New Cloud Atlas

JOSEPH DAHMEN
University of British Columbia

BEN DALTON
LEEDS Beckett University

AMBER FRID-JIMENEZ
Emily Carr University

Access to information is a vital element of a free society.[1] However, contemporary information infrastructure is largely constructed, operated, and maintained by major corporations. These multinational entities have a reach at once more extensive and less transparent than the national governments that have traditionally provided other forms of public infrastructure.[2] The New Cloud Atlas (http://newcloudatlas.org) is an open-source participatory platform that maps and catalogues the physical assets of the cloud—warehouse data centers, Internet exchanges, connecting cables and switches—in an open and accessible way, rendering the physical infrastructure of information open to public scrutiny. The project is modelled on the International Meteorological Organization, which at its founding in 1873 was one of the first transnational entities engaged in the global exchange of information.

THE FIRST CLOUD ATLAS

Prominent meteorologists and government delegates from around the world gathered in Vienna, Austria during the Fall of 1873. The group, which comprised the first International Meteorological Congress, sought to lay the groundwork for a global network of observatories, equipment, and standards to gather and share meteorological data,[3] as documented in the opening address of the meeting:

> If there be any branch of science in which work on a uniform system can be especially useful and advantageous, that branch is the inquiry into laws of weather, which, from its very nature, can only be prosecuted with a hope of success by means of very extensive observations embracing large areas, in fact, we might almost say, extending over the whole surface of the globe.[4]

The Permanent Committee of the first International Meteorological Congress published the International Cloud Atlas in 1896, expressing its ambitions for gathering information on a global scale. The Atlas consisted of a series of images taken from photographs and paintings that enabled observatories to amass their knowledge in a unified format so they could communicate weather data more effectively.[5] The congress planned for a network of global observatories linked by telegraph. The network would be capable of sharing consistent observations of weather systems whose scale stretched beyond national boundaries. The combination of standard communication protocols and infrastructure was an early example of a transnational approach to data generation and sharing.

The cooperation was short-lived. As colonial empires expanded, the most powerful nations controlled territories sufficiently vast to gather their own meteorological data. The onset of the First World War in Europe shifted emphasis away from international collaboration as meteorological data became a matter of national security.[6] Nevertheless, by the early years of the 20th century, the data gathering networks of colonial powers had begun to deliver on an earlier prediction by a teenage John Ruskin, who had imagined a global meteorological society as early as 1839:

> The meteorologist is impotent if alone; his [sic] observations are useless; for they are made upon a point, while the speculations to be derived from them must be on space. … The Meteorological Society, therefore, has been formed not for a city, nor for a kingdom, but for the world. It wishes to be the central point, the moving power, of a vast machine, … It desires to have at its command, at stated periods, perfect systems of methodical and simultaneous observations; it wishes its influence and its power to be omnipresent over the globe so that it may be able to know, at any given instant, the state of the atmosphere on every point on its surface.[7]

The global weather data gathering networks of the late 19th century, which preceded the advent of Big Data by one hundred and fifty years, anticipated the centralization of command over ubiquitous data collection networks, hinting at the power of controlling the collection, archiving, processing and visualization of data.

DATA CLOUDS

One hundred and fifty years later, we live in a cloud of information. The world's data increases ten-fold every five years.[8] Data is processed, broken down, archived, and repackaged, moving across distributed networks. 'The cloud' is available everywhere to contemporary users fortunate enough to have access to information networks. Indeed, the ease with which information moves seems to render data placeless: the cloud is invisible territory that is simultaneously everywhere and nowhere.

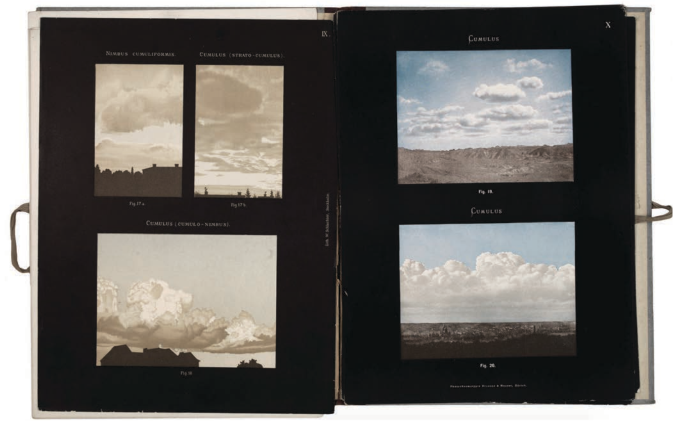


Figure 1: A spread from the *International Cloud-atlas*, 1896.

However, processing and storing this vast sea of virtual information requires physical infrastructure. Data centers are its primary nodes. An ever-expanding global network of these assets has arisen to meet the demand for information, fuelled by an equally large amount energy required to keep all the bits flowing freely. The energy required to operate the global network of data facilities collectively outranks the energy demand of all but five countries.[9] Inchoate as it seems, the virtual cloud is tethered to physical installations and kept aloft by massive infusions of conventional energy.

Owing to their key role in the flow of information as well as their extensive energy requirements, data centers have been the subject of territorial disputes and international conflict. The half-million data center locations spread across the globe can be read as indicators of contemporary geopolitical and economic forces. At the intersection of global politics and free-market forces, these locations are determined by a balance of favourable data regulation, access to affordable energy, and tax incentives created by municipal, state, and national governments. Computation and energy for cooling are paramount, but the physical and regulatory freedom to operate are no less important: an overly controlling government or the absence of sufficient connection infrastructure renders sites less desirable to the corporations that manage them.

As computation speeds increase and globally located markets grow, distributing these processing centers evenly across territory will become increasingly important, exacerbating extant geopolitical issues. Faster connection speeds can mean a few microseconds are sufficient to provide an advantage in trading or tracking. Calculation load can also be balanced around the world. Current data centers cater to economies of scales, favoring cheap boxes on cheaper land, but these are unlikely to be the only solution in the future. Plans are being developed for data freighters and platforms at sea[10], data drones[11] and data zeppelins in the sky.[12] Companies are already developing data centers housed in shipping containers, the standard unit of intermodal transport.[13] In the future, infrastructure designers may colonize currently uninhabitable places with data. The depths of the sea, arctic regions, offshore platforms,[14] liquid natural gas sites,[15] mountain tops, and deserts are all being explored as potential sites for data centers.

WHO OWNS THE CLOUDS?

Once the domain of national governments, information infrastructure is increasingly constructed, operated, and maintained by major multinational corporations. These corporations, which include the familiar Google, Facebook, Amazon, Apple, and Microsoft in the US, Yandex and mail.RU in Russia, and Baidu, Tencent, and Alibaba in China, have a similar vested interest in maintaining control over of the flow of goods and information once exercised by national governments. However, their reach is at once more extensive and less transparent. Keller Easterling has pointed out that many of the physical assets of information infrastructure are intentionally hidden from public view.[16] In keeping with her observation, the planning for data centers is cloaked in secrecy. Where large amounts of land are required, it is often acquired under various pseudonymous companies that help maintain the anonymity of the underwriters. Where governments run data centers, they take pains to conceal data center locations for security reasons. Unmarked buildings hum through the night. Older buildings are gutted and retrofitted with racks of switches and drives, leaving anonymous offices and factory façades intact but for obscured windows. Even as increasing access to information holds the promise of more free and open societies, physical information assets are hidden in plain sight, the better to assume control over information flows. Grasping at the physical assets, yields a vaporous handful of nothing: the cloud.

The cloud was used in early diagrams of the Internet to stand in for complexity. The Internet was designed to be ‘end-to-end’, so computers are meant to be able to connect to each other without interference as the message passes through a network of interconnections. Only the end points are meant to matter. The clouds here represent ‘something in the middle that is too confusing or complex to draw here’, a kind of neutral space through which information passes on its way out the other side. It is an act of simplification, but it also contains an implicit statement

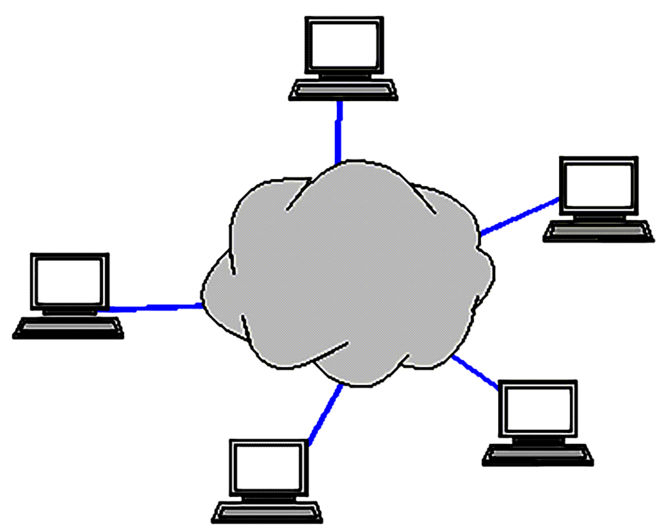


Figure 2: Schematic diagram of the Internet from the Computer Destop Encyclopedia published in 1998.

Figure 3: New Cloud Atlas workshop participants observing and cataloging cloud infrastructure at FACT Liverpool, 2016.

that 'the cloud will look after itself' — that this thing is going to carry on being there.

The use of the cloud has shifted in digital systems. The idea that the cloud is 'too complicated to think about' has been converted into a business model, shedding its innocence along the way. Companies becloud online digital systems, deliberately making them seem more confusing, in order to obfuscate or conceal a more functional description. Through a sleight of hand, the cloud appears now as a platform, later as a material. This narrative rests on the idea that the services provided are to be trusted, and they will take care of themselves. We trust them with our emails, our childhood photographs, our meeting plans, and everything in between. This new definition of the cloud asserts that 'it is too complex to deconstruct or critique'. You shouldn't try to look into the cloud and see what's there. It's made up of vapour, and it's not to be interrogated. Better to simply observe it from a distance and admire it at sunset.

CLOUD INFRASTRUCTURES

The original International Meteorological Organization was formed to track phenomena larger than national borders. Contemporary data clouds also exceed national borders. We therefor call for a reprise of the original organization to catalyze global collaboration to monitor these emergent 'weather systems' in a systematic way: the New Cloud Atlas. A system that declares it is not to be deconstructed has the making of a worthwhile challenge. Our model of the digital cloud is as ephemeral, ethereal, intangible, and aspirational as the original. Although the virtual cloud will remain forever beyond reach, investigating the infrastructure that sustains it lends a different impression. Viewing a Google data center through Google Maps satellite imagery suggests a giant warehouse. This is infrastructure not on a human scale, but rather post-industrial infrastructure, rapid construction surrounded by mud and dirt.

A typical data center is a simultaneously dynamic and lifeless machine. Like an organism that constantly replaces its cells with new growth, it is in a process of constant replacement of hardware and improvement of technology. Nevertheless, each facility is operated by just a handful of technicians. The central architectural volume of each warehouse is an undifferentiated storage space for hard drives and processors. It contains vast amounts of stored data, but also acts on that data, processing it by taking the information and doing the things that make it more useful, searching, reformatting, interpreting. The large sections on the sides of a data warehouse are generally dedicated to cooling. Processing information produces heat, a phenomenon familiar to anyone who has ever placed a laptop on their lap. That heat has to be dissipated, and so the unseen layer on a Google data center map is the sheer quantity of energy channelled in to this system as electricity and pumped back out as heat. Other than the data itself, the production of vapour from the heat exchangers is their only tangible output.

Data centers invite interrogation of their architectural form, which is notable in its disregard for human ergonomics. Fire is suppressed by injecting with carbon dioxide or other gaseous agents into the space, which avoids damaging sensitive servers with water but threaten human life. Data mechanics and access to cheap energy replace accustomed patterns of physical behaviour. Cool climates are especially sought after, so that passive sources of air or water can help cool equipment, reducing power consumption and improving the economics.[17] The connection speed provided by national or local infrastructure competes with tax breaks and potential political instability when locating data facilities. They are often sited by hydroelectric dams or in old warehouses that have had consumed large amounts of power historically.[18] Google's Dalles data centers is next to a hydroelectric dam which used to serve an aluminium smelting facility before it closed down.[19]

The second significant physical infrastructure of the cloud are its switches, known as Internet exchange points. An anonymous looking switch building at 111 Eighth Avenue in New York may serve as an example. The building contains many interconnections. There are many data lines running underground into the basement, and various organizations and companies agree to connect their networks to each other within the building. The building and the connections within is a reminder that the 'Inter-net' really is a connection of many networks, and switch buildings are where they join, at these key infrastructural points. Though limited in number, a vast quantity of information passes through these narrow constrictions.

The final element of cloud infrastructure is the connections themselves. Ironically, much of the cloud is actually underground, buried as cables. Attempting to bury a cloud reads like an ancient myth, and the technology used to accomplish it dates back to the prehistory of telegraph lines. Sail a ship between continents, spooling a cable into the sea astern as it travels. When the line connects one continent to another, bury it in the mud at each end. Voilá: the major connections of the Internet.

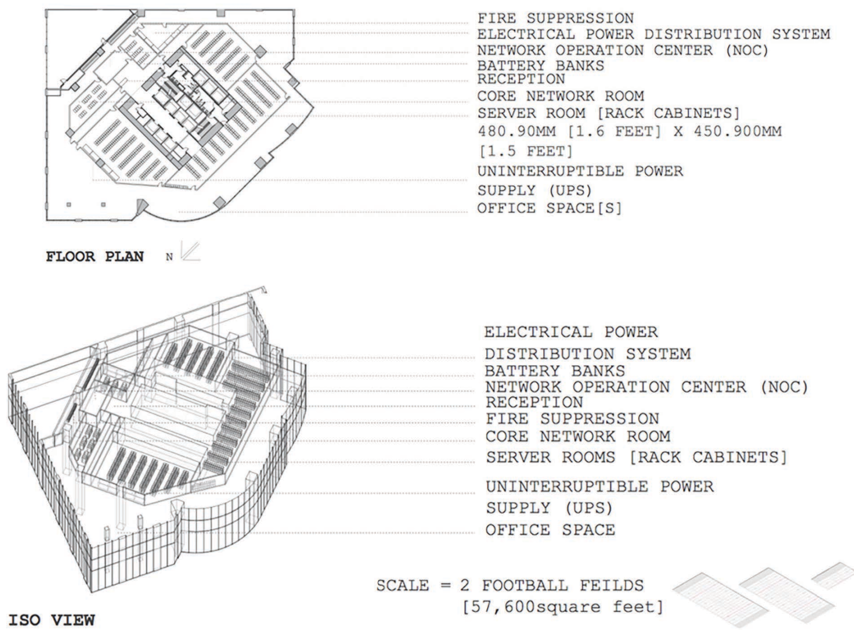


Figure 4: Architectural analysis of a data center in downtown Vancouver, Canada

Multinational corporations currently control a significant amount of data infrastructure, and the information flows that require it, with minimal oversight. The half-million data center locations spread across the globe are a barometer of global politics and free-market forces, determined by a balance of favourable data regulation, tax incentives, sufficient connectivity, and access to abundant and affordable energy. The physical and regulatory freedom to operate is paramount; overly controlling government regulation renders sites less viable. Territorial disputes, international conflicts, and perceived threats have led corporations to shroud the locations of information infrastructure in secrecy. Even as increased access to information at all levels holds the promise of more free and open societies, the physical assets required for the access to information are hidden from view, enabling the possibility of large scale control over access to information flows.

NEW CLOUD ATLAS

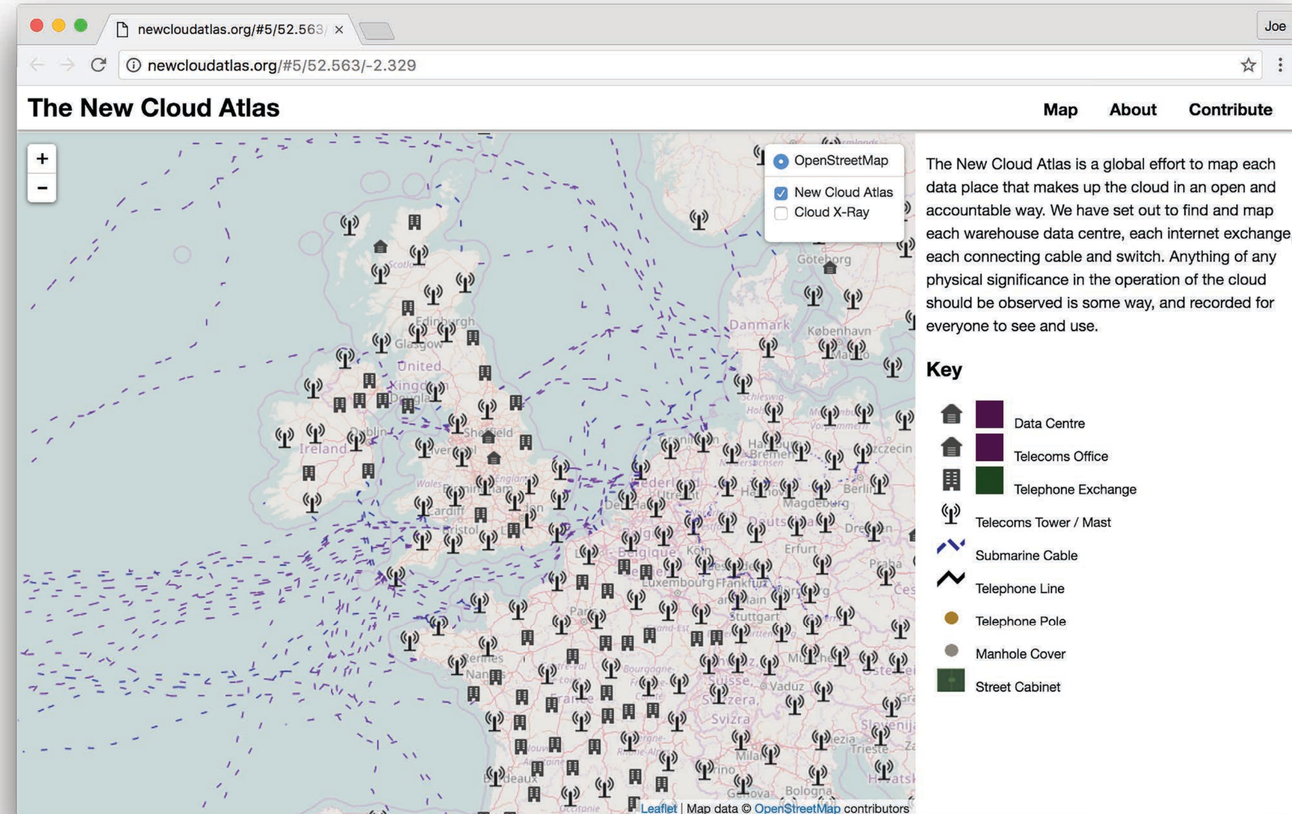
Taking the International Meteorological Organization as its inspiration, the New Cloud Atlas responds to the current control over access to information infrastructure by mapping and cataloging the physical assets essential to the flow of information. The New Cloud Atlas is a an open online platform that enables participants to map anything of any physical significance in the operation of the cloud— each warehouse data center, each Internet exchange, each connecting cable between those points. Using the OpenStreetMap (OSM) community, the mapping tool creates shared protocols within the the OSM platform in conjunction with an Open Data Commons Open Database Licence. Contributors can add to the New Cloud Atlas using a customized OSM editor. Changes made to database are parsed and imported into the New Cloud Atlas in fifteen minute intervals. All New Cloud Atlas data are available to developers who want to build on the resources, ensuring that the work of mapping the global cloud infrastructure is kept openly accessible.

Capturing the cloud poses difficult forensic challenges. For example, power consumed by data centers is a useful parameter of the Cloud Atlas, but power consumption has been classified as a corporate secret in some states within the United States, making this data hard to come by. Taking a page from history, the New Cloud Atlas responds to these challenges through crowd-sourcing. In its early days, the International Meteorological Organization enlisted fishermen living along the remote coast of Norway to help fill in gaps between observation stations. The individual contributions from participants in far-flung locations were added to those generated by official meteorological posts to create a complete picture, in a manner akin to a picture perceived from an aggregation of individual dots in a halftone image. The New Cloud Atlas invites people from across the world to develop a new form of dérive based on the physical movement of information.

The effort to map the physical infrastructure that makes up the cloud and the network is the start of a process of deconstruction and critique necessary to the production of meaning and value).[20] The scale of the undertaking is global and relies on a diverse and distributed network of citizen participants who collect information from many different sources. Companies are often secretive about aspects of their cloud infrastructure, citing security concerns as the reason for obfuscating locations of infrastructure. While we acknowledge that security risks are real, the risks of leaving potential bottle-necks and weaknesses in our digital commons undocumented and unaccounted for are greater. We maintain that the physical infrastructure of information is a matter of public concern and should be treated as such.

FUTURE DIRECTIONS

The current New Cloud Atlas platform maps the physical infrastructural elements of the cloud. In the future, the project will be extended to



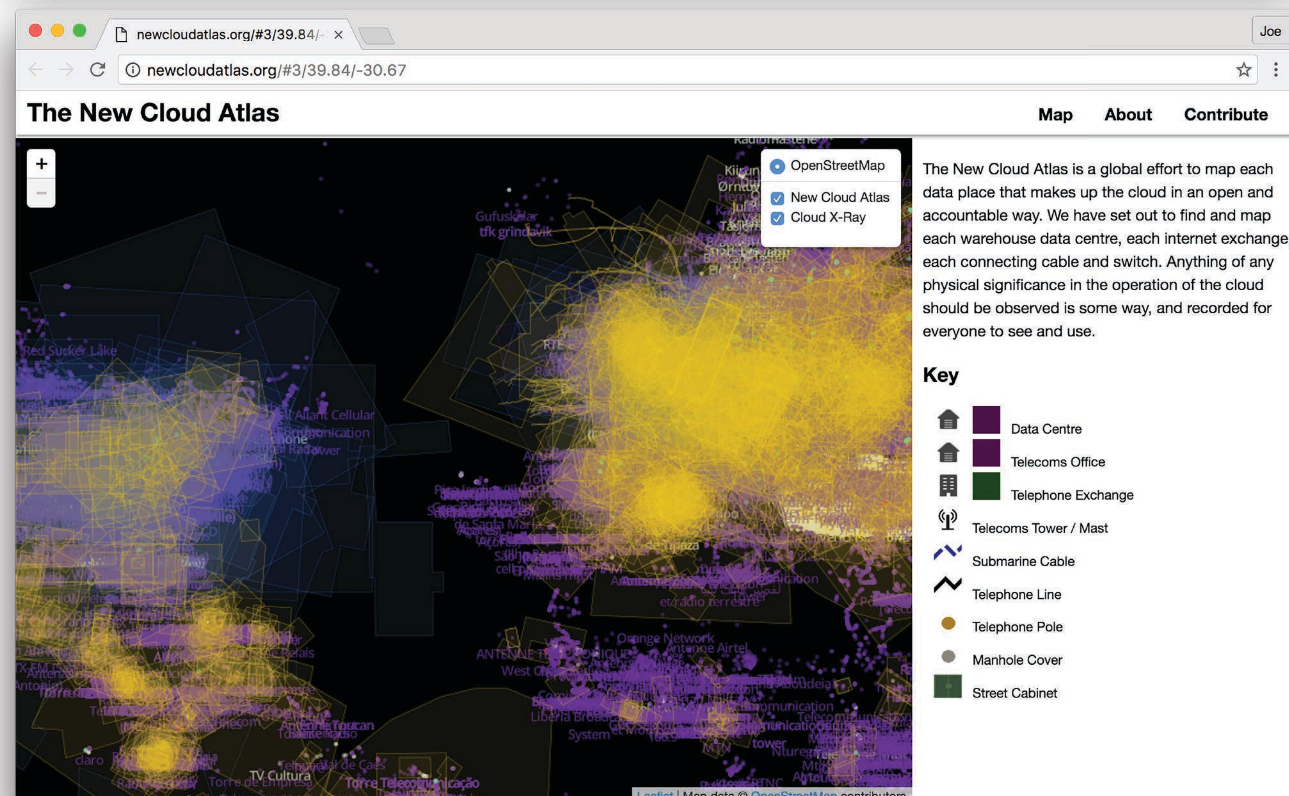


Figure 5: Screenshots from newcloudatlas.org

include its inchoate qualities. Foremost among these will be the energy use. Current estimates indicate that the infrastructure of the Internet requires approximately 30 billion Watts of electricity.[21] In individual terms, one average size data center requires 10^15 Joules of energy in a year to process information, which is about the same amount of energy contained by one cloud in a thunderstorm. As the scale of the virtual cloud grows, we can imagine a data center consuming one thunderstorm of energy per month, or per week, a level that may affect public welfare and should be the subject of public debate. Including metadata on energy use in the New Cloud Atlas will contribute to enabling public debate on this issue.

Future versions of the New Cloud Atlas will also include experiential and phenomenological attributes of the cloud, such as the building materials, the acoustical quality of the hum of the servers, and the relationship of the infrastructure to its architectural context. The effects of infrastructure on public space, and the opportunities it presents, has been the subject of much comment and debate.[22,23] The future of the project will extend beyond cartography to build a comprehensive understanding of the physiccal and discursive forces of the various valences of the facilities that create the cloud.

The New Cloud Atlas was designed and developed by AFJD and Ben Dalton with programming assistance provided by Tim Waters. The project has been exhibited as part of *Information Everything*, curated by Gillian Russell & Katherine Gillieson at the Concourse Gallery, Emily Carr University of Art and Design in Vancouver, British Columbia in 2016, and at the Foundation for Art and Creative Technology (FACT), Liverpool, UK in 2016. An earlier version of the project was presented at the Sensuous Knowledge Symposium, at the Bergen National Academy of Art & Design in Bergen, Norway, in 2013.

ENDNOTES

1. Chomsky, Noam (2003). *Chomsky on Democracy and Education*. C.P. Otero, Ed. New York: Routledge Falmer.
2. Easterling, K. (2009) “Cable” in *New Geographies 1: After Zero*. Stephen Ramos and Neyran Turan, eds. Cambridge, MA: Harvard Univerity Press.
3. Luedecke, C. (2004) “The First International Polar Year (1882-83): A big science experiment with small science equipment.” *Proceedings of the International Commission on History of Meteorology* 1.1 p, 55.
4. Bruhns, C. Wild, H. and Carl Jelinek (1873). “Invitation“, Meteorological Committee, *Report of the Proceedings of the Meteorological Conference* at Leipzig, London,E. Stanford, 5-9, p. 5.
5. Sarukhanian, E.L.; Walker, J.M. (2004). “The International Meteorological Organization (IMO) 1879-1950,” in: *Proceedings of the International Seminar to Celebrate the Brussels Maritime Conference of 1853: An Historical Perspective of Operational Marine Meteorology and Oceanography Under the High Patronage of HM King Albert II of Belgium*, Residence Palace, Brussels, Belgium, 17-18 November 2003 [CD-ROM]. JCOMM Technical Report, 27: pp. 1-8
6. Luedecke, C. (2004) “The First International Polar Year (1882-83): A big science experiment with small science equipment.” *Proceedings of the International Commission on History of Meteorology* 1.1 p, 55.
7. *Quarterly Journal of the Meteorological Society* (1881). London: Edward Stanford, Vol VII. p. 72
8. Cukier, K. (2010). “Special report: Data, Data Everywhere.” *The Economist*. Feb. 25.
9. Glanz, J. (2012). “The Cloud Factories: Power, Pollution and the Internet.” *New York Times*, Sept. 22.
10. Carol, R. (2013, Oct. 30) Google’s worst-kept secret: floating data centers off US Coasts The Guardian (UK).
11. Kay, J. (2014) “Hurricane Edouard the perfect storm to test new data-seeking drones.” *Minneapolis Star Tribune*, September 18.
12. Wall, R. (2014) “The New Zeppelins: Giants of the Sky.” *Wall Street Journal*. July 11.10.
13. Morgan, T. (2011) “Place your data centre in a handy container: The future is modular” *The Register*. London, Sept. 15
14. CNET (2013) “Is Google building a hulking floating data center in SF Bay?” Oct 25. accessed Sept. 1, 2014 at: http://www.cnet.com/news/is-google-building-a-hulking-floating-data-center-in-sf-bay/
15. LaMonica, Martin (2014) “To Zero Out Data Center Air Conditioner Bills, Build It Next to an LNG Port.” *IEEE Spectrum*, (online) August 21. Accessed Sept .15, 2014 at: http://spectrum.ieee.org/energywise/computing/it/to-zero-out-data-center-air-conditioner-bills-build-it-next-to-an-lng-port
16. Easterling, K. (2009) “Cable” in *New Geographies 1: After Zero*. Stephen Ramos and Neyran Turan, eds. Cambridge, MA: Harvard Univerity Press.
17. Ladurantaye, S. (2011). “Canada called prime real estate for massive data computers” *The Globe and Mail* (Canada), June 22
18. Ori, R. (2013, Oct. 4) “Chicago: City of Big Servers.” *Urban Land: The Magazine of the Urban Land Institute*.
19. Beck, B. (2008, June 4) *Welcome to Googleville: America’s newest information superhighway begins On Oregon’s Silicon Prairie*. Portland, Or: Willamette Week
20. Derrida, Jacques. *Of Grammatology*. Trans. Gayatri Chakravorty Spivak. Baltimore: Johns Hopkins UP, 1976. p. 30
21. Glanz, J. (2012). “The Cloud Factories: Power, Pollution and the Internet.” *New York Times*, Sept. 22.
22. Bélanger, P. (2009) “Landscape As Infrastructure” *Landscape Journal* Jan. 1,. Vol 28, no. 1. pp. 79–95
23. Shannon, Kelly and Marcel Smets (2010) *The Landscape of Contemporary Infrastructure*. Rotterdam: NAi Publishers.