

- **Data Libraries Topic Outline:**

- *Performance: Data Formats (Andrea)*

- *Usability: Configuration/Access (Ben)*

- *Discussion (All)*

Improving Usability of Data Libraries

Ben Morgan

THE UNIVERSITY OF
WARWICK

Usability Issues with Data Libraries

- Accessed in Geant4 code via `getenv()` of environment variable pointing to the root directory of the library.
- With N libraries, this means a user of Geant4 has to set N env vars correctly just to get default behaviour.
- Though we can generate scripts to do this, it's awkward and cannot cover all config management systems.
- Versioning of data is also a bit unclear with regard to Geant4 versions and which modules use what data.

Data Library Access Proposal

- Single PATH-style environment variable for data library locations
- Get root path to dataset NAME in C++ code:
 - `string G4Data::GetPath(string NAME)`
- Implementation searches for first occurrence of “NAME” dir/file under elements of the PATH-style variable
- NB: Lots of details for discussion here! Mild dependence on format of libraries (file vs dir)

Example Behaviour

- `G4DATADIR=/mydat:/g4dat`
- `GetPath("G4F001.4")`
resolves path to
 - `"/g4dat/G4F001.4"`
- `GetPath("G4BAR1.0")`
resolves path to
 - `"/mydat/G4BAR1.0"`

```
+ - /g4dat
|   +- G4F001.4
|   |   +- dat1.dat
|   |   +- ...
|   +- G4F002.0
|   +- G4BAR1.0
+ - /mydat
    +- G4BAR1.0
```

Benefits

- PATH-style var allows multiple data libraries to be accessed, and additional paths to be added (e.g. test a new version by prepending its location) but only ever one thing to configure.
- **Could add UI commands to add/remove elements from the search path**
- **Could add a default path determined from runtime location of Geant4 libraries (remove requirement to set env var) using binreloc/GetModuleFileDir**

Gotchas

- Where does API go? Global, Processes...?
- How should data library versions be used in API?
 - GetPath(id), with id holding version, e.g. “G4Foo1.4”
 - GetPath(id, version)
 - How to make version consistent across clients of a given data library?
- Anything else?

Discussion