Using Blockchain Technology and Artificial Intelligence in geospatial data sharing

Elio Decolli (1), Kejdi Lleshi (2), Elva Leka (3), and Luan Arapi (4)

(1) Universiteti Politeknik i Tiranes, Fakulteti i Gjeologji-Minierave, Albania(eliodecolli@gmail.com), (2) Universiteti Politeknik i Tiranes, Fakulteti i Gjeologji-Minierave, Albania (klleshi97@gmail.com), (3) South East European University, Faculty of Contemporary Sciencesand Technologies, Tetovo, Macedonia (el23618@seeu.edu.mk), (4) Universiteti Politeknik i Tiranes, Fakulteti i Gjeologji-Minierave, Departamenti i Gjeoinformatikes, Albania (luan.arapi@fgjm.edu.al)

ABSTRACT

Lately Blockchain Technologies as well as Artificial Intelligence have received considerable attention from many researchers and government institutions. There are many articles circulating around the internet about howboththese technologies can be applied on a geoscientific scenario. In this article we will discuss a way to implement aplatform on which scientists can share their work/study in a secure and efficient manner.Using the blockchain to prevent two of the major flaws in scientific research as well as data sharing and collection: ownership-rights, and based on the model described here, equal rights and a democratic eco-system for publishing your data, the last is completely automated; by doing so we also remove the need for a third-party, thus removing bureaucratic short-comings, i.e. corruption and inefficiency. Based on this we also introduce a model detailing how such platform can be programmed to mimic an economic market in order to also produce competition betweenauthors, leading to better studies. Artificial Intelligence is used in order to support this model by examining similarities between projects, therefore detecting potential plagiarism or unauthorized redistribution. We believe that this is a small but necessary step in order to solidify a concrete base for further development and migration of computer science with geosciences. Moreover, the model presented here can be modified and expanded in order toachieve near-real-time distribution of data, thus helping distribute and analyze potential disaster-occurrence data,as well as share previous data sets regarding past disasters information. By keeping this in mind we can integrate a smaller network inside our model, made by nodes which receive data from other participant nodes, and work towards DRM strategies. Lastly, we present a hypothetical scenario, on which the model has been successfully implemented and how the results would look like.

Disaster Risk Management Improvements Proposal

We'll tackle DRM in three main categories:

1. Risk Identification

While not every geophysical disaster is foreseeable, many other disasters can be identified or at least be taken under consideration while designing risk management strategies. We believe that overcoming bureaucratic short-comings[1], data archiving, practitioner's point of view and connecting people's experiences with the scientific community are an essential part of this process.

2. Preparedness

Evaluating "preparedness" is a process which for us should have a focus on the previous efficiency of the actual policies. We think that in order to advance our solutions, our wisdom must be "widened", this is done through a linear progress that starts from raw data, knowledge and ultimately wisdom[2], and while data is already stored and distributed in databases worldwide, its access is limited and its controlling is narrowed down to single instances in most cases. If we work towards wisdom, a community-like approach to data collection and organization is essential.

3. Resilient Reconstruction

Nowadays "resilient" is frequently associated with "adapting" and "changing" rather than "resisting" and "stability"[3], this means that practitioner's point of view and public awareness-spreading are necessary means in order to "build-back stronger". We'd also like to include developers incentives to study a specific zone for potential hazards as a mean to achieve better resilience and preparedness.

Model Overview

The model described here is based on the Ethereum Blockchain, using Smart Contracts, which are essentially a piece of code executed on a decentralized virtual machine, EVM. The main goal of this is to build a decentralized and democratic community where users are able to share their studies, projects or other materials while contributing to filter and categorize works done by validity. A deep learning model is used in order to extract the context of a document, then a Vector Space Model classifies their similiarities based on the streams of terms extracted by the deep learning algorithm. We use a method called *Cosine Similiarity* to do the latest.

$$Sim(d_1, d_2) = \frac{\sum_{i\, t \in d_1 \cap d_2}^{n} (W_{t_i \in d_1} \times W_{t_i \in d_2})}{\sqrt{(\sum_{t_i \in d_1}(W))^2 \times (\sum_{t_i \in d_2}(W))^2}}$$

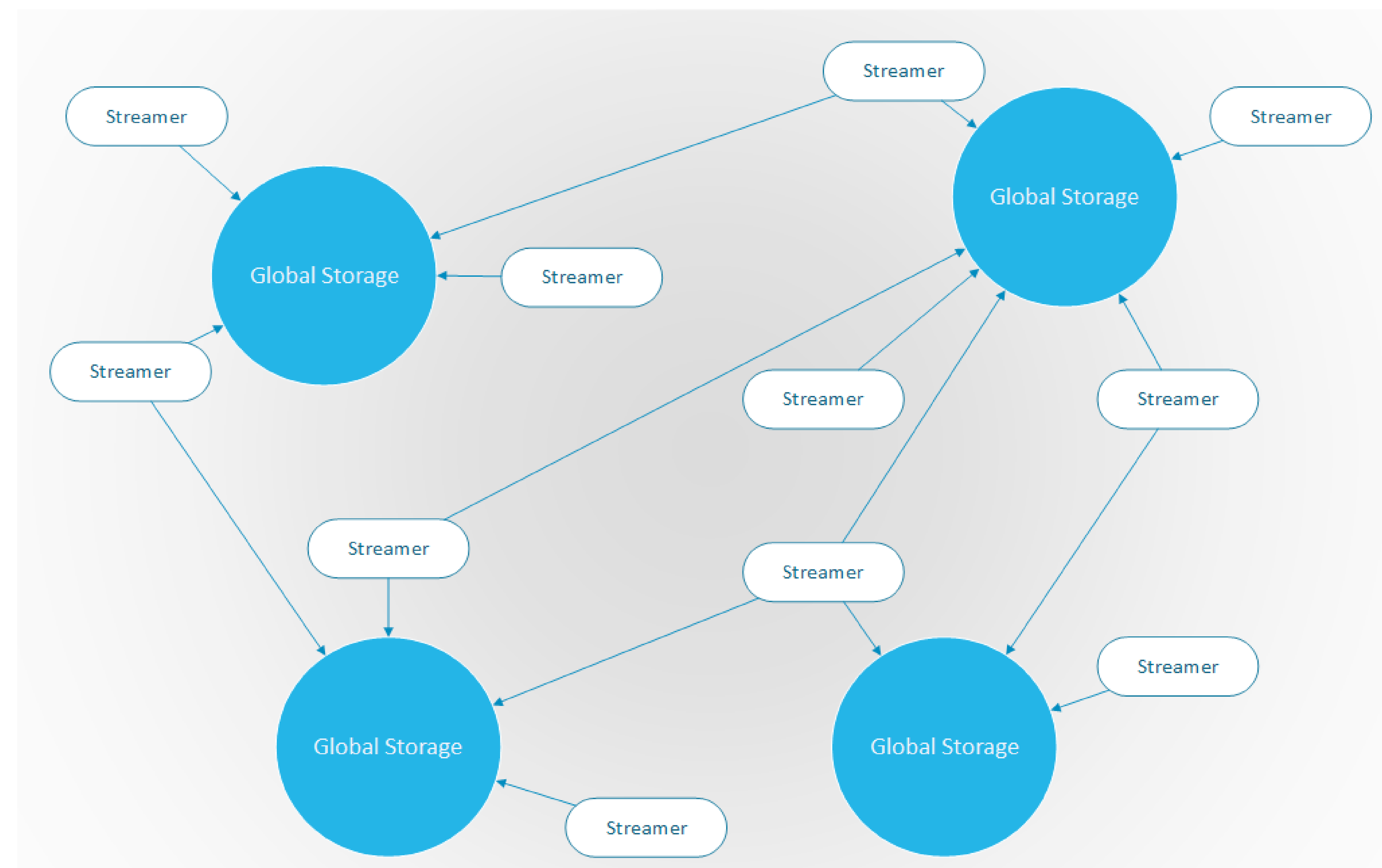

Figure 1.0. *Shows a high-level overview of the network, divided into Global Storages, every GS contains several Local Storages, each of them owned by a single Streamer. The latest are able to connect to several GS and communicate with the data being uploaded there. Note that Local Storages as well as Global Storages are represented by different Smart Contracts, so in reality they're not stored inside each other, but rather assigned somewhere inside a reserved address in the blockchain, completely decentralized.*

The network consists of three main entities:

1. Local Storage(s) - A meta-data representing the actual material being shared on the network, i.e., a geo-referenced map or a natural hazards information dataset.

2. Streamer(s) - An individual registered on the network. Streamers also can own one or more Local Storages.

3. Global Storage(s) - Sub-networks containing Local Storages sharing common attributes, such as position or type of materials uploaded. Each GS has its own rules and regulations established based on the productivity and overall contribution of Streamers accessing it.

Reputation across the network is governed by the amount of contribution given to the network and reviews of the user's work so far, moreover, we use virtual tokens in order to measure this reputation, the more tokens a user has the more influential his/her vote is on the Global Storage.

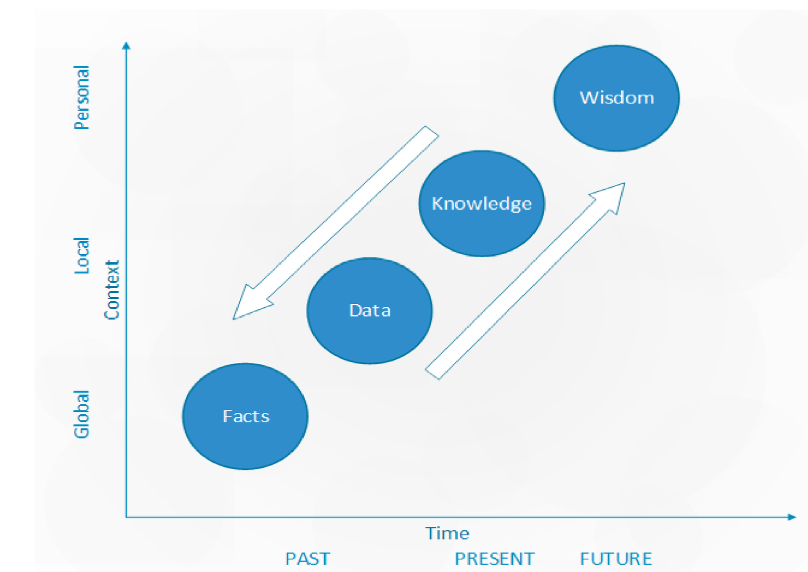

Figure 1.1. *The continuum of understanding. Shows the relation between time and context applied to data and wisdom[1].*

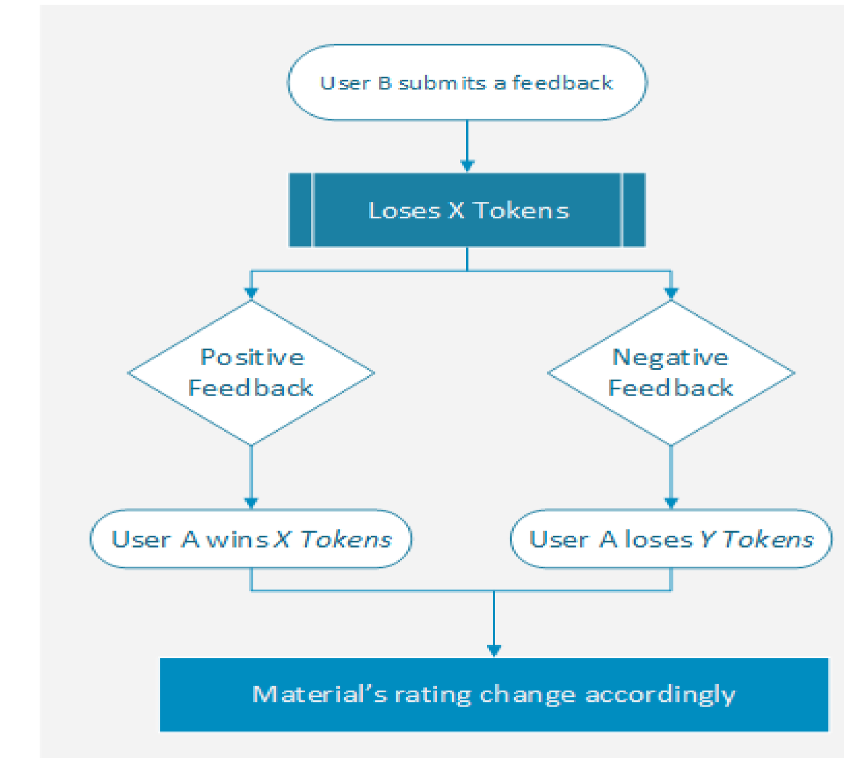

Figure 1.2. *A simplified internal operation when a user reviews another user's material. Tokens are being exchanged accordingly based on the type of feedback received on the uploaded data.*

Conclusions

Even though the model presented here is still a work in progress and the ideas behind it are being tested out, our main goal is to create a solid platform which will act as a backbone for future communication between the scientific community, people and government agencies. We believe that by solidifying and merging different disaster reduction approaches while taking advantage of modern technologies such as the Blockchain and Artificial Intelligence, we can take the next step to a better and more informed community where people and users can better learn from past mistakes.

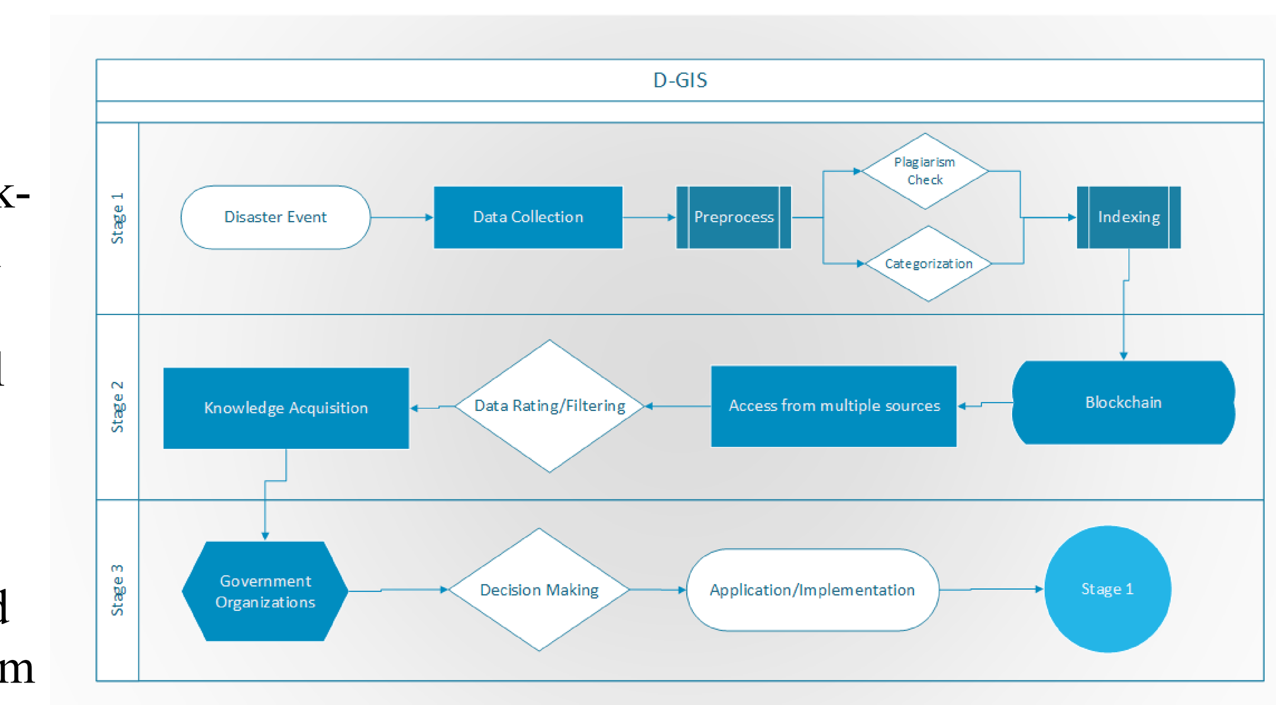

Figure 2.0. *Shows a real-time working model which connects scientists, regular users, and government agencies in order to test out, review and contribute towards a better DRM approach based not only one a single instance but on different practitioners' point of view.*

References

1. HUPE, P., & HILL, M. (2007). STREET-LEVEL BUREAUCRACY AND PUBLIC ACCOUNTABILITY. Public Administration, 85(2), 279-299.

2. Weichselgartner, J. & Pigeon, P. Int J Disaster Risk Sci (2015) 6: 107. https://doi.org/10.1007/s13753-015-0052-7

3. Aldunce, P., Beilin, R., Howden, M., & Handmer, J. (2015). Resilience for disaster risk management in a changing climate: Practitioners' frames and practices. Global Environmental Change, 30, 1-11. https://doi.org/10.1016/j.gloenvcha.2014.10.010