An automated data collection protocol using open source applications for large sampling campaigns

Elise Lepage, Catherine Wipf-Scheibel, Marion Szadkowski, Loup Rimbaud, Karine Berthier, Sylvain Piry

HAL Id: hal-04195727
https://hal.inrae.fr/hal-04195727

Submitted on 4 Sep 2023

An automated data collection protocol using open source applications for large sampling campaigns

Elise Lepage, Catherine Wipf-Scheibel, Marion Szadkowski, Loup Rimbaud, Karine Berthier, Sylvain Piry

EL : AgroParisTech, Palaiseau, France; EL : INRAE – BioSP, Montfavet, France ; EL, CWS, MS, LR, KB, SP : INRAE – Pathologie Végétale, 84140, Montfavet, France ; KB, SP : INRAE - UMR CBGP, 34988, Montferrier sur Lez, France

Large sampling campaigns are often essential to getting insights into the spatiotemporal dynamics of epidemics at the scale of agricultural landscapes. For example, to understand the key drivers of the recent emergence of the cucumber mosaic virus (CMV) on pepper (Capsicum annuum) in the Espelette region (South-West of France), twice a year our team collects about 2500 plants in a zone stretching on a total area of 15x15 km². To accommodate the needs of such sampling campaigns in terms of data diversity and traceability, we developed an automated data collection protocol based on the open source applications OsmAnd and KoboToolBox, coupled with two-dimensional barcodes and our internal database Virobase.

Thanks to these apps, users need no more than (solid) smartphones or digital tablets for offline data tracking onsite. Indeed, based on a GPS and the cartography from OpenStreetMap, OsmAnd allows offline navigation between fields and an accurate positioning of samples. KoboToolBox provides flexible templates to create forms to record, in a step by step procedure, all relevant data relative to the samples (e.g. species, symptoms, photos, spatial coordinates). Back in the laboratory, the standardised format of KoboToolBox forms allows the automatic import (through an R script) of all data into the Virobase database, an ad hoc database developed with PostgreSQL/PostGis which includes a web2py user friendly interface. In the end, this protocol ensures reliable data tracking "from the sampling bag to the storage tube", thanks to a unique and standardised index encoded in a two-dimensional barcode attributed to each sample.