

# F# Data

Making structured data first-class citizens

**Tomas Petricek**, University of Cambridge

Project homepage: <http://fsharp.github.io/FSharp.Data>

Get in touch: [@tomaspetricek](https://twitter.com/tomaspetricek) | [tomas@tomasp.net](mailto:tomas@tomasp.net)

software stacks

trainings teaching F# user groups snippets

mac and linux cross-platform tutorials

# F# Software Foundation

F# community open-source MonoDevelop

<http://www.fsharp.org>

contributions research support

consultancy mailing list



## **F# Data type providers**

First-class data

CSV, REST, WorldBank...



## **R Type provider**

Statistics & visualization

5000 tested packages

[www.fslab.org](http://www.fslab.org)



## Deedle data frame

Data exploration

Indexing and aggregation



## F# Charting library

Simple & composable

Interactive style

[www.fslab.org](http://www.fslab.org)

# **What are type providers?**

## **Integrating WorldBank and R**

# The confusion of languages

# What are type providers?

$$\emptyset \vdash e : \tau$$

# What are type providers?

$$\pi(\text{) \vdash e : \tau$$



# Type provider research questions

Data vs.  
Schema

Laziness

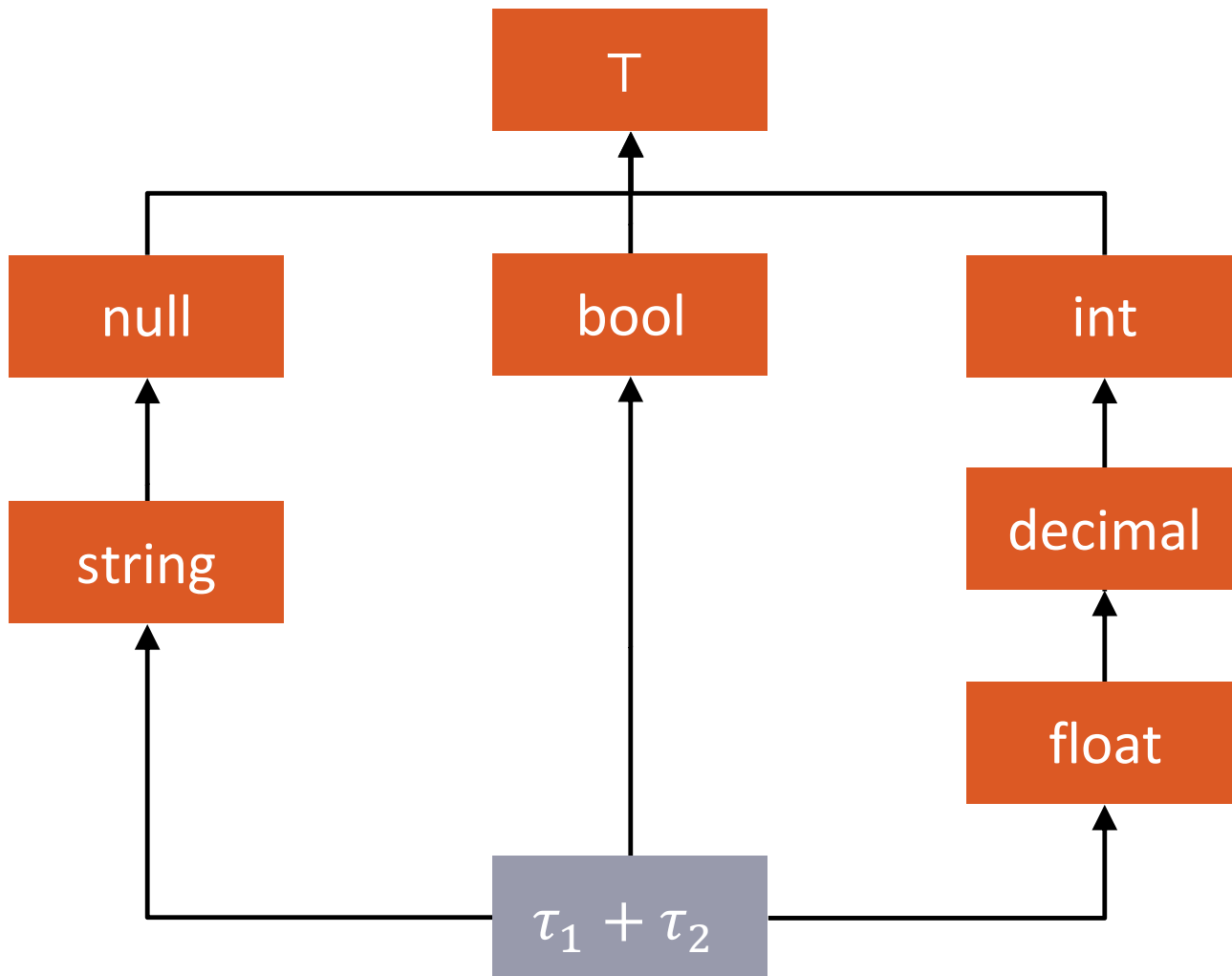
Mapping  
to types

Schema  
inference

# Schema inference

Loading Titanic data from CSV

# Inferring primitive types



# Structure inference

Working with XML and JSON data

# Inferring structured types

person { name : string }

person { name : string, age : int }



person { name : string, age : int option }

[ { num : int } ]

[ { str : string } ]



[ { num : int option, str : string option } ]

int

{ value : int }



int + { value : int }

**Does it scale?**

Query movies using Apiary provider

# Conclusions

Inference from small-scale samples works!

Schema is (very) often missing

But data is (very) often regular

Check out F# Data and contribute!

Project homepage: <http://fsharp.github.io/FSharp.Data>

Get in touch: [@tomaspetricek](https://twitter.com/tomaspetricek) | [tomas@tomasp.net](mailto:tomas@tomasp.net)