



gbif-indexingtoolkit

The GBIF Harvesting and Indexing Toolkit (HIT)

Search projects

- [Project Home](#)
- [Downloads](#)
- Wiki**
- [Issues](#)
- [Source](#)
- [Export to GitHub](#)

Search Current pages for

BackgroundInformation

Background Information

Featured

Updated Jun 11, 2009 by kyle.br...@gmail.com

Background Information on the HIT

The Harvesting and Indexing Toolkit (HIT) was originally designed with the intention of trying to simplify the lives of data aggregators (persons responsible for harvesting occurrence records from several data publishers).

The act of harvesting is in and of itself a complicated procedure that has traditionally involved sending and receiving xml messages. Even though new protocols continue to evolve that have simplified things greatly, data publishers are not likely to adopt a new protocol if an existing one is in place and working. Therefore, data aggregators are in a position where they must continue to adopt old protocols, while adapting to new ones as they become available.

Certainly then, the HIT is a tool that aims to ease the complexity of such heterogeneous networks of data publishers. In brief, it does so by shielding the end-user from the complexities of the different protocols so that only a common set of operations is ever visible to them. In that way, one only has to concern themselves with the quality and quantity of data coming back, without having to necessarily know what protocol is actually being used at particular end-points.

Ultimately, it is hoped that the HIT will allow data aggregators to manage extremely large networks of data publishers and increase the amount of data that they are capable of mobilising.

► [Sign in](#) to add a comment

[Terms](#) - [Privacy](#) - [Project Hosting Help](#)

Powered by [Google Project Hosting](#)