era7 bionformatics

Aλhambra day

Eduardo Pareja-Tobes, CTO

2019-05-06

ERA7?

- Founded in **2006**. Granada, Cambridge MA
- *Biological* **data analysis** *Genomics, NGS*
- *Wet lab* too: Sequencing
- **Team** mix of Biology, IT, Maths backgrounds

- **Research** Bioinformatics, Biology
- **100% Open Source** *AGPLv3*, **no** dual licensing
- **Cloud Computing** *AWS*, running on it since ~2007

SOFTWARE

Data Analysis

- *Algorithm* design and implementation
- Low-level DNA sequence analysis
- AWS deployment

Graph Databases

- **Bio4j**: since 2008. Graphs for biological data.
- strongly typed graph query language
- Working on our own graph data engine

- **Webapps**: Mostly for clients using our services.
- **Elm** client-side, datavis

Scala

In production since Scala 2.7/2.8 ~2010. 9 years of experience. We have used it for basically everything:

Use II

- Data Analysis low-level code
- Language/eDSL implementation (graph data query languages)
- Generic Libraries (ohnosequences/cosas, ohnosequences/stuff, …)
- Infrastructure (AWS, mostly EC2, S3, SQS, DynamoDB)
- Web Applications (Play)

Team had been working in Java before; FP obvious
given team background. JVM interop a plus (Graph
Databases, good AWS support, etc).

Learning Scala

Training people has been really hard.

- *Which* Scala?
- Corner cases everywhere
- No good docs anywhere
- ...

Happy?

Not really. Currently using it only as a better Java. No time for going into much detail though.

Performance

- awful collections
- really hard to avoid object creation
- generics and fast: pick one (specialization = more bugs)
- value classes = more bugs, no real value
- …

As a better Java

Nicer syntax, but... mostly that

- interop actually not that good
- Explosion of features/patterns confusing
- Need to know Java anyway
- Java (minus syntax) really close to Java wrt FP

Thinking about moving to Kotlin (web apps, infrastructure) and F# (data analysis).

As an Expressive Language

- Bugs everywhere
- Type inference is awful
- fake dependent types: more problems than what they solve
- Scala3/Dotty not looking good (dropping type projections, focus on singleton types, ...)

Since a year ago doing everything in **Idris** (compilers for graph data languages, distributed graph DB, ...)

Hiring?

We're *always* hiring.

- Languages, qualifications not important
- skills, interest, willingness to learn –critical

Thank you!