

**Position paper on VGI and the collection of natural sciences information.
Brian Klinkenberg, Department of Geography, UBC**

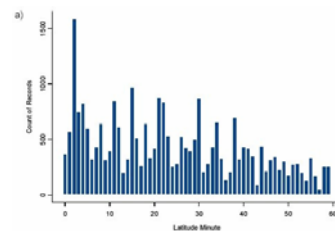
There is a long tradition of volunteer-collection information in the natural sciences, from observation-based data to geo-referenced collection data. A significant proportion of collections in natural history museums / herbaria / (in)vertebrate museums have been provided by people with no formal relation to the institution. Without those collections, and their associate collection or label information, accurate knowledge of the biodiversity of many regions would be far more limited than it is today. However, the concept of “collection” is changing today, fuelled by new technologies such as the digital camera. Documentation of new records or species can be caught on camera, and can provide a new means of vouchering, a concept that lies at the heart of museum collections. And this vouchering can be done by non-experts/scientists, lying more and more in the realm of the “volunteer”.

E-Flora BC and E-Fauna BC are electronic atlases of the flora and fauna of British Columbia (Canada) that use natural history collections to provide distributional information, and on volunteers who contribute expertise and data to build the databases and atlas pages that are key to the project, including the interactive mapping component. Developing these atlases has provided a unique opportunity to work with both historic and contemporary sources of VGI through taxonomic collections and photographs.

Traditionally, the material that is housed in a museum (e.g., a dried plant specimen or the preserved remains of an insect or mammal) is identified by an expert before being added to a collection database. However, while much attention is paid to the taxonomic aspects of the material, far less attention has been paid in the past to the description of the geographical location of the material.

Reclaiming our past—adding reliable geographical coordinates to the large number of specimens that already exist in museums and herbaria—should be one important component in any examination of the use of VGI today¹. While programs such as GeoLocate (Tulane University Museum of Natural History) have made that process much easier, significant issues exist with such automated routines since place names are far from unique and correctly deciphering the syntax of often-vague location descriptions is fraught with difficulties. Reasonably precise geographic locations can be automatically associated with some collections, but other location descriptions cannot be resolved so easily, for example: i) 35 miles S of Quesnel; ii) Deadmans Creek road, 1 mi South of Criss Creek; iii) E of Fording River, Rocky Mountains.

Collection data that have geographic coordinates recorded (data often transformed into a plane coordinate system without due attention being paid to the choice of datum) can still be problematic, as the coordinates may have simply been recorded to the nearest 5' (as demonstrated in the figure showing the latitudes derived from collections of butterflies in BC; note that spikes in the graph appear every 5' minutes reflecting the increased numbers of collections at those coordinates).



When considering contemporary sources of natural history VGI, the technical sophistication of the community of concern must be considered. Society is clearly split into several groups—those that are technically savvy, those that have embraced some aspects of technology but not all, and those that remain largely ignorant of the

¹ This is particularly important as societal attitudes towards the ‘collection’ of zoological specimens increasingly precludes the actual taking of the specimen.

technologies currently available. I would suggest that the overlap between the domain-specific groups (e.g., knowledgeable amateur botanists) and the technologically-aware groups needs to be explored in some detail. Those that are currently fully participating in VGI initiatives such as Wikimapia are a self-selected group that are fully aware of, and capable of using, all of what technology has to offer. However, based on our experience in developing the electronic atlases of BC, certain domain groups, such as naturalists, are likely to be far less technologically sophisticated. Obtaining accurate spatial information from them will require education and considerable easing of the technological hurdles currently associated with modern technological devices.

Of course, some technologies such as digital cameras and computers have widespread acceptance. Widespread acceptance does not necessarily translate into sophisticated use of the technology, however. For example, while all users of E-Flora / E-Fauna BC must necessarily use a computer, we have observed that not all are not fully conversant with scroll bars on their browsers (e.g, if the material isn't present on the screen then it will potentially remain hidden from their view).

Digital cameras have fundamentally altered the collection of natural history materials, and significant numbers of digital photos are being uploaded into E-Flora and E-Fauna. While for many species groups digital photos can provide an equivalent record to a physical specimen (e.g., most birds can be reliably identified using a photograph), for some groups a physical collection must be made—identification can only be made with the specimen in hand (e.g, most insects require microscopic inspection before they can be reliably identified to the species). And, of course, taxonomic studies require access to genetic materials, so the need for physical specimens is a constant. Overall, the level of uncertainty associated with some collection records is increasing as digital pictures replace physical collections as the primary source of biodiversity information. We have also observed that very few of our volunteers provide geographic coordinates when they upload their photos (less than 1% of E-Flora and E-Fauna submissions), so the geographic specificity of the 'collection' is also going down (noting that those that collect a physical specimen and deposit it at a museum will almost certainly provide an accurate set of coordinates for that collection).

However, these observations may reflect the age of the community involved in the project at present and the fact that for many people photographs are not yet considered collections per se, and therefore the need to include geographic coordinates is not apparent to them. Most of the people submitting digital photos to E-Flora and E-Fauna did not grow up immersed in technology, and therefore some technologies, such as GPS, have not been as actively embraced by them as they have been by younger generations. The future role of VGI in the natural sciences looks promising, however. GPS devices are now being employed in undergraduate botany classes, so future generations of natural scientists will likely be as comfortable with GPSs as they are with microscopes.

The integration of GPS into cameras (and both increasingly integrated with a cell phone) means that in the near future a geographic stamp will become as commonplace as a time stamp on digital photos. This will enhance the use of photos *as collections*. Finally, when the Barcode of Life project produces its first 'tricolor', the ability to identify the specimen to the species will reach the level of certitude associated with the GPS-derived coordinates. At that point, natural sciences VGI will truly become a reliable and accurate source of biodiversity information throughout the world. However, until that time uncertainties in the species identification will remain a problem, and the need to encourage the use of technologies such as GPS will remain paramount.