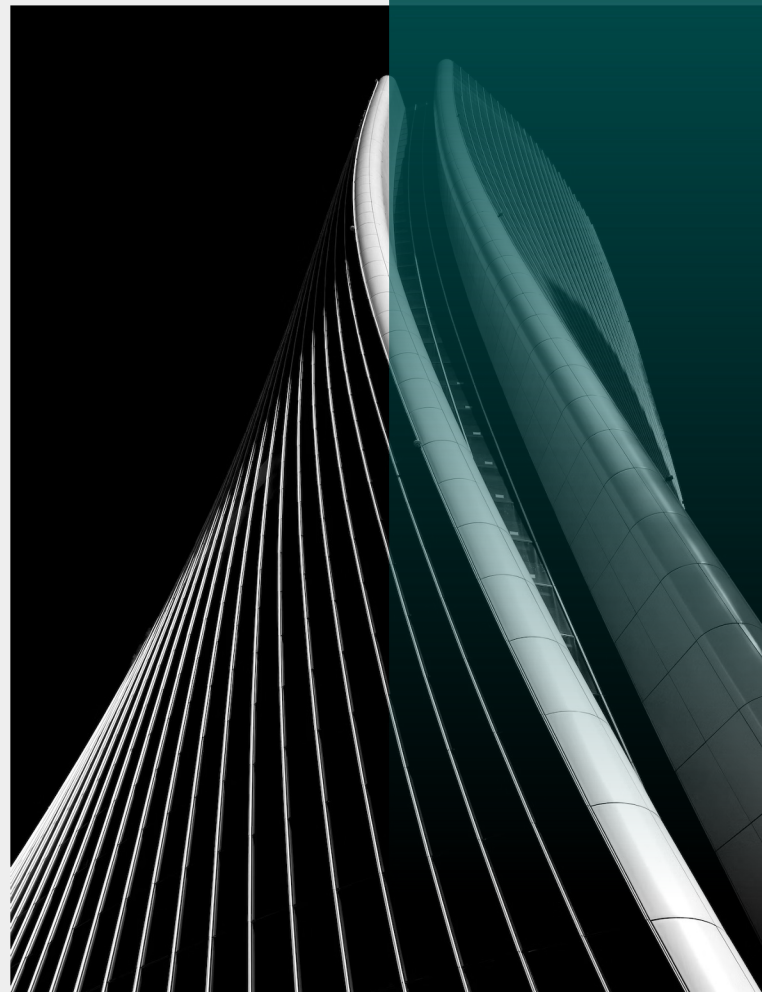


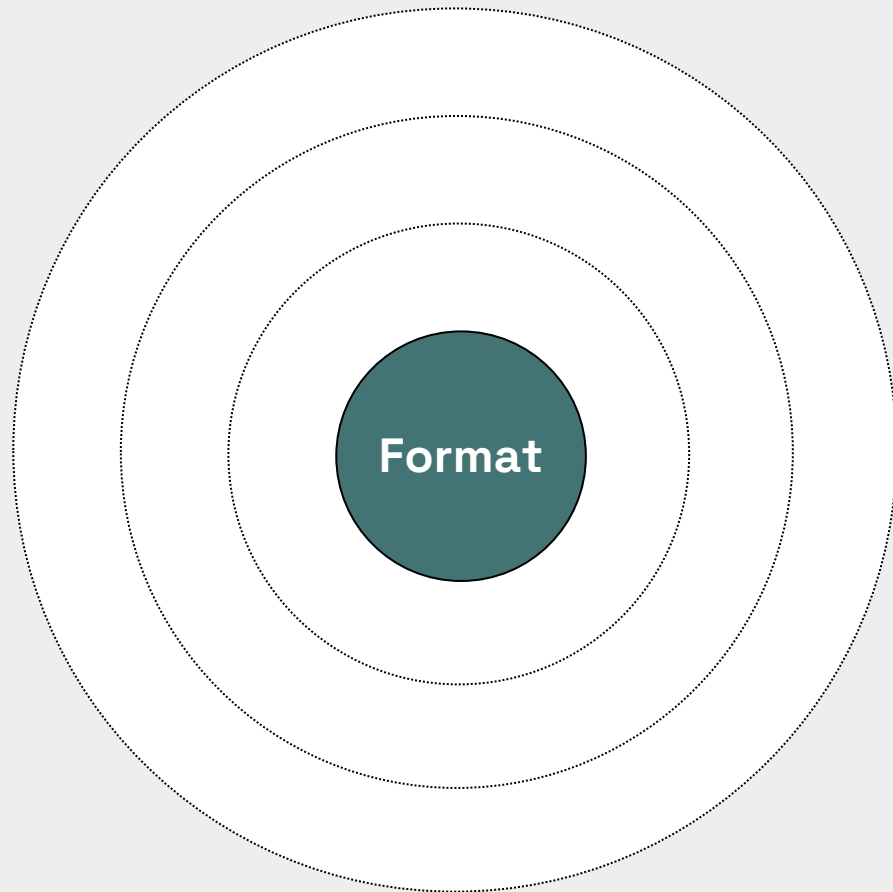
Ian Cook
Director of Product Management
ian@voltrondata.com

Apache Arrow



What is Arrow?

At its core:
a standardized
in-memory
columnar
format for
structured
data



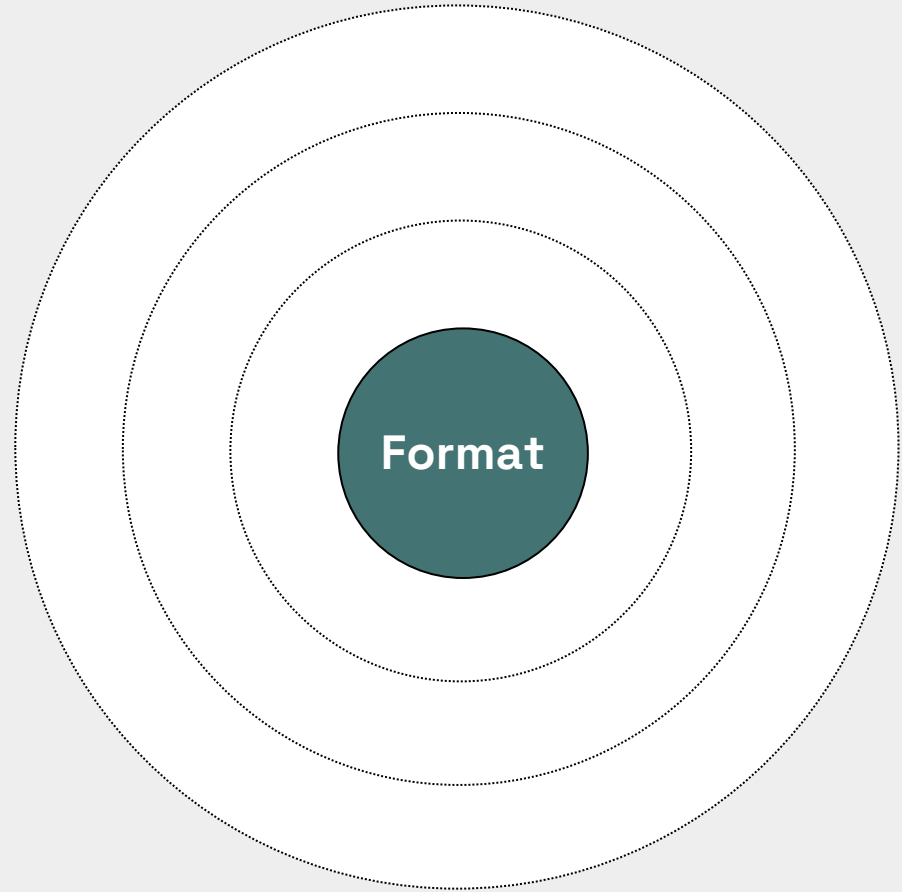
Why Arrow?

**Eliminates data
conversion costs**

**Reduces redundant
software
development**

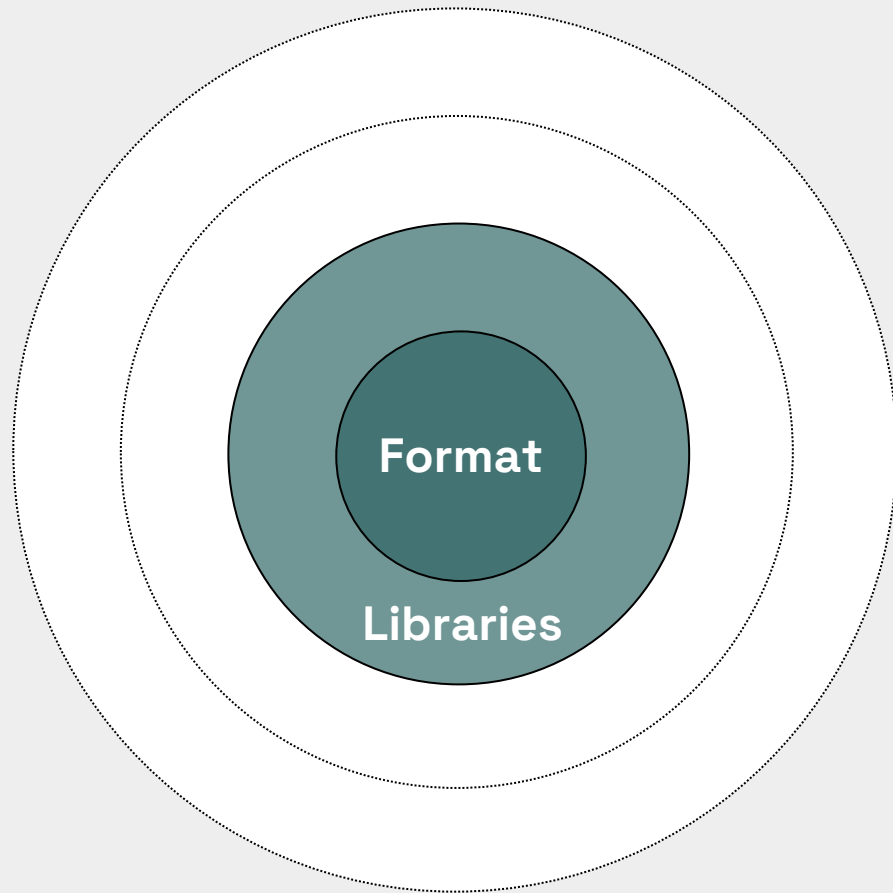
**Enables modular
and composable
data systems**

What else is Arrow?



What else is Arrow?

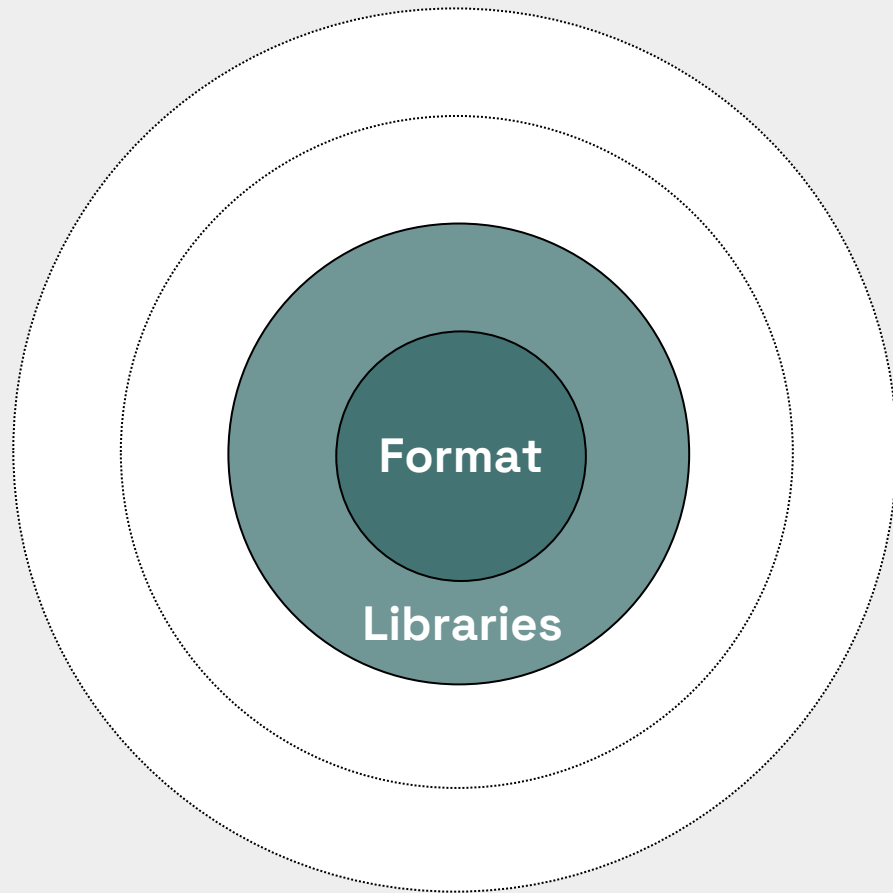
A set of **libraries** in 13+ languages that implement the format and provide useful tools



What else is Arrow?

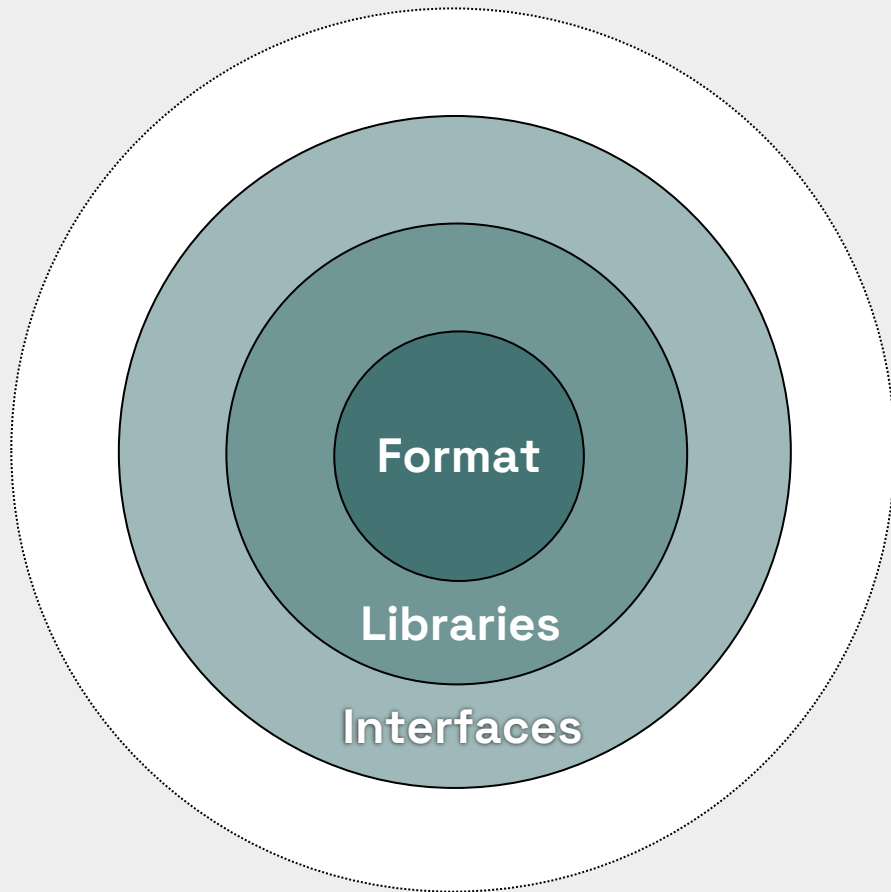
A set of **libraries** in 13+ languages that

C, C++, C#, Go,
Java, JavaScript,
Julia, MATLAB,
Python, R, Ruby,
Rust, Swift



What else is Arrow?

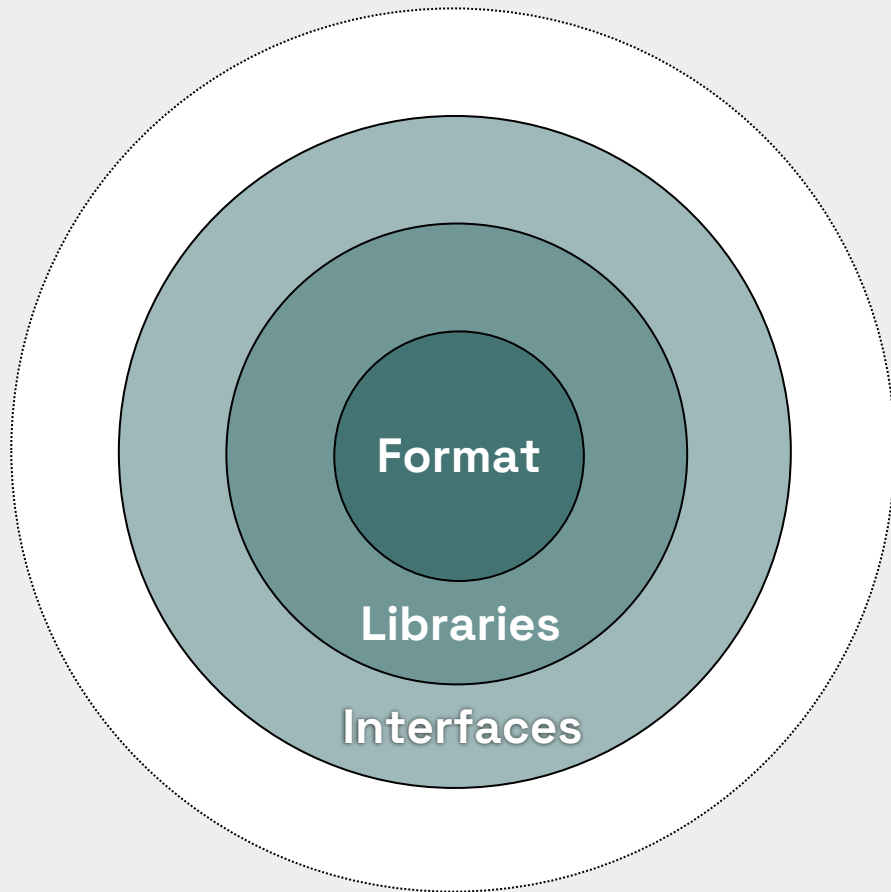
A set of **interfaces** for exchanging Arrow-formatted data



What else is Arrow?

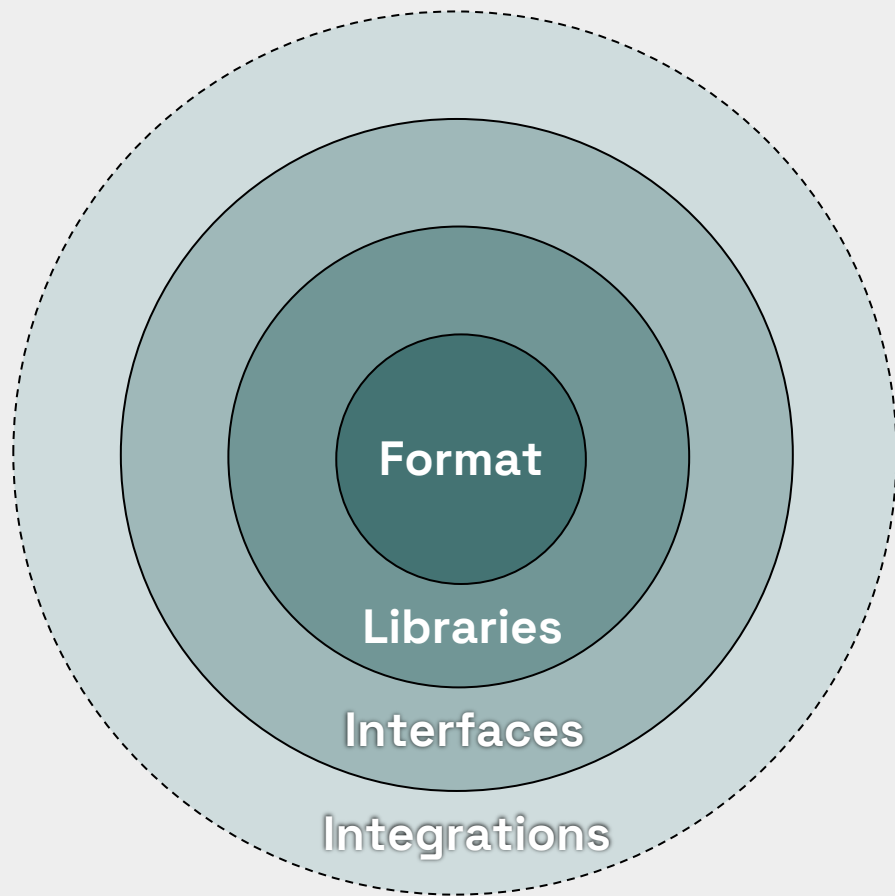
A set of **interfaces** for

...
C data interface, C stream interface, IPC file format, IPC stream format, Flight RPC, Flight SQL, ADBC



What else is Arrow?

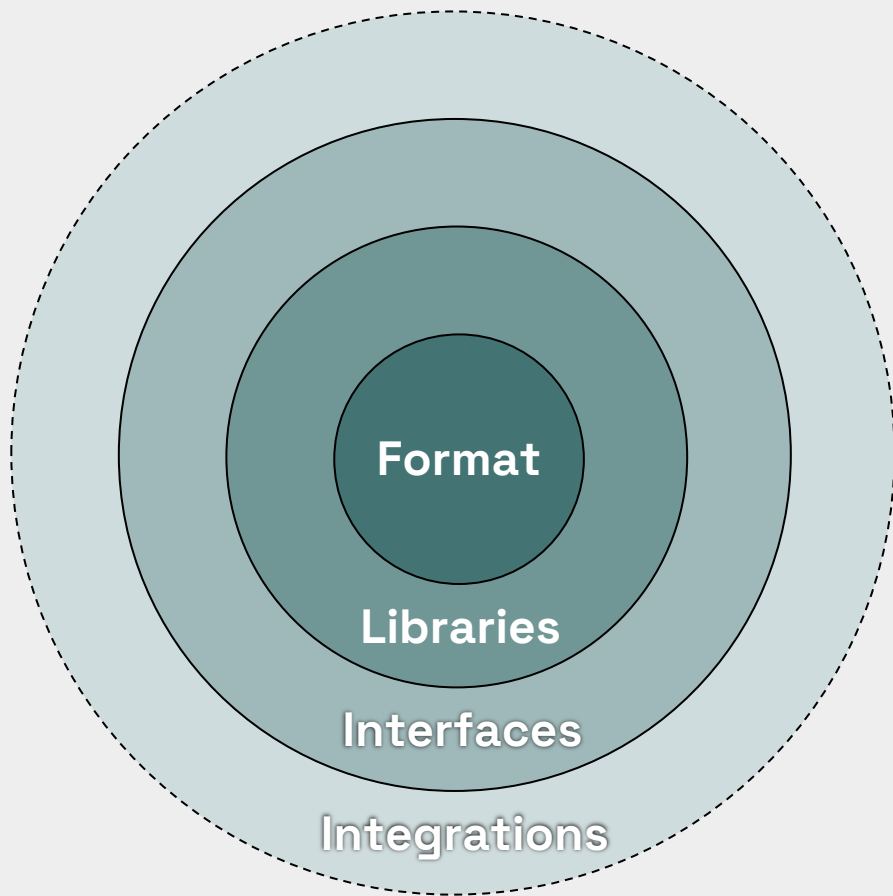
An ecosystem
of **integrations**
of Arrow into
third-party
products,
projects, tools,
frameworks, ...



What else is Arrow?

An ecosystem
of **integrations**

BigQuery, ClickHouse,
Dask, Dremio, DuckDB,
Hugging Face, InfluxDB,
pandas, Parquet, Polars,
Ray, Snowflake, Streamlit,
Spark, Velox, ...





Thank you