Ideas for genomics project: signing and identifying genome [SafeGenome\GenomeSigner].

In a series of posts tagged with ideas-for-genomics-project [http://scalaakka.blogspot.com/search/label/ideas-for-genomics-project] I am describing ideas for business that can be build in area of genomics. If you like those ideas and you would like to join as an investor, co-founder or consultant please contact me at LinkedIn [https://pl.linkedin.com/pub/artur-stanek/a5/643/691] or just drop me an email, stanek.artur@gmail.com [mailto:stanek.artur@gmail.com] .
You may be also interested to check previous post "Filter out what is not needed [http://scalaakka.blogspot.com/2015/10/idea-2-for-genomics-project-allow.html] ", next post "Sharing genome information [http://scalaakka.blogspot.com/2015/11/ideas-for-genomics-project-genome-broker.html] " or the bigger picture [http://scalaakka.blogspot.com/2015/10/ideas-for-genomics-project-diagram-of.html] .

Idea "GenomeSigner" is about signing (and later identifying) the genome file.

Let's imagine that client would like to give genome to some company which claims that the file (or any of its part) will be deleted right after finishing service like some very complicated analysis. Naturally company also guarantees that file will never leak to the Internet. Client trusts that but would like to double protect as this genome belongs to the most expensive horse at the stud.
If everything would go well client also would like to send her/his genome for analysis but as she/he is very famous actor double protection will also be needed (please find that genome protection & safety will be soon the new intimacy [http://scalaakka.blogspot.com/2016/04/genome-new-intimacy.html]).

To achieve that client opens our page, grants access or uploads the file (GenomeBank [http://scalaakka.blogspot.com/search/label/GenomeBank]), performs signing, gets file back.

Active/passive signing of the genome:

- active signing means that genome delivered by customer will be slightly modified [*] by us (we can use non coding areas without breaking any base pairs rules or any other kind of rules so modifications will be almost impossible to detect by anybody),
- passive 'signing' is just about making hash code, store it locally and/or give back to the customer, file with genome (initially delivered by the customer) will stay unchanged.

The customer can later come back, send full (or just piece of) some genome file and:

- we can say if it was signed by our company/our software,

- if yes and if it was done by the same customer then we can provide more details like when it was signed, what methodology was chosen, details of the original file,
- in case of active signing: maybe we could even revert changes and give back to customer original genome.

We can also provide active identification by scanning the Internet and when signed file (or its part) would be found we can contact customer and send full details like where file was found, date, screenshots. We can grab and send all details that would allow the customer to protect her/his proprietary in the court.

* - TL;DR, implications here are very interesting as well as very danger; as we are actively modifying source file should we also put information (that genome was modified) somewhere in that file? What if company, that had the file to make some analysis over it, will also perform wider analysis on it and leave for internal use? (Let's skip here the case if they really had a right to do that.) That means that this second analyse will be broken. What then if it will be forwarded to scientists to work on it as on regular one? What if they will produce virus or bacteria based on that (modified, manipulated) input?

Posted 3rd November 2015 by Artur Stanek (kermitas)

Labels: bioinformatics, GenomeSigner, genomics, ideas-for-genomics-project, non web find, SafeGenome