VOLTRON DATA

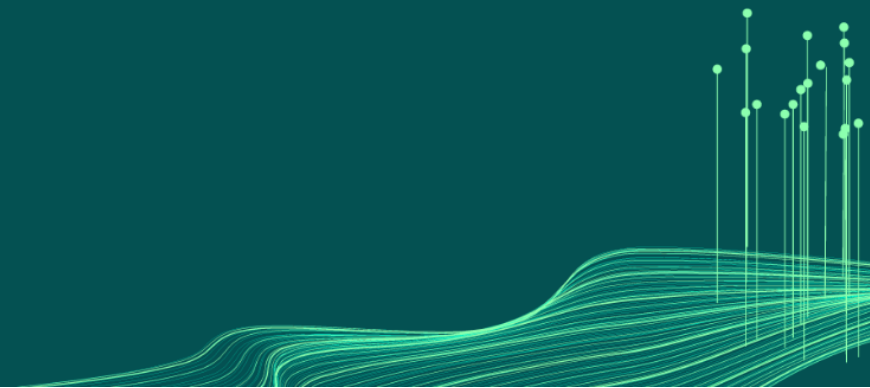
library(geoarrow)

```
geoarrow::read_geoparquet()
```

(currently ~5x faster than reading a .gpkg or shapefile in R)

```
geoarrow::write_geoparquet()
```

(currently ~5x faster than writing a .gpkg or shapefile in R)



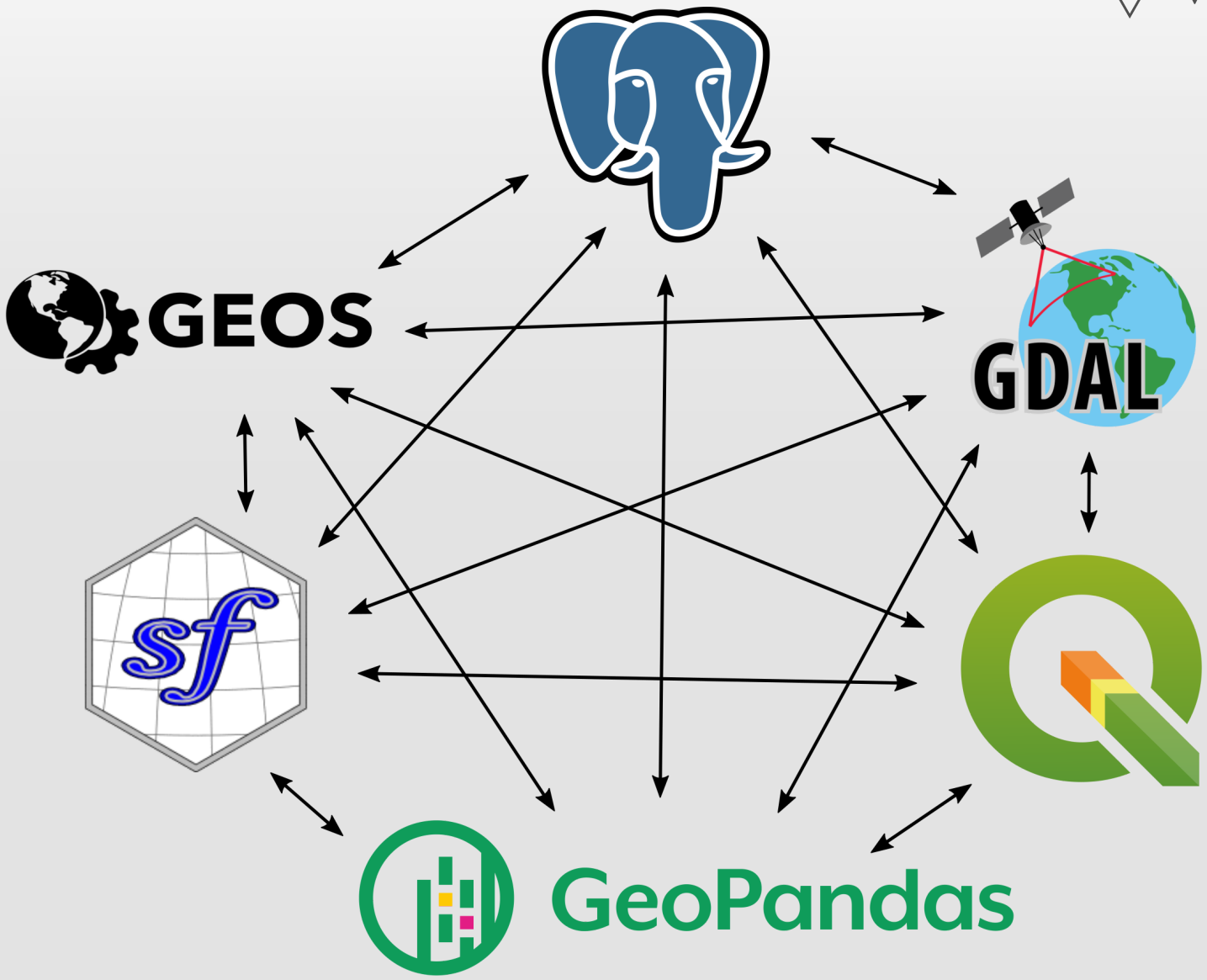
VOLTRON DATA

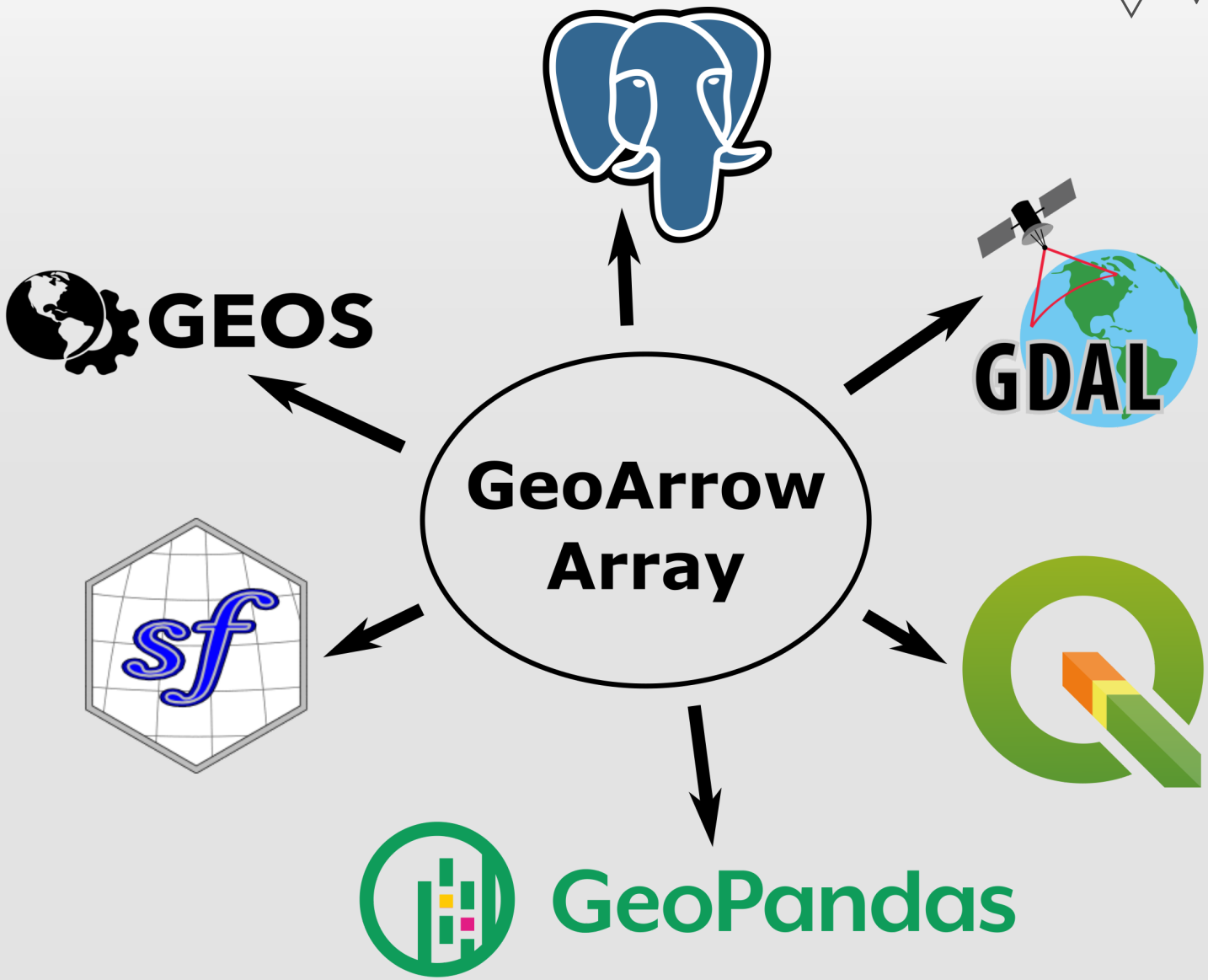
arrow::open_dataset()

```r
library(arrow)
library(geoarrow)
library(dplyr)

bucket <- s3_bucket("voltrondata-public-datasets")
ds <- open_dataset(bucket$path("phl-parking"))
ds %>%
  select(issue_datetime, violation_desc, fine, geometry) %>%
  filter(fine > 300) %>%
  head() %>%
  collect()
```

```
# # A tibble: 6 × 4
#   issue_datetime      violation_desc       fine geometry
#   <dttm>              <chr>               <dbl> <grrw_pnt>
# 1 2011-12-31 22:03:00 HP RESERVED SPACE     301 POINT (484642.5 4419268)
# 2 2011-12-31 20:52:00 UNREG/ABANDONED VEH   301 POINT (487284.6 4421182)
# 3 2012-11-07 20:00:00 UNREG/ABANDONED VEH   301 POINT (488181.6 4434546)
# 4 2012-11-12 20:00:00 UNREG/ABANDONED VEH   301 POINT (nan nan)
# 5 2012-11-07 20:00:00 UNREG/ABANDONED VEH   301 POINT (485706.1 4430103)
# 6 2012-11-07 20:00:00 UNREG/ABANDONED VEH   301 POINT (484298.8 4432491)
```


GEOS
GDAL
sf
GeoPandas


GEOS
GDAL
GeoArrow
Array
sf
GeoPandas

VOLTRON DATA

library(arrow)

VOLTRON DATA

`library(geoarrow)`

- `write_geoparquet()` and `read_geoparquet()` are fast
- `arrow::open_dataset()` + `library(dplyr)` is awesome
- Using Arrow in geospatial packages can make spatial workflows faster and easier to maintain

twitter.com/paleolimbot
github.com/paleolimbot
dewey.dunnington.ca