

# Data Haskell SIG Workshop

ICFP

Oxford  
September 9, 2017

# Program

- A Haskell for Numerics Manifesto - D.Steinitz
- Probabilistic Programming with monad-bayes - A.Scibior
- Probabilistic Programming with Hakaru - P.Narayan
- Accelerate - T.McDonell
- Why I don't use Haskell (so much) at work - B.Komuves
- A Numerical Haskell FAQ - M.Gajda
- Datahaskell.org - M.Zocca
- Discussion



Marco Zocca  
[github.com/ocramz](https://github.com/ocramz)

ICFP

Oxford  
September 9, 2017

# An open-source organisation

- Started in September 2016 by Nikita Tchayka (@nickseagull), supported by theam.io (who built & fund datahaskell.org)
- 220 users and growing on chat, ~40 active on avg. week

# What/where

- [github.com/datahaskell](https://github.com/datahaskell) : library examples and benchmarks, website, HCAR entry...
- [datahaskell.org/docs](https://datahaskell.org/docs) : curated list of libraries
  - ... from basic tasks (numerics/frames) to specialised algorithms (e.g. classification, regression)
- [gitter.im/dataHaskell/Lobby](https://gitter.im/dataHaskell/Lobby): chatroom (github account needed)

# An idea whose time has come?

- “How do I do data science in Haskell?”
- “Where’s `pandas` for Haskell ?”
- ...
- “How do I build/run a neural network in Haskell?”

# Opportunities

- “Self-documenting”, easy to write/read scientific codes
  - Types, DSLs
- High-performance/easy to parallelise methods
  - Multicore/cluster architectures (see `par`, distributed-process)
- Property-based testing, formal verification ..

# Hurdles/what's missing?

## *Technical*


- (Consensus on) numerical substrate
- Dataframe/matrix/grid-of-scalars thing? “Type providers”?
- Knowledge on time/space complexity of common (numerical) algorithms doesn't carry over trivially to Hs; mismatch of competences

## *Social*

- “Library glue”/standardisation of interfaces
- TUTORIAL MATERIAL, DOCS, BLOG POSTS
- “Batteries-included” framework (à la Matlab/numpy etc.), IDE, “value of tinkering” (ergonomics of exploratory programming)



# Collaboration is vital !

- Publish/document your work. Openly discuss interface design
- haddocks !
- Join  !
  - Conversation on Gitter. Many timezones covered (EU, Australia, US east & west)
  - ping me or Nick for repo push access