.

Overall, we expect the experimentations to bring in the first million in sales, targeting essentially local governments. Then the ETL offer would bring in the next 5 million from both local governments and companies invested in digital twin technology, construction, mobility, telecom… Finally, we would introduce the visualization-only offer to complete the ETL one and expand to new sectors.

4.2. What is our current traction?

Since the beginning of our project we have reached out to 50 potential data customers, mostly local governments but also large and small companies, digital twins solutions providers and gaming studios, from France, UK, USA, Japan, Italy, Germany, Netherland and Finland. We hold interviews with about half of that number to better understand how they use 3D data and what they need in a source of data. We had the same process with data productors and interviewed 5 of them.

One of these data productors, named Helix Drone Corsica, ended up signing our first contract in December 2023. This small contract (1000€) served as a proof-of-concept for our decentralized computing infrastructure, applied to the processing of data for a project commissioned by the French Forest Office. The project aimed at providing a better picture of the damage caused to the beds of Corsican rivers by the floods of October 2023. The test allowed us to validate the functioning of our compute-over-data infrastructure and the benefits in terms of cost reduction.

We are currently in the process of signing a second, much larger contract (50-70k€) with the Ile-de-France region (around Paris) to run an experiment of the protocol principles at a small scale. This corresponds to the end of phase 1 of our development and the production of a working prototype. This project will take place in a high school and focuses on incentivizing high schoolers to collect pictures of their school using their smartphones. The goal is to reconstruct a 3D model of the school using a NeRF model that runs on decentralized computing services. We expect this contractualization to be complete in mars and the project itself to take place in may-june. Other contracts with the Ile-de-France region could follow if the experiment is successful, in order to scale the protocol to incentivize the collection of data at the region scale (Phase 2).

We are also in discussions with the French Geographic Agency to start testing the protocol's potential together. Our aim is to deploy the first national version of the protocol in France. This would provide a testing ground for our phase 2, before moving on to a global version of the protocol.

This logical sequence of French river then high school then region then country will form the backbone of our development. We also have sales processes underway in other countries and plan on conducting the same experimentation with different partners in order to expand our learnings, get revenue and expand our network of future suppliers and data consumers.

4.3. Who are our customers?

During the development phases of our protocol, we will focus on the segment of urban planification. This segment is dynamic, open to new solutions and does not require large-scale or near real-time capabilities. The target customers for this segment are cities, regions, and even national administrations. The ideal consumer profile is that of a city with a population of over 50,000, that have already conducted projects related to 3D modelisation and LiDAR surveying. Through these projects, they have firsthand experience with the cost of these programs and the challenges of data obsolescence. While our commercial efforts are not limited to a specific country or continent, we have observed that European cities are often more advanced on the digital twin reflexion.

After the development phase of our protocol, we will expand our commercial efforts to companies using 3D geospatial data for digital twin related use cases. These are companies in construction/architecture, telecom, event planning, transport… As an example, we had positive discussions with EDF (multinational electric utility company) and Generation Park in Houston (commercial development).

Then we'll adapt our offer (see business model) to target companies in gaming, AR/VR, navigation and other consumer-facing use cases. Our goal is to offer a data source that can be realistically utilized by these applications, both from a technical and economic standpoint. We believe that precise and comprehensive 3D maps could greatly enhance mobile and desktop games, navigation applications, social applications and delivery services. The developers most likely to utilize our data source are those who are already working with 3D engines such as Unity or Unreal Engine.

4.4. How do we acquire our customers?

In the early stages of our development, we contact potential customers directly or through trade fairs such as the Smart City Expo World Congress in Barcelona. Extra is also an active member of the Urban Technologie Alliance, a global association that organizes experiments between innovation players and local governments, and Datalliance, a network that connects companies working on geodata challenges with the french government.

In addition, we are building relationships with geospatial developers working for local governments by participating in open source conferences (FOSS4G) and starting a small community through a mailing list and working groups.

As we move towards a self-service model, we will be strengthening our sales and marketing strategy to reach more potential consumers. We will maintain the bottom-up principle of targeting developers themselves to evangelize their employers. This will translate into a strong presence at engineering schools, hackathons and technology events, as well as an extensive documentation strategy and content creation.

4.5. How will the distribution work?

Beyond our experimental phases, data consumers should progressively be able to connect directly to the protocol without interacting with salespersons. Our aim is to enable them to consume data in a self-service way, by implementing a 3D engine plugin, SDK or API in their application. The integration and payment experience will be designed to be as close as possible to the familiar Google Maps experience. The entire process of purchasing and using tokens will be transparent to the data consumer.

In addition, we'll be working on the development of plugins to integrate our data source into Unity and Unreal Engine.

4.6. What is our moat?

Our protocol is designed to offer the best 3D maps data available at the lowest price. It relies on a network effect powered by its token. Each new contributor adds more data to the protocol, resulting in a better map with updated information, more details, and wider coverage. This, in turn, attracts more data consumers and distributes value to all contributors, which further attracts more contributors. We believe that this network effect would soon become our main defensibility mechanism. A competitor would face difficulties in convincing data providers to grant them exclusivity, unless they were willing to pay a high price. Paying for exclusivity would limit their ability to scale quickly and offer competitive prices to their customers.

5. Supply Side

5.1. Who are the supply providers?

The supply side consists of two categories of actors: those who already collect data and those who have the potential to collect data given the right circumstances and financial incentive.

The actors who already collect data include:

- **Aerial and terrestrial surveying companies** such as the companies involved in the LiDAR HD project in France: Geofly, Sintegra, PixAir, BlueSky, and Eurosense. This category also includes companies flying drones. Industry experts estimate that more than 20,000 hours of drone flying were carried out by professionals every day in 2023. These companies typically collect data based on customer orders. With their customers' agreement, they could contribute the same data to our protocol. Additionally, they could utilize their unused sensors to collect data specifically for the protocol. This data collection model, where data is collected first and access is sold later, is already implemented by Jakarto, a surveying company leader in North America with which we have contacts.

- **Autonomous mobility companies** such as Uber, Cruise, and Google use LiDAR in their journey towards driverless cars. Even mainstream companies like Tesla, BMW, Honda, Mercedes-Benz, and Hyundai-Kia utilize LiDAR to enhance the driving experience and safety. In the future we could partner with these companies to get their data into the protocol.

- **Drone hobbyists** are becoming increasingly numerous. 5 millions of new consumer drones were sold in 2020, a number that should double by 2030. These people often film remote locations that other data providers wouldn't pay attention to. Their drones rarely carry LiDAR sensors but their video could produce 3D point clouds through photogrammetry. They could bring much value to the protocol assuming the contribution process would be smooth.

The actors who could collect data with the right incentive are:

- **Delivery and transport companies** (post office, cabs, Chronopost, FedEx, UPS...) could equip their vehicles with sensors to collect data along their routes. Some of these companies have already diversified their revenue streams by selling advertising on their vehicles. In Japan, the company Symmetry and the Ministry of Land, Infrastructure and Territories have led an experiment in 2021 where they placed sensors on top of buses and cabs and produced a 3D model of the city through multimodal reconstruction. We had discussions with their teams and received the full rapport of this experiment to support our own development.

- **Citizens** as well could equip their personal vehicles with sensors and contribute data.

5.2. What is the value prop to the supply side?

The main incentive for all actors, regardless of whether they already collect data or not, is money. The protocol will ensure that these actors receive financial value for their contributions as long as data consumers are willing to use the data. The balance of the protocol, which means ensuring that the value received by the providers is sufficient to generate a benefit from their costs, will be a focal point for us.

Note : we estimate the price of the sensor installation on a vehicle to be around 1000€/vehicle (Hivemapper have two dash cams at $300 and $650 at time of writing, with a delivery time of up to 4 months).

Another incentive is that supply providers may have their own 3D data needs. By contributing to the protocol, they can earn a token to access datasets from others.

5.3. How do we acquire supply?

Our first action will be to integrate the datasets available in open data, products of public acquisition programs. This will produce a basis for our 3D map and a reference point to compare new datasets. A lot of free data is available worldwide, often several years old, but still 80% correct. The 20% difference is what interests data consumers, particularly in the field of urban planning, because it reflects the impact of local government work: new infrastructure, new buildings, new transport. Using the 80% as a basis for our map allows us to focus on the high-value 20%.

During experimentations, our strategy is to acquire the demand side first then to identify local delivery and postal services to invite them to join the experiment. The support of the local government will be useful to convince the potential data providers and to partially fund the sensors.

Speaking of the sensor price, we plan to sand down this friction through partnerships with LiDAR sensors producers and funding programs. We'll focus our direct efforts on companies with a fleet of more than 5 vehicles. We'll also invest in professional and hobbyist events, especially targeting those around mobility and drones.

We also have an open source strategy, especially targeting drones operators and surveying companies. We leverage our research team to produce new processing pipelines for the data they collect. These innovant, more efficient pipelines will have the double mission to attract interest from these companies and to facilitate the contribution of their data by bringing it to the right place.

5.4. How do we onboard the supply?

Supply providers will be able to integrate an API into their data storage infrastructure, simplifying the process of submitting data to CIRCUM. They would also have the option of interacting with a front interface, for a less technical contribution.

We'll invest in making sure that the contribution experience on the blockchain is as transparent as possible for the contributor, using account abstraction and a dedicated front interface.

6. Vision

6.1. What does the future look like if we succeed?

CIRCUM has the potential to revolutionize the funding, collection, and sharing of geospatial data. By providing a detailed, dense, up-to-date, and cost-effective data source for 3D maps, we can bring our industries into the realm of 3D and make a significant impact on urban planning, mobility, and consumer applications.

However, the impact of CIRCUM could be even greater in developing parts of the world. This includes 85% of the global population and most of the fastest-growing cities, such as Delhi, Shanghai, Dhaka, Kinshasa, Lagos, and Cairo. These cities face rapid development but struggle to provide their residents with basic services like housing, transportation, water access, electricity, and waste management. In these areas, the traditional method of collecting large-scale LiDAR data every few years is highly impractical due to the rapid obsolescence of data and the limited investment capacity of local governments.

A new funding and collection model like CIRCUM's could bring significant improvements in how local governments, NGOs, and companies understand and monitor these cities. This could be a game-changer for urban planning and urban services, as well as for other sectors of the economy, like e-commerce. Indeed, accurate 3D maps have the potential to solve the last-mile problem in a world where 4 billion people lack addresses.

Beyond 3D cartography, Circum introduces a new approach by using decentralized technologies to facilitate efficient collaboration between humans and machines in creating a public good.This permissionless method of distributing tasks, ensuring honesty, and rewarding contributions allows for the production of a more efficient and scalable service than what could be achieved in a centralized manner.

While CIRCUM has been initially designed for the use case of 3D maps, its infrastructure could actually be utilized for a wide range of other purposes. In the future, we plan to fork it to collect various types of data for digital twin purposes, including indoor and underground mapping, pollution, temperature, and more. Our work will be open source, aiming to inspire other projects in this direction.

7. Near-term strategy

7.1. How have we been funded until now?

We were granted 90k€ by the French Public Bank of Innovation (BPI France) through the "Bourse French Tech Emergence" program. Additionally, we obtained 50k€ in debt. We have also started to bring in revenue.

7.2. How much are we raising now? When do we expect to close the fundraising?

We are looking for 700k€ to 1,5M€ for 20% post-money. We are aiming at closing in the spring of 2023.

7.3. Why are we raising now?

We feel we have made sufficient progress in our understanding of the market, customer needs and revenue-generating opportunities to be able to invest further in developing our product. We want to find the right balance between rentability and growth and that requires external capital to accelerate our experimentation/development phase. We could keep developing the company without it, but it would mean that the protocol wouldn't be live for 4 years. With funding we expect to deploy the first countrywide version in 24 months.

We want to invest in the team first and foremost, hiring solid people to build the best technical and operational models.

7.4. What are our goals for the next 3/6/12 months?

In the next 3 months:

- Complete the first contractualization with the Ile-de-France region to conclude our phase 1
- Complete fundraising before April 2024
- Hire a senior DevOps and a second smart contract developer

In the next 6 months:

- Everything above
- Start phase 2 of the development process
- Hire a deep learning engineer and 2 research engineers
- Launch an innovation program on the comparison and estimation of the respective value of 3D representations (supported by the french government)

In the next 12 months:

- Everything above
- Hire a postdoctoral researcher on 3D data processing and a second deep learning engineer
- Sign a contract with the French Geographic Agency

Market references and specific knowledge

Here is some more information about the geospatial / 3D map technologies and market.

About the market and the competition

- A Minds behind Maps podcast episode about Hivemapper
- Another Minds behind Maps podcast episode about Google Maps and Niantic
- A Messari DePIN report that outlines the potential of decentralized mapping
- About geospatial product made for geospatial folks and about the consumer map market
- About the amazing but a bit top-heavy 3D map rolled out by Zenly a few months before Snapchat killed the company. Detail: This map has made a big impression on the geospatial ecosystem. We got in touch with the Zenly team, who explained that many companies had contacted them to ask if they could reuse this map for their own projects. However, the map was so highly customized and the infrastructure so delicate that it was not possible to make a product out of it. Following the announcement of Zenly's closure, several unsuccessful offers were made to Snapchat to buy the map division.

About the technology

- Simple introduction to point clouds
- Explicit vs implicit 3D modeling
- Report of the project between Symmetry and the japanese Ministry of Land, Infrastructure and Territories (in japanese, sorry, but there is pictures)