Lean GHTorrent: Github data on demand

Georgios Gousios, Bogdan Vasilescu, Alexander Serebrenik and Andy Zaidman
{g.gousios, a.e.zaidman}@tudelft.nl
{b.n.vasilescu, a.serebrenik}@tue.nl

Want to do research with GHTorrent data?
It is now as easy as:

1. Filling in the form at ghtorrent.org/lean.html

No need to care about this
(but ask if you do!)

2. Getting the data!
In the package, you will find:
- A MySQL dump (to query like a boss)
- MongoDB collection dumps (all Github API data)
  for all repos specified in step 1!