

GARDIAN PII checker to minimize risk of personally-identifiable information in datasets

Project Title: P1536 - IFPRI Cross-Cutting Module 1

Description of the innovation: The PII checker is available as a containerized system for easy deployment by CGIAR Centers and others. It is a service built on state-of-the-art algorithms from Stanford University to flag the presence of personally-identifiable information (currently only names), and intended to minimize risk for the CGIAR System and other entities using it.

New Innovation: No

Stage of innovation: Stage 3: available/ ready for uptake (AV)

Innovation type: Research and Communication Methodologies and Tools

Geographic Scope: Global

Number of individual improved lines/varieties: <Not Applicable>

Description of Stage reached: The PII checker is available via GARDIAN Services, and has been tested by a couple of CGIAR Centers.

Name of lead organization/entity to take innovation to this stage: SCiO - Big Data in Food Systems

Names of top five contributing organizations/entities to this stage:

- IFPRI - International Food Policy Research Institute
- ICRISAT - International Crops Research Institute for the Semi-Arid Tropics
- Bioversity (Alliance) - Alliance of Bioversity and CIAT - Headquarter (Bioversity International)

Milestones:

- 1.2.5. GARDIAN officially launched with a PII detector, a data upload functionality for partners without repositories, indexing external data from at least two partners' repositories.

Sub-IDOs:

- 43 - Enhanced institutional capacity of partner research organizations
- 44 - Enhanced individual capacity in partner research organizations through training and exchange

Contributing Centers/PPA partners:

- IFPRI - International Food Policy Research Institute

Evidence link:

- <https://hub.docker.com/r/sciogardian/piiengine>

Deliverables associated:

- D9078 - New interface and functionalities added to GARDIAN, including tools and services to mine data effectively, taking into account data privacy, ethics, and licensing issues (<https://hub.docker.com/r/sciogardian/piiengine>)

Contributing CRPs/Platforms:

- BigData - Platform for Big Data in Agriculture