



Valuing data and Storecoin *datacoins*

A process to determine the *value* of the tokenized app data in order to value the *datacoins* they back



April 2019

Why does this matter?

By 2025, it is estimated that world will emit 175 zettabytes of data, 5x more data^[1] than today.

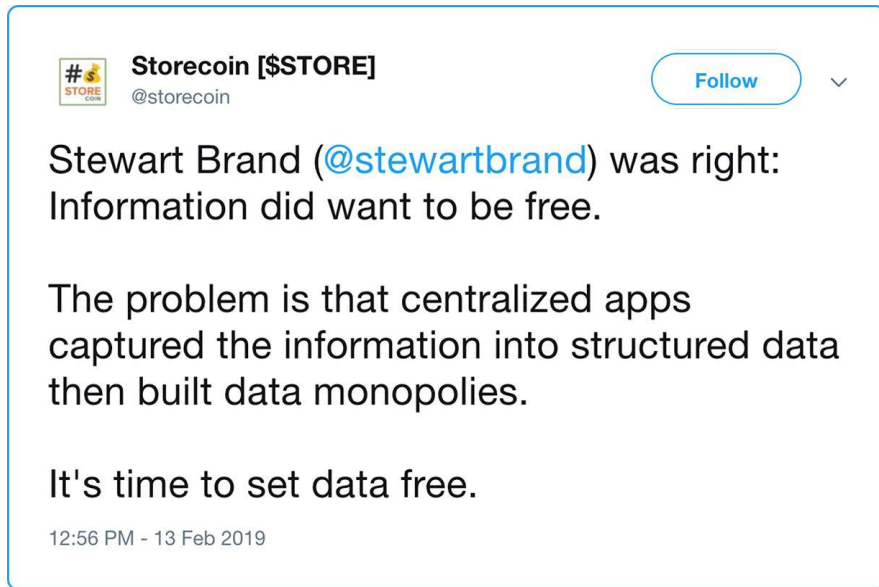
Tokenization that creates incentives for data to be structured, discovered, and tradable could transform computing.

The Storecoin p2p cloud computing platform will enable data to be tokenized, open, and tradable. Storecoin will transform data into p2p money -- into data-backed tokens called *datacoins*.

Each time a third party wants to crawl, query, or search the tokenized data, they'll pay Storecoin miners (miners) and the developers building on top of the Storecoin public blockchain for access to the tokenized data using the datacoin. Developers can then, if they wish, pay users, creating an opening for new business model experiments.

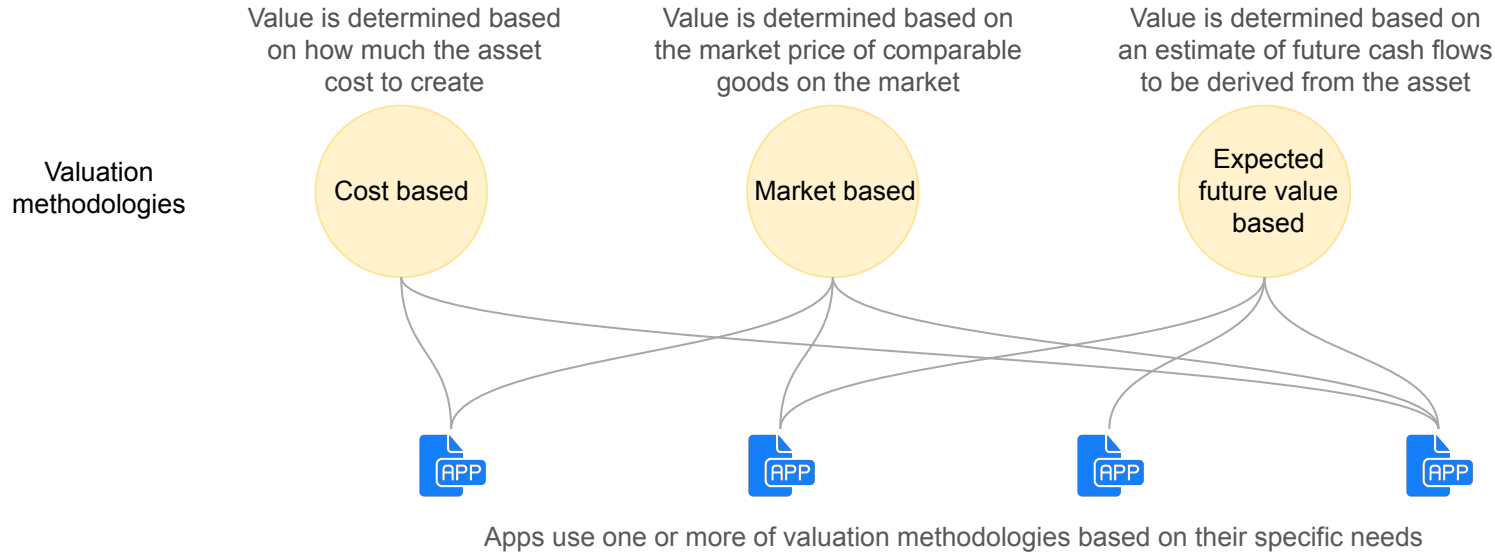
Tokenized, open data will limit the control data monopolies have on future innovation while ushering in a new era of computing.

Datacoins will set data and information free.



Why do we need to value data?

Storecoin's p2p cloud computing platform facilitates the transformation of data into programmable money (*datacoins*). To help datacoins be valued, we need to define a process to value the data underlying them.

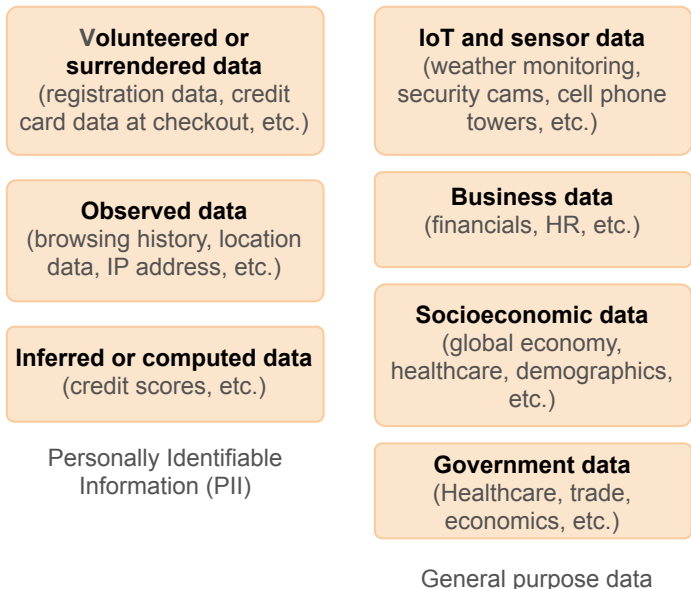


Key point: One size doesn't fit all. Valuing data in Storecoin Platform will be *app specific* and use one or more of the above methodologies.

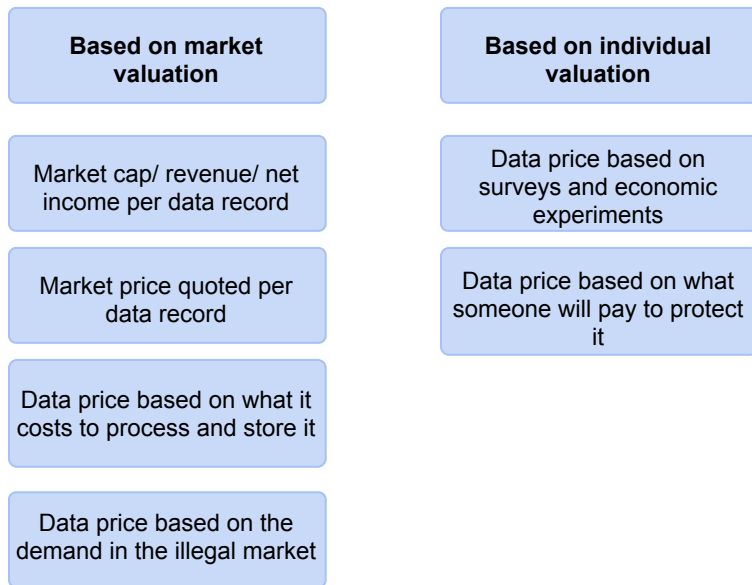
Data as an asset is massively valuable — especially structured data

Yet, there are no methodologies for determining the monetary value that would work for all types of data and in all situations. So, a *contextualized process* is necessary for determining the monetary value of the data.

Data types*



Monetary value determination**



* This list is not exhaustive, but is meant to show different types of data in different domains and with different interests.

**<https://www.oecd-ilibrary.org/docserver/5k486qtxldmq-en.pdf?expires=1552263628&id=id&accname=guest&checksum=2E0EF5B6557470927E8ABD6FFC347DDE>



No data valuation methodology is perfect

Each valuation methodology has its benefits and drawbacks. Sometimes, the data price is determined based on what the market determines. In other cases, it is based on individual evaluation.

Example:

Financial results per data record have some benefits as a methodology including ease to identify, but are likely reflect too many additional factors to be taken alone.

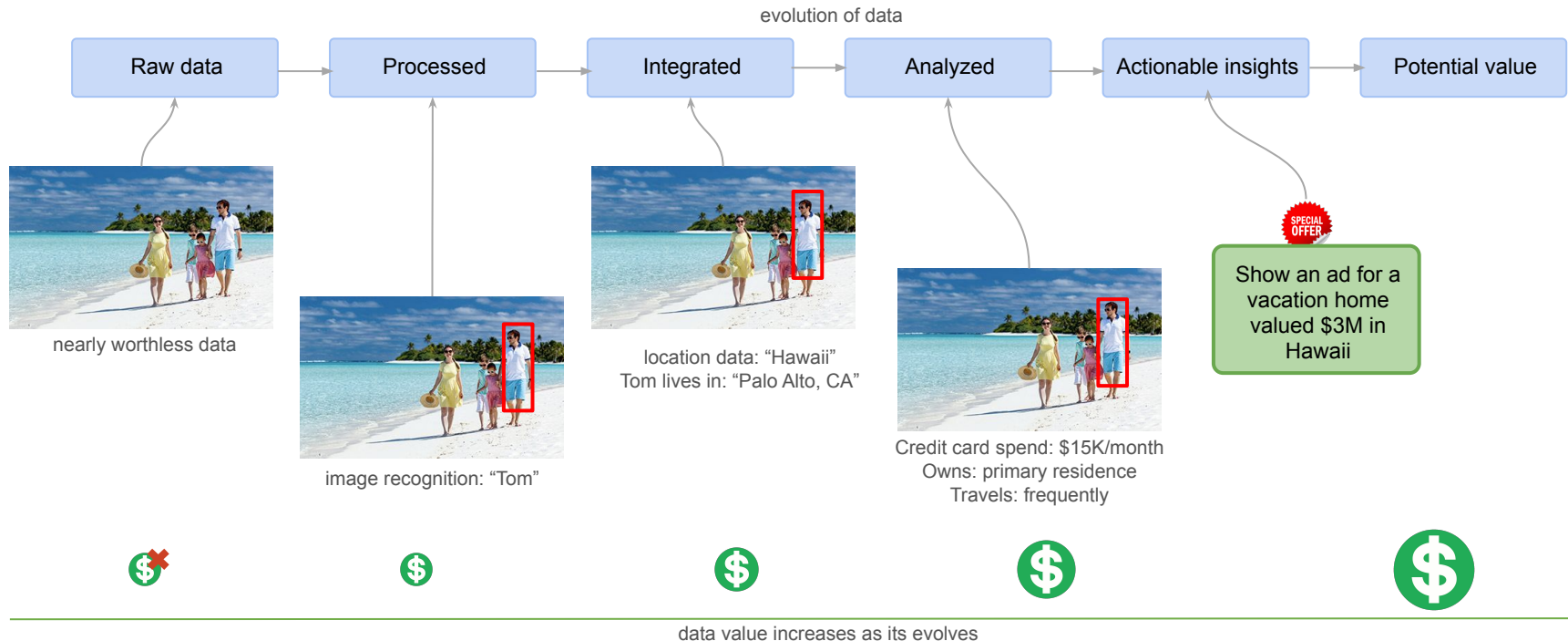
Indicator	Description	Benefits	Potential Drawbacks
Indicators based on market valuation			
<i>Financial results per data record</i>	Aggregated market cap (revenues, or net income) of a company divided by the total number of personal data records used by this company.	<ul style="list-style-type: none">- Relatively easy to identify.- Reflects actual economic value added generated through personal data.	<ul style="list-style-type: none">- Likely to be inaccurate, as numerous other components impact market cap / revenues / income of a company.- Possible synergy effects could lead to overestimates for firms with larger datasets. Appropriateness of this approach depends on what portion of turnover is directly tied to personal data.

See appendix for full data methodology breakdown



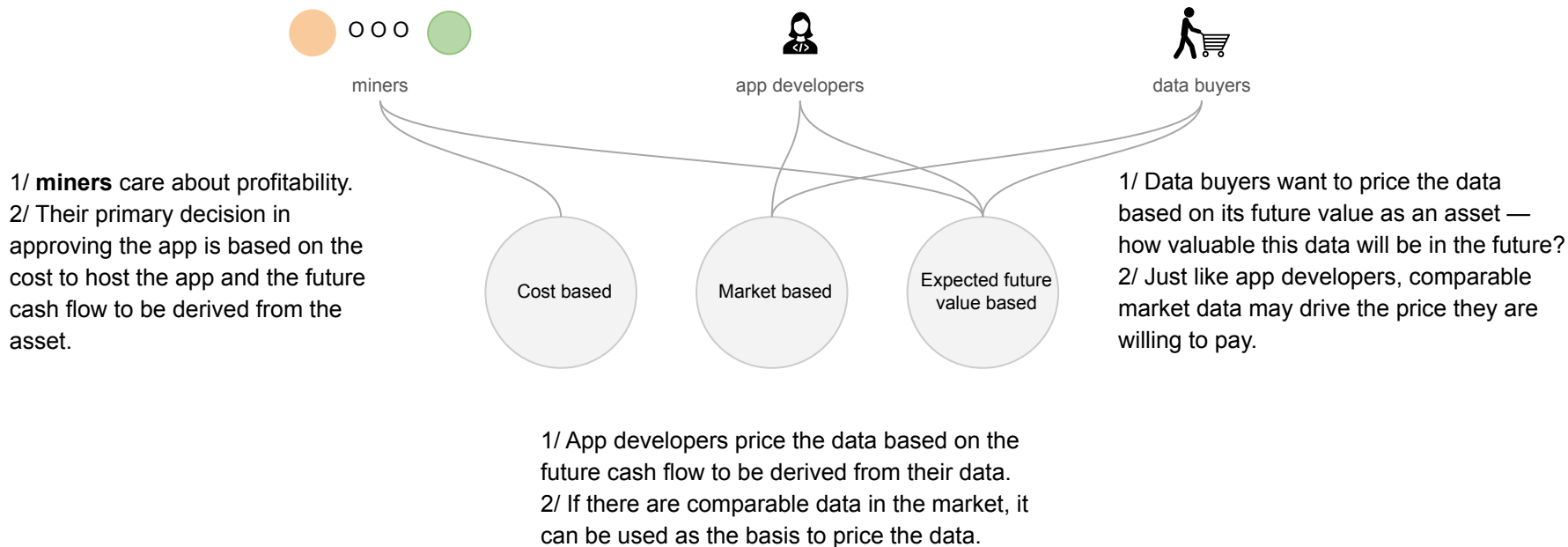
Data value increases as it evolves through its lifecycle

Raw, uncategorized, unstructured data is rarely valuable. As the raw data is processed, aggregated, integrated and analyzed, its value increases. If we facilitate this evolution, we can ensure the highest value for the data.



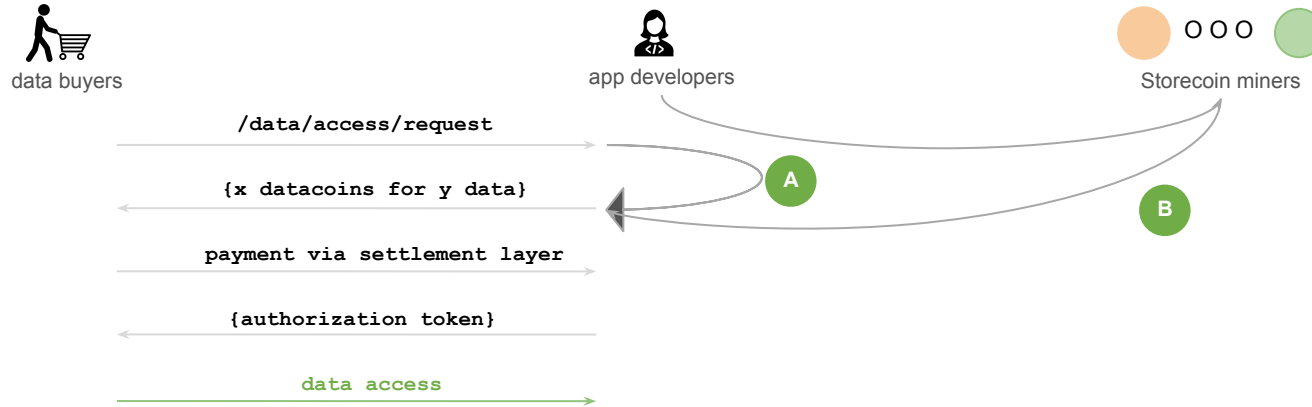
Hybrid approach to valuing data

Since one-size-doesn't-fit-all, a hybrid approach is devised where Storecoin miners, app developers, and potential buyers participate *together** to value the data.



Hybrid approach to valuing data — what does this mean?

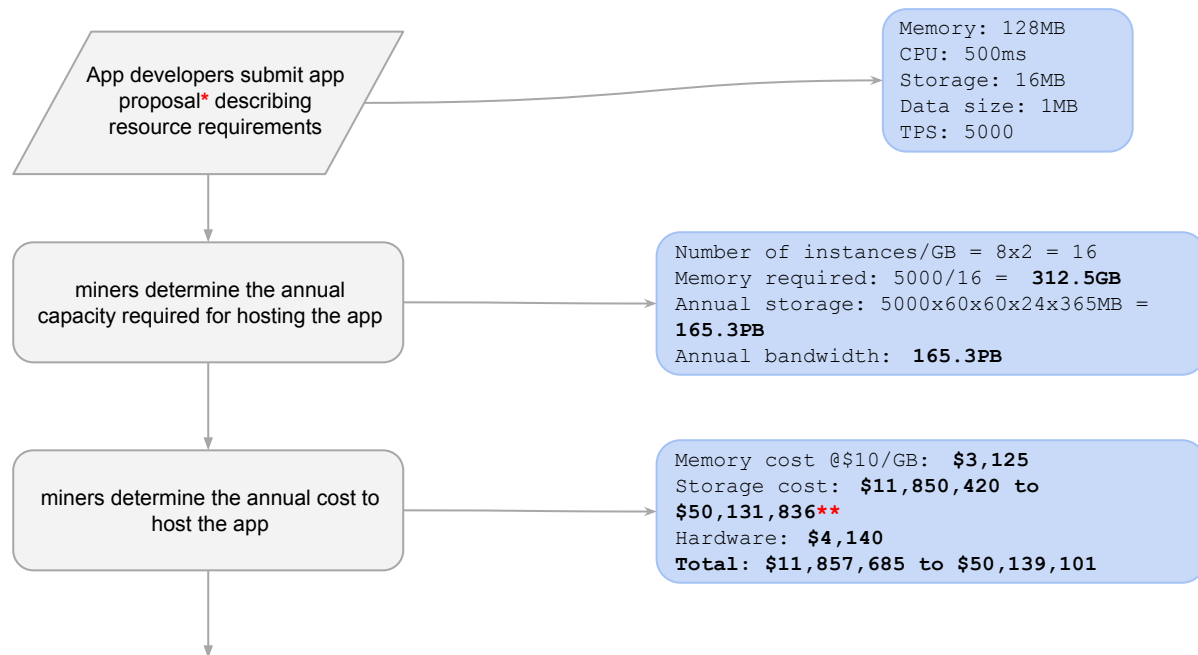
Storecoin miners, app developers, and potential buyers are likely price the data based on different *tiers*. Custom pricing is also possible for bulk requests or data requests across multiple apps.



- A** Data access request is made for one of the predefined *tiers*. The buyers complete necessary payments for the selected tiers of data access to get the authorization.
- B** Data access request cannot be satisfied with predefined tiers of access. In this case, a custom access is requested or the request spans across multiple apps for *similar* data. miners negotiate such access requests and price the data.

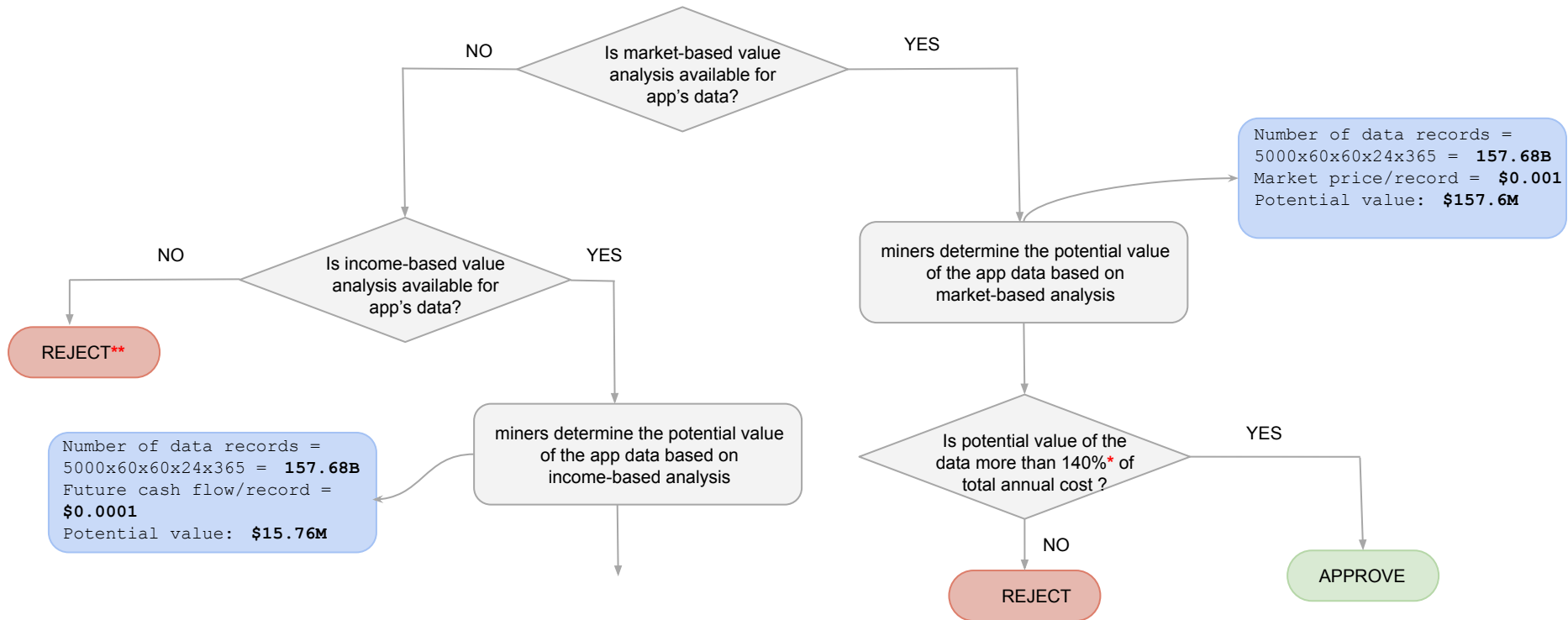
Valuing app data — the workflow

Storecoin miners use the following process to arrive at a value for the app data.



Valuing app data — the workflow (continued)

Storecoin miners use the following process to arrive at a value for the app data.

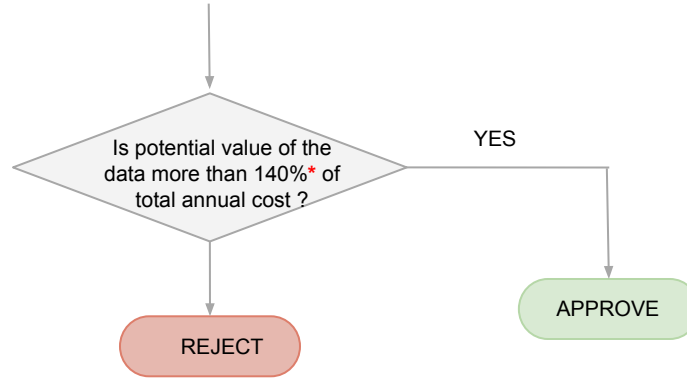


* This assumes a margin of 40%.

** Since no reliable methodology is available to value data, the proposal is rejected. However, the app developer can negotiate to compensate miners in STORE.

Valuing app data — the workflow (continued)

Storecoin miners use the following process to arrive at a value for the app data.



Determining the cost of hosting the tokenized app

Storecoin miners determine the cost of hosting (runtime, storage, and bandwidth) the tokenized apps from the app proposals submitted by app developers.

Used to determine the compute capacity required for Validators.

Used to determine the storage capacity required at the Messagenodes.

Used to determine the bandwidth capacity required between the Validator and Messagenode networks.

1. Memory required to run an instance of the app.
2. CPU time required to run an instance of the app.
3. Storage required to persist the data produced by the app instance.
 - a. Categories of data produced per app instance. Ex: "personal_data", "financial_data", etc.
 - b. Size {min, max} of data for each category.
4. Bandwidth required per app instance. By default, this is the same as storage requirement, but it can be different, if the app runtime uses external/additional datasets for its computation.
5. Number of instances expected per second (throughput or TPS)

Total annual cost to host the app = SUM (
- memory required for the specified TPS,
- CPU capacity required for the specified TPS,
- bandwidth required for the specified TPS,
- annual storage required* based on the specified TPS
)

Used to determine the annual compute, storage, and bandwidth required to host the app.

Seven laws of information — why data decentralization is profitable

Answers why would app developers want to decentralize their data when they can freely sell the data with existing centralized services.

1. Information is (infinitely) shareable.
2. The value of information increases with use. Information has no real value on its own. It only becomes valuable when people use it.
3. Information is perishable.
4. The value of information increases with accuracy.
5. The value of information increases when combined with other information.
6. More is not necessarily better.
7. Information is not depletable.

Seven laws of information — why data decentralization is profitable

Law of information	How Storecoin Platform enables it
1/ Information is (infinitely) shareable.	The data is discoverable across apps, so “similar” data can be discovered, negotiated for, paid for, and accessed across apps.
2/ The value of information increases with use.	Data from different apps don't live their own silos, resulting in increased use. See discoverability above.
3/ Information is perishable.	Storecoin Platform is data-centric and exists exclusively for that purpose. So, this fact is highlighted in the data pricing approaches it uses, thus encouraging apps developers to share as much of the analyzed and structured data as they can while they still have the shelf life.
4/ The value of information increases with accuracy.	Discovering and accessing data across apps allows for the creation of new generation of startups that continually improve the quality of the data, thus improving the overall value of the platform.



Seven laws of information — why data decentralization is profitable

Law of information	How Storecoin Platform enables it
5/ The value of information increases when combined with other information.	See above.
6/ More is not necessarily better.	The platform focuses on value creation with laws 2, 4, and 5. The focus is not on quantity of data produced.
7/ Information is not depletable.	See 1, 4, and 5.

Structured data is *discoverable* data

Unstructured data is not discoverable and hence is not usable. If the data is not usable, it has no value. So, data *discoverability* is the core tenet of Storecoin Platform.

Examples of data discoverability in Storecoin Platform:

```
https://apis.storeco.in/api.version/data/shape?categories
```

Anyone can query for available data categories available on Storecoin Platform.

```
https://apis.storeco.in/api.version/data/shape?category=<category1>
```

Users can query for the details of one or more categories (what the data looks like, which app provides this data, how much it costs, etc.)

```
https://apis.storeco.in/api.version/data/shape?category=<category1>&app=<appID>&class=data_tier1
```

Users can query for the details of a specific data tier for a specific app.

App developers categorize their data into an existing category, special case an existing category, or create a new category for their data. The Platform APIs automatically make these categories visible across the network.

Appendix

Data valuation methodologies

Different methodologies to value data:

- **Cost-based:** value is determined based on how much the asset cost to create. This method may be highly imprecise for data, because data is often created as an intermediate product of other business processes. Cost could be estimated based on cost of storage and other data infrastructure, but this may not capture the full value of the data nor reflect the differences in value between vastly different data sources.
- **Market-based:** value is defined based on the market price of comparable goods on the market. In most cases, comparable data sources may be non-existent. A market would require a consistent notion of what makes data more or less valuable. Moreover, even if there were a well-formed data market, data sources rarely have the same content or quality.
- **Expected future value-based:** value is defined based on an estimate of future cash flows to be derived from the asset. This approach may be useful for valuing data for a very specific use.

No data valuation methodology is perfect - Full Descriptions

Each valuation methodology has its benefits and drawbacks. Sometimes, the data price is determined based on what the market determines. In other cases, it is based on individual evaluation.

Indicator	Description	Benefits	Potential Drawbacks
Indicators based on market valuation			
<i>Financial results per data record</i>	Aggregated market cap (revenues, or net income) of a company divided by the total number of personal data records used by this company.	<ul style="list-style-type: none"> - Relatively easy to identify. - Reflects actual economic value added generated through personal data. 	<ul style="list-style-type: none"> - Likely to be inaccurate, as numerous other components impact market cap / revenues / income of a company. - Possible synergy effects could lead to overestimates for firms with larger datasets. Appropriateness of this approach depends on what portion of turnover is directly tied to personal data.
<i>Market prices for data</i>	Price per personal data entry offered on the market by data brokers.	<ul style="list-style-type: none"> - Relatively easy to identify; - Reflects market value of a given, specific data entry. 	<ul style="list-style-type: none"> - Apart from the data value, it includes the cost of data search and processing. It also neglects the context in which the data is sold, which has a large influence on the demand (and price) for data.
<i>Cost of a data breach</i>	Economic cost of a data breach (for firms and individuals) per data entry.	<ul style="list-style-type: none"> - Reflects a real market value and a portion of the risk that companies must protect against. 	<ul style="list-style-type: none"> - Captures market costs of damage caused by data breach rather than value of data themselves. Does not include the costs of damage to a firm's reputation.
<i>Data prices in illegal markets</i>	Estimation of prices of personal data (per data entry) in illegal markets.	<ul style="list-style-type: none"> - Reflects market value of a given, specific data entry. 	<ul style="list-style-type: none"> - Difficult to measure and only applies to the context where the data is used again to obtain other benefits illegally. Because criminals must balance the risk of detection and punishment, the value of the personal data is likely undervalued by such an approach.

Indicators based on individual (data subjects') valuation			
<i>Surveys and economic experiments</i>	Valuation of personal data in monetary terms are reported / revealed by individuals in surveys / economic experiments.	<ul style="list-style-type: none"> - No ambiguity in data identification. - Captures the pure economic value of personal data from an individual perspective. - Results usually can be used for comparative studies (across economies and across various types of data). 	<ul style="list-style-type: none"> - Hypothetical value not verified by the market. Previous research shows that a person's valuation of their own personal data is highly sensitive to context, meaning that the way questions are phrased could significantly alter the responses.
<i>Individual willingness to pay to protect data.</i>	Amounts that individuals are ready to spend to protect their personal data.	<ul style="list-style-type: none"> - Captures the pure economic value of privacy from an individual perspective. 	<ul style="list-style-type: none"> - Captures individually perceived aggregate costs of damage caused by data breach, rather than value of data themselves.

Source: <https://www.oecd-ilibrary.org/docserver/5k486qtxldmq-en.pdf?expires=1552263628&id=id&accname=guest&checksum=2E0EF5B6557470927E8ABD6FFC347DDDE>

