Supporting Information

Information Decentralization and Concatenation: A Mobile-based Bidirectional Traceability System for Sustainable Food Supply Network

Kai-Yuan Lin[1*], David Chavalarias[2,3*], Maziyar Panahi[2], Tsaiching Yeh, Kazuhiro Takimoto, Masaru Mizoguchi[1]

Publications : Lin, K., Chavalarias, D., Panahi, M., Yeh, T., Takimoto, K., Mizoguchi, M., 2020. Mobile-based traceability system for sustainable food supply networks. *Nature Food* 1–7. https://doi.org/10.1038/s43016-020-00163-y

[1]Laboratory of Agro-informatics, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 7 Chome-3-1 Hongo, Bunkyo City, Tokyo 113-8654, Japan

[2]Complex Systems Institute of Paris Île-de-France, CNRS

[3]Centre d'Analyse et de Mathématiques Sociales, EHESS

*Corresponding authors: Kai-Yuan Lin, Tel: (+81) 030-5841-1606, E-mail: rgpstu15@kmd.keio.ac.jp ; David Chavalarias, Tel: (+33) 09-7263-7922, E-mail: david.chavalarias@iscpif.fr

Supplementary material

Supplementary material 1: Technological supplementary information

(A) P2P network

Our system is built upon the GUN, which is a fast data sync and storage system which it is fully decentralized (peer-to-peer or multi-master), meaning that changes are not controlled by a centralized server (learn more about GUN at: https://gun.eco/). Traditionally, Data Base Management Systems (DBMSs) are centralized servers belonging to one organization to manage all data. In contrast, P2P DBMS utilizes a decentralized network of servers around the world instead of huge central servers owned by a large organization, which allows all data to be distributed automatically throughout the entire network. Building a decentralized system, the information in which can be completely trusted, is the obvious development pathway for the industry30. All of the separate communications and the entirety of the data is encrypted to ensure safety and elimination of any man-in-the-middle attacks.

Due to this feature, even in the offline scenario, the user can temporarily retain the data and share it with other peers after returning online, therefore even in the unstable network environment, it can still complete transactions for food and goods. Information will never disappear as long as the peer does not permanently disconnect from the network. For versatility, we propose a mobile-based solution instead of servers. As a normal peer running in a web browser, they will not start downloading the whole database but only the parts they request. Data will sync it to other peers without needing a complex consensus protocol. The peer application is built with JavaScript and has P2P connections via WebSocket. The WebSocket protocol reduces Internet communication overhead and provides efficient, stateful communication between Web servers and clients31. The whole database is considered to be the union of all peers' graphs.

(B) Graph database

Shown as supplementary fig. S1, the prototype uses GUN DataBase (DB) as a storage system (Code refers to https://github.com/amark/gun). It's a graph database stored across all peers participating in the network, thus on the client's devices. By using a graph database, the relationality provided between data can be used to track from a specific point toward the upstream (producer) side or downstream (consumer) side, and reveal the entire traceability chain instantaneously. With conventional databases such as RDB (Relational DataBase), huge overhead of the reading and writing side of the data query means the data contains many relationships requiring joining of large tables32. It is extremely complicated for RDB, resulting in an inefficient structure. We propose a methodology to convert a relational to a graph database by exploiting the schema and the constraints of the source33. There is no theoretical limit for the total size of a GUN DB. The amount of data that a peer has locally available is limited by the memory constraints of the host environment, like operating system, browser, etc., therefore the amount of data that can be persisted beyond the running process depends on the storage engine. Central storage is used for bootstrapping the initial state of DB and caching. As a super-peer dedicated GUN peer running on NodeJS with Radix file, storage can persist with much bigger amounts of data. They will receive all data that is broadcast to the network and be able to retrieve it on request. For user authentication, we use the same GUN DB SEA (Security, Encryption, Authorization) framework, which allows having P2P identities, instead of server-based authorization. The platform is compatible and easy to deploy into public Clouds such as AWS, Azure, or Google Compute Engine to allow seamless scalability and agile development.

(C) 2D barcode

We propose two-dimensional barcodes for product identification and media for traceability which allows users to store valuable information both into the platform and into the barcode itself, which provides traceability back to the source.

They are geometric patterns in two dimensions which can store more data than one-dimensional barcodes while using the same or smaller space since they can store data in both vertical and horizontal directions34. In addition, two-dimensional barcodes usually include error correction codes. If some parts of the two-dimensional barcode are deformed or damaged, it remains possible to decode the barcode correctly. Generally, two-dimensional barcodes contain black squares on a white background. Currently, two-dimensional barcodes are becoming very popular because they feature both increased capacity to store data and increased reading speed.

Supplementary material 2: Protocol of choice for consumers

The organization of most of the world's major economies is based on the belief that perfect consumer information and the unconstrained collective behavior of competing sellers will lead the systems to the best state in terms of offering consumers with the ideal price for each good. This traditional, and now obsolete, conception of economics ignores the fact that information on products is often scarce, expensive or even sometimes impossible to obtain. This results in imperfect information for consumers, a situation known to lead to significant market failures[1].

Besides, the information access issue is not only a matter of cost: the very definition of what is a relevant information for a consumer is not stable over time and across consumers. While the prices, the freshness or the life expectancy of products are often mentioned as a relevant information, the set of dimensions along which some consumers might need some information in order to choose what to buy is undefined: it is impossible to foreseen in advance what will be needed. Let's take for example food traces in products. Nuts are among the most frequent allergens so that their presence, even in trace amounts, is now indicated on most products. But there are much more situations where even traces amount of a product could lead to severe health problems for somebody in the population. Not to mention religious beliefs that require some arbitrary goods not to be in contact with some others even if they do not appear in the same end-product. It is impossible to cover all the information needed by any individual in the population on a product's paper label.

The question of food composition is only one example of what someone might need in order to buy a product according to one's preferences or ethics. We can observe that consumer concerns have extended broadly, even to the non-food producers, which are related to the supply chain. For example, consumers can be concerned about process and production that attributes organic, natural, genetically modified and origin; about environmental and sustainable issues such as the use of energy, resources and the emission of Green House Gases (GHG), or about packaging and recycling ; as well as about some social issue such as consumer justice, animal and human welfare and agricultural operation[2,3,4,5].

These issues cannot be addressed with traditional food and products labeling but can easily be addressed with future extensions of our P2P system. Although it is not the place here to expose such system in details, we would like to give the broader picture that could be built upon our traceability system.

The open nature of the protocol will also allow a wide variety of actors to propose new applications and interfaces to easily subscribe to different recommenders data, access to them when needed and aggregate their information and judgments. One of the best available option to implement such aggregation is the *majority judgment* algorithm[6]. The majority judgment voting method, praised by the economics Nobel Prize Kenneth Arrow[(1),7], makes it possible to rank items that have been evaluated by a population with mentions similar to those proposed above, while addressing the well-known shortcomings of standard methods, such as plurality voting. Last, it is important to give a specific statute to the occurrences of at least one 'to reject' mentions among the subscribed recommenders in order to spot cases where an aspect of the product is incompatible with the consumer's lifestyle.

Supplementary fig. S3 gives an example of a distributed Label Enrichment System that complies with the three requirement mentioned in the discussion. It articulates four different types of actors around a common protocol for consumer good labeling enrichment:

1. A community of contributors providing evidence-based enrichments of products labeling. This step is reinforced by the traceability system we have proposed in the first sections, which considerably broadens the spectrum of evidence-based enrichment.

2. A community of recommenders that publish recommendations about products with some arguments referring to facts. The recommenders might use one or several facts proposed by the contributors. They can themselves provide their own fact about product as contributors and refer to them in their evaluation. Last, they give a global judgment on the product on the scale of the open protocol: 'to reject', 'unsatisfactory', 'fair', 'good', 'very good', 'excellent' ;

3. An ecosystem of apps that propose different ways to aggregate information and display the details. We recommend to use the majority judgment method[6] to aggregates the multi-criteria evaluations that steam from the different recommenders but some apps might propose other methods. To summaries roughly the majority judgment method, all evaluations from the different recommenders are collected and ranked in increasing order

---

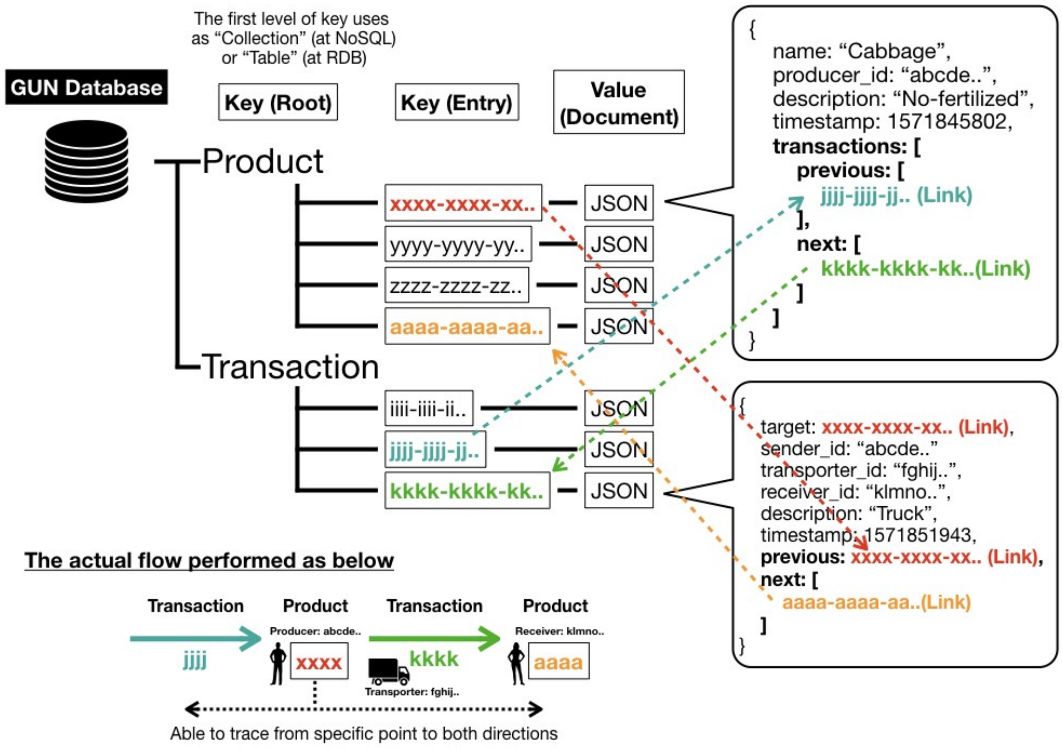
1 ()Arrow won the Nobel Prize in 1972 for his impossibility theorem, a social-choice paradox illustrating the flaws of ranked voting systems. It states that a clear order of preferences cannot be determined while adhering to mandatory principles of fair voting procedures. In his endorsement of Balinsky and Laraki book, he highlights that the Majority Judgment methods remedies this paradox.

(ex. Fair - recommender 3), good - recommender 1, good - recommender 2, excellent - recommender 4), and the mention attributed to the product is the median mention (here 'good'). The detailed procedure guaranty to be able to rank all the products, event in case of ties.

4. A population of consumers who choose the most appropriate app for their needs and subscribe to their favorite recommenders. They can get instant feedback about the compatibility of a product with their values and ethics by scanning its barcode.

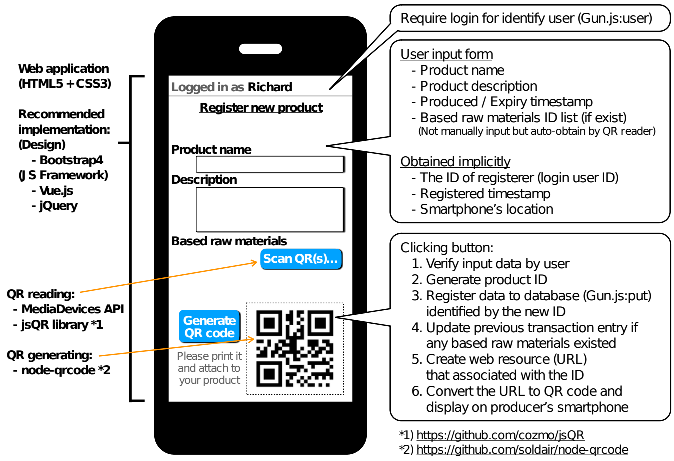
Additional References

1. Stiglitz, J.E., 1979. Equilibrium in Product Markets with Imperfect Information. *The American Economic Review* **69**, 339–345.
2. Aung, M. and Chang, Y. Traceability in a food supply chain: Safety and quality perspectives. *Food Control* **39**, 172-184 (2014).
3. Golan, E., Krissoff, B., Kuchler, F., Calvin, L., Nelson, K., & Price, G. Traceability in the U.S. food supply: Economic theory and industrial studies. *Agricultural Economic Report Number* **830** (2004).
4. Caswell, J.A. How labeling of safety and process attributes affects markets for food. *Agricultural and Resource Economics Review* **27**, 151-158 (1998).
5. Nelson, P. Information and consumer behavior. *Journal of Political Economy* **78**, 311-329 (1970).
6. Balinsky, M., Laraki, R., Majority Jugement - Measuring, Ranking, and Electing, MIT Press. (2010).
7. Arrow, K. J. A difficulty in the concept of social welfare. *Journal of political economy* **58**(4), 328-346 (1950).

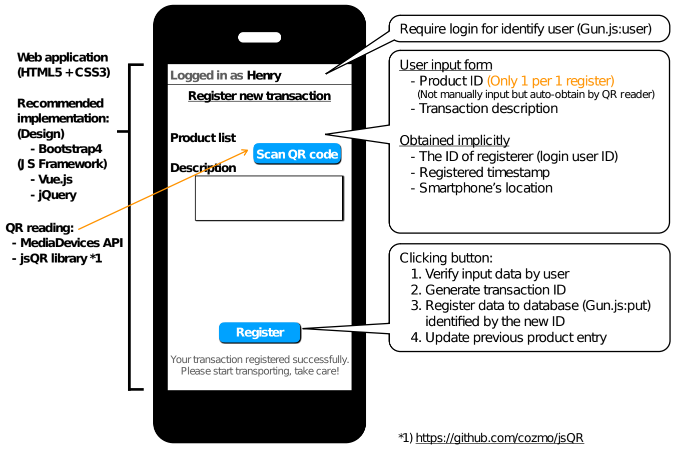


Supplementary Fig. S1 Database Structure with GUN ecosystem.

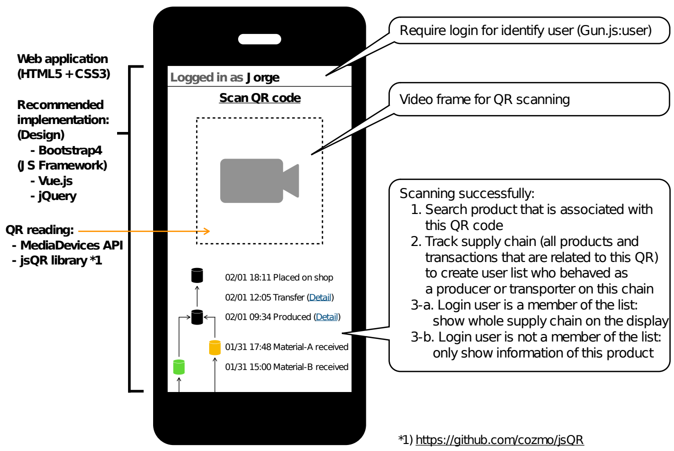
(a) For producer



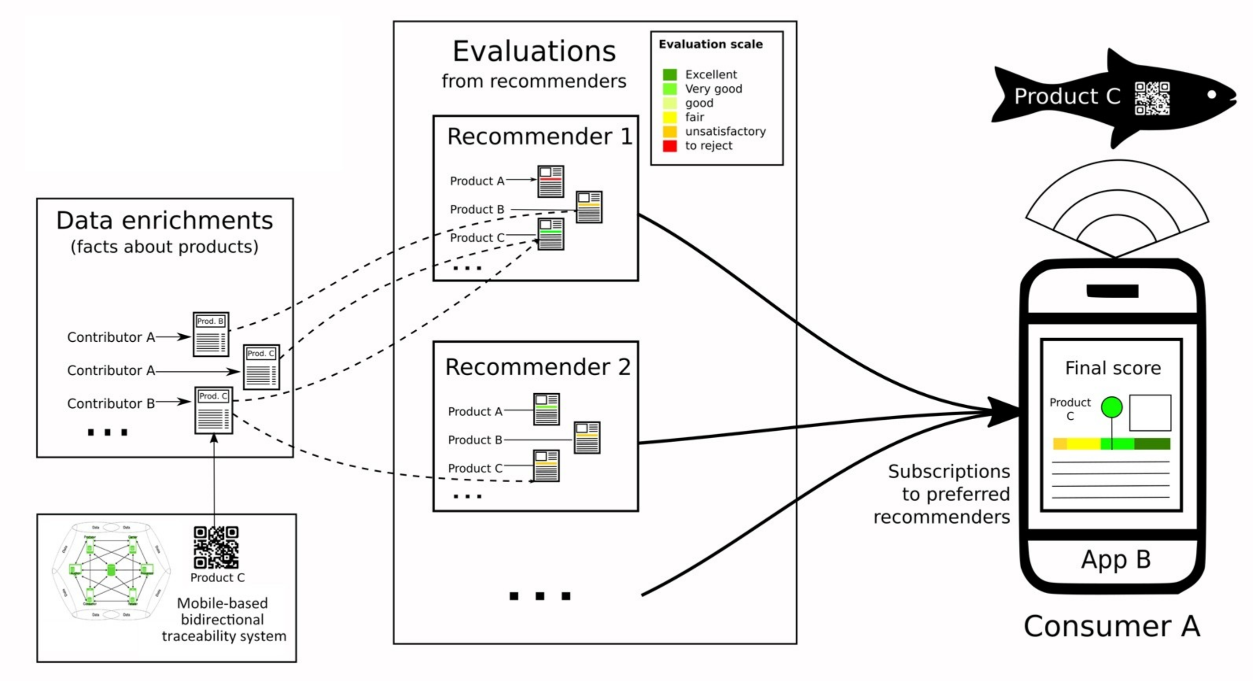
(b) For transporter



(c) For all (Scan Product QR to track)



Supplementary Fig. S2 Three user interface on the personal devices with suggested development language and functions.



Supplementary Fig. S3 Example of a decentralized and participatory Label Enrichment System. It coordinates four different kinds of actors: (1) Contributors to evidence-based products labeling enrichment. This step is enhanced by the traceability system we have proposed in first sections; (2) Recommenders who publish recommendations about products with some arguments referring to facts from contributors. They can themselves be contributors; (3) App providers. Apps makes it possible to subscribe to recommenders and aggregate of their evaluations. They allow their users to subscribe to one or several recommenders and get aggregated information and evaluations about the products they scan. Here, the majority judgment method is used to give a score (the green circle); (4) Consumers who choose the most appropriate app and subscribe to their favorite recommenders. Consumers can themselves be contributors or recommenders. In this example, consumer A uses the app B to scans the product C. She has subscribed to recommenders 1, 2 and some others. Recommender 1 uses products C facts from contributors A and B, while recommender 2 only uses products C facts from contributor B. Contributor B uses the Mobile-based traceability system to establish his facts.