

Researchers call for support for data in the cloud to facilitate genomics research

July 9 2015

Today in the journal *Nature* prominent researchers from Canada, Europe and the U.S. have made a powerful call to major funding agencies, asking them to commit to establishing a global genomic data commons in the cloud that could be easily accessed by authorized researchers worldwide.

This would increase access to the [data](#) for researchers, reduce the time and cost associated with transferring and storing data on local servers and accelerate genomics research worldwide. Storing data in the cloud has been shown to be as secure, if not more secure, than storing it locally.

With a typical university connection it can take months to download datasets from major international projects like the International Cancer Genome Consortium (ICGC) and the hardware costs associated with storing and processing those data can also prove quite expensive.

With cloud computing a data set from a big genome project can be executed in days, at a fraction of the price.

The authors propose that funding agencies request that major [data sets](#) be uploaded into the cloud and that they pay for its long-term storage. Data would then only need to be copied once and researchers would only have to pay for temporary storage while the analysis was in progress. Access would only be provided to authorized researchers.

"Currently a great deal of valuable time and money is spent by researchers transferring data from a repository to their own preferred server, instead of easily and cheaply tapping into a global data commons whenever they need to," said Dr. Lincoln Stein, Director of the Informatics and Bio-computing Program at the Ontario Institute for Cancer Research, leader of the ICGC's Data Coordination Center in Toronto and a lead author on the paper. "We encourage a larger investment in the cloud in order to use public funds more effectively and to help accelerate the pace of genomics research."

"Having authorized access procedures in place ensures respect for the wishes of data donors, including that their data be used safely and securely," said Dr. Bartha Knoppers, Director of the Centre of Genomics and Policy, McGill University. "[Applying the Framework for Responsible Sharing of Genomic and Health-Related Data](#) is a first step in enacting the human right of citizens to benefit from scientific advances and of scientists to be recognized for their work."

"The complexity of cancer biology means that we need huge data sets - basically, the bigger the better," said Dr. Peter Campbell, Head of Cancer Genomics at the Wellcome Trust Sanger Institute. "We have now reached a stage where these data sets are too large to move around - cloud computing offers us the flexibility to hold the data in one virtual location and unleash the world's researchers on it all together."

"The amount of genomic data is growing at an amazing rate. Moving data and analysis tools to the cloud will democratize access to data and to the computational resources required to analyze that data," said Dr. Gad Getz, Director of the Cancer Genome Computational Analysis Group at the Broad Institute of MIT and Harvard. "The expanded access will accelerate tool development, grow the population of researchers analyzing these rich data sets and ultimately increase the pace of scientific discovery. These cloud-based analysis platforms will also

enable the testing of new distributed computing paradigms which expand both the scale of the analyses and the sophistication of the computational algorithms. We are now building a pilot of such a cloud platform."

"The establishment of novel powerful [cloud computing](#) frameworks enabling us to store, share and analyze data across borders will open new perspectives in cancer research," said Dr. Jan Korbel, group leader at the European Molecular Biology Laboratory (EMBL). "These will take into consideration developments in science and policies for the distribution and sharing of data sets as sensitive as patient genetic data ensuring a safe environment to serve the interests of both sample donors and researchers."

Cloud computing is most widely associated with consumer products, such as storing music, photos or editing documents in real time. But in fact a great deal of research is already conducted in the cloud, safely and securely. Cloud computing is shared resource, giving researchers access to storage and computing power as needed, instead of making a long term investment in computer infrastructure. This also maximizes the use of the infrastructure as it can be used by many researchers instead of just one.

Provided by Ontario Institute for Cancer Research

Citation: Researchers call for support for data in the cloud to facilitate genomics research (2015, July 9) retrieved 27 April 2024 from <https://medicalxpress.com/news/2015-07-cloud-genomics.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.