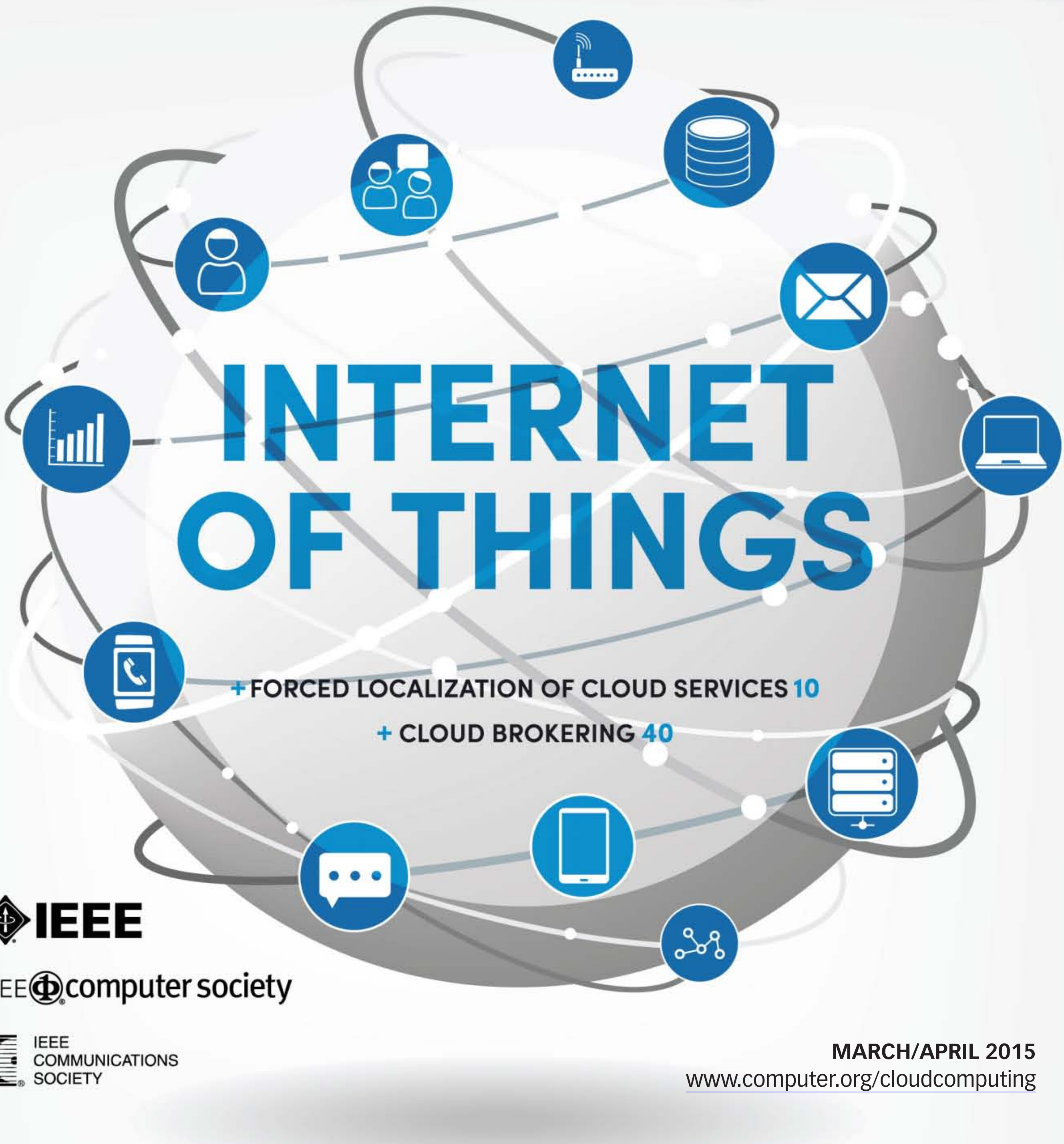
IEEE CLOUD COMPUTING



INTERNET OF THINGS

+ FORCED LOCALIZATION OF CLOUD SERVICES 10

+ CLOUD BROKERING 40

IEEE

IEEE computer society

IEEE COMMUNICATIONS SOCIETY

MARCH/APRIL 2015
www.computer.org/cloudcomputing

CLOUD ECONOMICS COLUMN

Data as a Currency and Cloud-Based Data Lockers

Omer Rana
Cardiff University

Joe Weinman

THE BIGGEST INFORMATION TECHNOLOGY TRENDS OF THE LAST FEW YEARS ARE THE RISE OF THE CLOUD, THE EXPLOSION OF BIG DATA, GLOBAL NETWORK UBIQUITY, SOCIAL MEDIA, AND THE EMERGING INTERNET OF THINGS. These aren't independent trends. It's estimated that 90 percent of all data ever created was generated in the last two years.[1] This massive increase in data is largely due to phenomena such as social sharing of videos over mobile networks and ambient video monitoring devices backed up in the cloud. A key accelerator of these trends has been consumer cloud and technology adoption—such as the exponential growth of Google search, Facebook, Twitter, and Instagram. Although these applications are largely based on free and freemium models, their growth is creating challenges for all participants in the ecosystem.

Monetization strategies that are apparent to consumers include virtual goods, in-app purchases, and advertising-supported models, as the increasing number of pop-ups, pop-unders, banners, mobile, text-based, in-game, and other ad types makes clear. Such ads can be used not only to sell products and services directly, but also to position brands and political viewpoints. A less obvious monetization strategy is the use of customers' personal data—data entered into forms, email text and attachments, the increasingly extractable data in images and videos, data from devices such as DNA sequencers or in-home "things" such as dishwashers and refrigerators, and mobile device sensors. In some cases, such data is used to better target advertising, with real-time auctions used to sell the impression opportunity to the highest bidder. But in other cases, the data is sold through data brokers, and consumers have little visibility into who has acquired their data or for what purpose. By one estimate, such personal data is valued at more than 300 billion euros, and this number is forecasted to treble by 2020.[2]

Ownership of this data is often buried in lengthy terms and conditions, or is less obvious than one might think. Human DNA, microbiome, and even weight or cholesterol levels would seem to be among the most personal data of all, yet hospitals, medical researchers, and DNA repositories argue that they have property rights to both cell tissue and the abstract DNA information.[3]

As is often the case, technology can both bring about new challenges and help solve them. These challenges—and opportunities—include fine-grained ownership of data, data portability, extensible (and retractable) rights to the data, opportunities for third-party cloud-based intermediaries to act as escrow agents, and the possibility of customers directly monetizing their own data through the use of virtual currencies.

EDITOR
JOE WEINMAN
joeweinman@gmail.com

2325-6095/15/$31.00 © 2015 IEEE

Opportunities for Cloud-Based Intermediaries

Given this context of the emerging "value" of data being generated through cloud-based services, we continue a key discussion point considered elsewhere—that is, whether data can be the basis of a new transactional relationship between people and companies in which both sides benefit from new products and services and increased economic growth.[1] However, our key distinction from previous discussions is whether the existence of a global cloud computing industry (resident in datacenters located in different parts of the world) can be used to facilitate such transactional relations, with awareness of data privacy and access management.

Currently, consumers of cloud-provisioned services don't fully see how providers use their data or aggregate it with data owned by other providers. Transferring and selectively migrating their own data stored in the cloud—in the context of using social networks, search engines, fitness tracking, and map search services, to the use of energy consumption and retail/purchasing patterns—should become an essential user capability supported by cloud providers. Hence, the availability of a cloud-based "locker," a key premise of the Locker Project (http://lockerproject.org), would allow individuals to aggregate data about themselves and release it based on their own interests. Such data might be physically held by different organizations and at different locations, but an individual or company could create an overlay to enable different datasets to be aggregated into a "virtual" container. The individual would control access to the container.

Figure 1 illustrates the general concept of a "virtual data container" implemented through an overlay network across different datacenters. Such a virtual data container provides a user-based view of data relating to and directly managed by an individual.

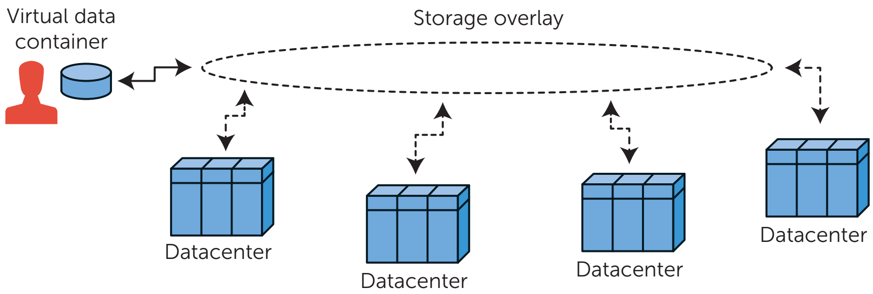


FIGURE 1. Virtual data container with encrypted content aggregated across multiple providers.

There is already significant interest in the use of "data as a currency," focusing particularly on how individuals can selectively use their personal data to acquire services. This emerging paradigm is based on the viewpoint that by providing sampled, preselected data, individuals can receive services for free from a potential service provider in return for giving them their data. The greater the quantity and frequency of data given to such a service provider, the greater the expectation (or duration) of capability/features expected in return. The service provider can use this data to fine-tune its own service provision, whether it involves selecting additional features to be made available or understanding deployment options (that is, the type of hosting infrastructure it'll likely need to offer its service in the future), often by aggregating and correlating data from a variety of consumers (both spatially and temporally). Although this type of provisioning already exists, whereby service providers offer a "lite" version of their service for free to consumers in return for acquiring statistics on usage patterns, the ability of data generators/owners to more actively participate in this ecosystem opens up significant additional potential for monetizing such data.

If data is a currency that individuals can use to purchase services from others, the value associated with the data will depend on the individuals generating the data and the devices they use to do so. For instance, someone using a personalized fitness monitoring device with more advanced capabilities (that is, it can monitor a greater number of parameters and at a greater frequency with higher accuracy) could request more services than someone using a more limited capability device. In this way, we envision the establishment of a consumer-device-data sharing economy, which could also involve the ability to store and archive this data in a cloud system. Individuals using devices with a higher bandwidth connection and support for real-time streaming can give a service provider access to their data locker at a higher potential value than individuals who use a batch mode to transmit data. This data generation and sharing economy could also act as a catalyst for creating a real-time data harvesting and management service ecosystem that consumers could embed within their own personal devices.

Ensuring that data remains portable and can be used by multiple providers remains an important requirement in such a marketplace. Identifying data standards to enable exchange would reduce exclusionary and exploitative behavior of particular providers and would prevent consumers from being locked into certain service-provision relationships. Given data management's history, it's unlikely that one or two standards will emerge or be adhered to universally. Instead, we consider the establishment of a data translation/mediation ecosystem as services that could be used directly by a provider or consumer. Such services would enable translation between data formats and could be made available as cloud-based services. As a particular data format became more popular, for instance, providers would be more likely to host such a translation service (compare this with the wide availability of PDF translators available today for documents). Such a translation service could also include, for instance, additional capability such as acquisition, cleaning, integration, cataloguing, and search over the data—enabling a service of this type to be made available at different prices depending on the functionality offered. The "data as a currency" marketplace could therefore also lead to the establishment of new companies able to respond to this data translation challenge. Companies could play a focused role (for example, acquisition or sale of data) or vertically integrate. They could offer services as SaaS providers, as PaaS or microservices providers (such as validating financial transactions or authenticating users), or sit on top of existing IaaS providers.

Complementary Currencies

Considering this from another perspective, the notion of "commodity money"[4] and the associated concept of "complementary currencies" have proven to be useful instruments for facilitating economic regeneration. Complementary currencies are recognized by their ability to add value to a community, and to improve the scalability of trade beyond bartering. In the context of a local economy (that is, where trading is restricted to participants that live nearby, and doesn't involve global entities such as central banks), goods or services can themselves represent tradable objects. Complementary currencies generally describe a group of currencies designed to be used alongside standard currencies. They can be valued and exchanged in relationship to national currencies but also function as a medium of exchange on their own. Complementary currencies have also found applicability in peer-to-peer (P2P) networks as powerful instruments to promote exchange, particularly to avoid the need for accessing central servers. Complementary currencies have different rates of exchange and scopes of circulation. The value of a complementary currency can change over time, and in most instances is also related to the value of some real resource (such as commodities—gold, oil, or services). The relationship of such a currency to a service is particularly interesting, because the currency's value is based on the time required to perform a service in hours, notwithstanding the service's potential market value. Such a currency has primarily been used for mapping human time to perform a task into an economic value in a local context. Bitcoin is a recent example of a complementary cryptocurrency (https://bitcoin.org). Bitcoin provides a peer-to-peer method of transferring money electronically, allowing users to send payments worldwide almost instantly. Because it operates without central authority or banks, it's decentralized, with all transactions carried out by the Bitcoin network itself. The network isn't controlled or owned by anyone in particular, and is based on the consensus of all participants.

Complementary currencies can also devalue over time (for example, through the use of negative interest) to stimulate market exchange, thereby encouraging greater participation in the market. Devaluation also prevents storage of wealth (hoarding) and encourages spending as the value continues to decay. Other experimental complementary currencies use high interest fees to promote heavy competition between participants, and the removal of wealth from long-term wealth-holding structures (natural/material wealth, property, and so on) to aid competitive innovation and better use of resources across various members of a community.

Many of the ideas in complementary currencies relate to work by Silvio Gessel in his *Natural Economic Order,* which refers to "free money" as an "instrument of exchange . . . to facilitate the exchange of goods."[4] Gessel identified the notion of a national currency office, which would involve collaboration between distributed (regional) currency offices able to understand the demand for local services within a given area, without the need to establish a central bank able to support a local currency through gold reserves. One key aim of free money was to facilitate local exchange, enabling participants who were geographically colocated to use local services. Kenji Saito modelled the concepts outlined in Gessel's work in the Internet WAT (iWat, a variation on the spelling of watt, as it was originally used in an energy context; www.lietaer.com/2010/05/the-wat-system-in-japan) system.[5] According to Saito's theory, each participant can generate a debit note identifying, as creditor, the provider offering a service. In return, the creditor can exchange the debit note with another to carry out a trade. A user's ability to issue a debit note and the value it holds within an exchange depends on both the reputation of the provider generating the debit note, and the likely need for the service that such a user can offer within a marketplace. Creditors are therefore only able to accept a debit note if they have some trust in the other party's ability to deliver what has been identified in the debit note, which can include the provision of another service in return or the provision of some physical good. In this way, it's possible to conceive of a network of debtors and creditors able to use services from each other within a certain time period (in Gessel's model, the amount of free money in circulation decreases 5 percent annually through depreciation, enabling any surplus to consume itself automatically, without direct involvement of the currency office).

Data as Free Money

It might therefore be possible to use data as a complementary currency, along the same principle as the iWAT system model. Because any consumer can generate data, the "value" associated with this data will be determined by demand. The data can be used to acquire services from a provider, the quality and capability of which will depend on the value associated with the data by the provider. The value will be influenced by demand for that data at that point in time, and the reputation of the consumer who generated the data. Demand in this case could be associated with the number of potential providers interested in the type of data being provided by the user (that is, for aggregation with data from other consumers, perhaps in a particular sector). Reputation, in this context, would be a multiattribute characteristic based on the identity of the consumer and the capability of the computational infrastructure the consumer uses to generate data. Over time, the "value" associated with particular data would depreciate—more recently generated data being of greater value than older data.

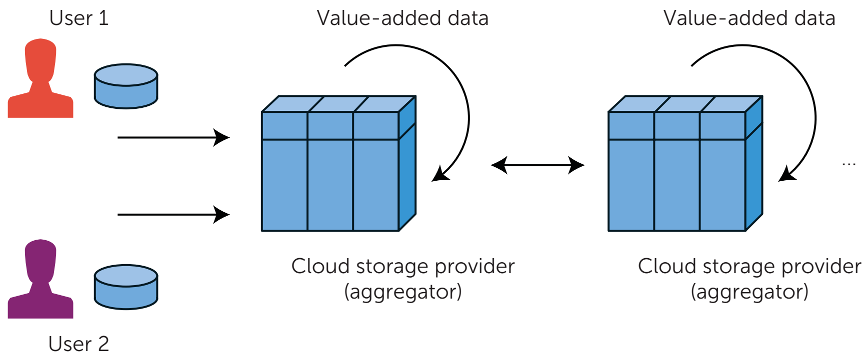


FIGURE 2. Data capture chain in which users generate data and selectively place it with a cloud storage provider. The provider can analyze all users' data in the aggregate to generate new, value-added data.

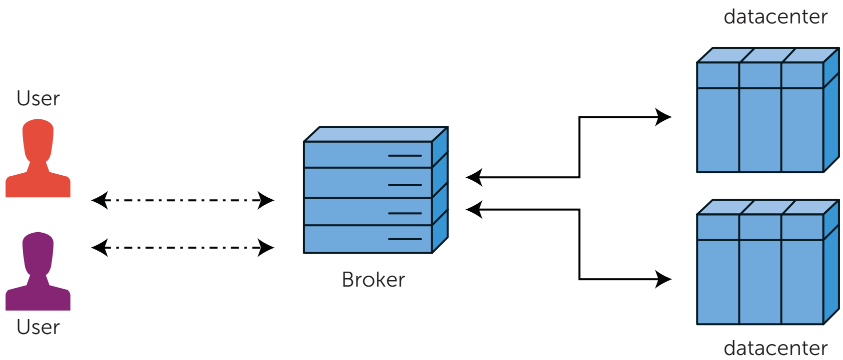


FIGURE 3. Interaction in a data economy. Brokers act as intermediaries between users and providers, and help both parties determine the value of and need for particular types of data.

Figure 2 illustrates the types of actors who could participate in such an exchange, with the users generating data and selectively placing their data at a cloud storage provider. Because the storage provider can view data from multiple users, it can perform general aggregate statistics and correlation on this data, leading to the addition of value-added data, which itself has value and can be exchanged with other cloud storage providers in a value chain.

Another role within such an economy is that of a broker, which can communicate with both users and datacenter providers. The broker could perform several roles, such as taking metadata from users and returning a valuation to them based on current demand from datacenter operators. The broker could also enable users to register interest in providers interested in particular types of data, allowing users to decide which data they want to make public.

Figure 3 illustrates the relationship between users, a broker, and multiple datacenter operators. As with cloud service brokers, these brokers would act as

CLOUD COMPUTING | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

intermediaries between users (offering data) and service providers (interested in acquiring data). However, rather than enable discovery of potential cloud services, such brokers would support transactions over data (primarily facilitated through metadata). For instance, a broker could support currency translation or determine the instantaneous worth of a user's data. Such a broker could offer a publish/subscribe event-based infrastructure to enable those interested in particular types of data to register their interest (based on topics or content), and those interested in offering data to infer potential demand. This would, in many ways, be similar to existing cloud brokering companies, such as PlanForCloud.com and Cloudharmony.com.

> The use of data as a currency opens up the potential to support new types of interactions between users and cloud providers.

THE USE OF DATA AS A CURRENCY OPENS UP THE POTENTIAL TO SUPPORT NEW TYPES OF INTERACTIONS BETWEEN USERS AND CLOUD PROVIDERS. This relationship enables users to place greater confidence in how the cloud provider uses its data and, at the same time, to monetize their data. Gessel's concept of a "free money" system (as implemented in the iWAT) provides a useful basis to further investigate how such data owned by an individual could be valued in such an economy. Users would then, essentially, have their own currency (depending on the type of data they're able to generate and make available to others), which they can exchange for other data sources or services from providers. This would create value chains that spread across users, device providers and manufacturers, cloud providers, and potentially a number of third-party companies offering data translation/mapping and brokering roles.

References

1. S. Taylor, "Data: The New Currency," *European Voice*, 2014; http://3kck0t38mmqdnrjfp23kzm9tz0.wpengine.netdna-cdn.com/wp-content/uploads/2014/06/Telefonica_research_paper.pdf.
2. European Data Protection Supervisor, *Privacy and Competitiveness in the Age of Big Data: The Interplay between Data Protection, Competition Law and Consumer Protection in the Digital Economy*, Mar. 2014; https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2014/14-03-26_competitition_law_big_data_EN.pdf.
3. M. Knight, "Who Owns Your DNA? It's Not Who You Think," Genetic Literacy Project, 9 Sept. 2014; www.geneticliteracyproject.org/2014/09/who-owns-your-dna-its-not-who-you-think.
4. S. Gessel, *Natural Economic Order*, translated by P. Pye, Peter Owen Ltd., 1958, part 4; www.silvio-gesell.de/neo_index1.htm.
5. K. Saito, "Peer-to-Peer Money: Free Currency over the Internet," *Proc. 2nd Int'l Conf. Human.Society@Internet* (HSI 03), LNCS 2713, Springer, 2003, pp. 404–414.

OMER RANA *is a professor of performance engineering in the School of Computer Science and Informatics at Cardiff University. His research interests include high-performance distributed computing, data analysis/mining, and multiagent systems. Rana has a PhD in neural and parallel computing from Imperial College (London University). He is a member of IEEE. Contact him at ranaof@cardiff.ac.uk.*

JOE WEINMAN *is the chair of the IEEE Intercloud Testbed executive committee. He also serves on the advisory boards of several technology companies. He has been awarded 21 patents in areas such as homomorphic encryption, pseudoternary line coding, adaptive bandwidth schemes, Web search, and distributed storage and computing, and is the author of* Cloudonomics. *Weinman has BS and MS degrees in computer science from Cornell University and the University of Wisconsin-Madison, respectively, and has completed executive education at the International Institute for Management Development in Lausanne.*