Types from data

Making structured data first-class citizens in F#

Tomas Petricek, University of Cambridge

tomas@tomasp.net | @tomaspetricek

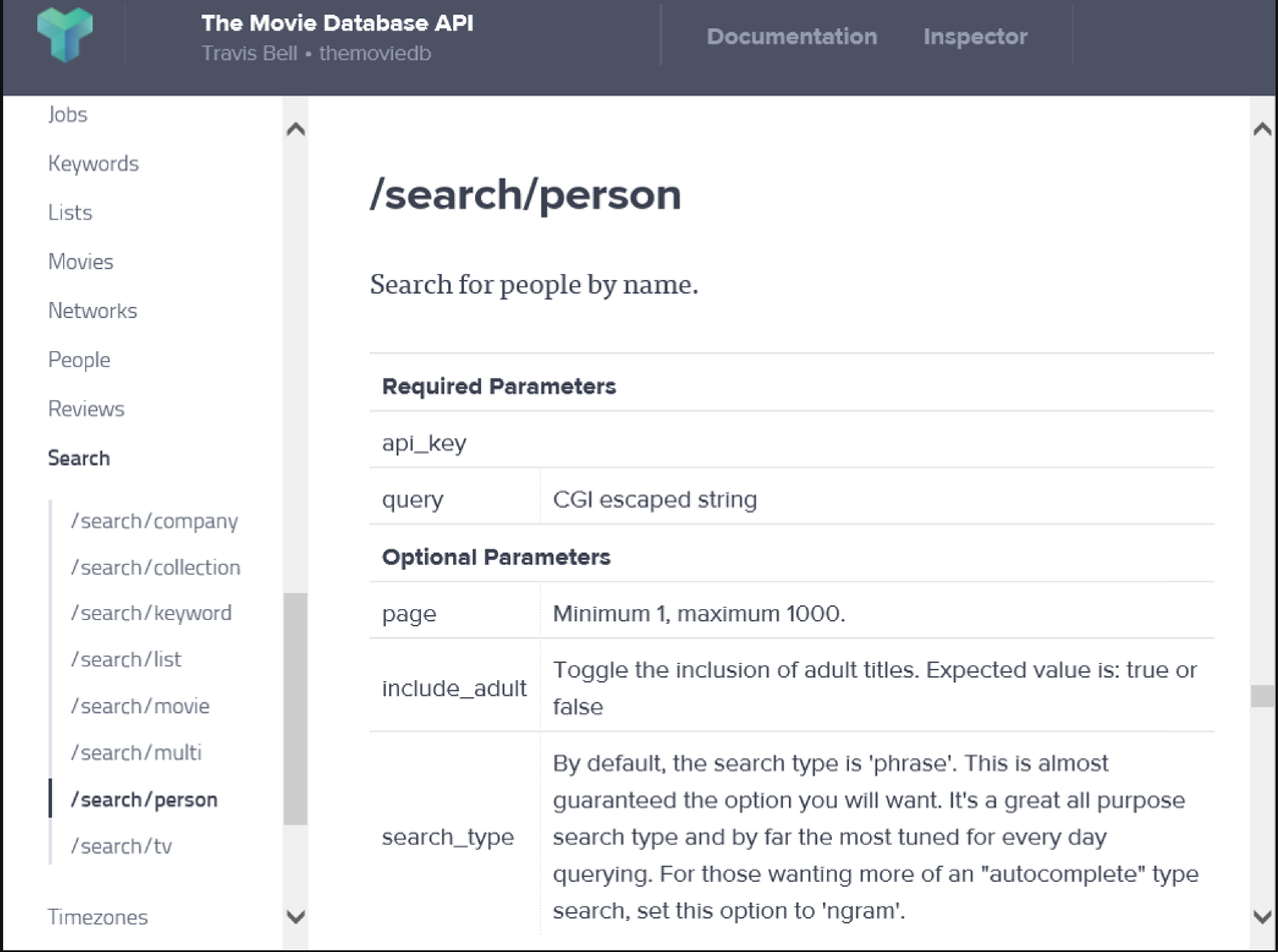
Unsafe dynamic access

```
var url = "http://dvd.netflix.com/Top100RSS";
var rss = XDocument.Load(topRssFeed);
var channel = rss.Element("rss").Element("channel");

foreach(var item in channel.Elements("item")) {
  Console.WriteLine(item.Element("text").Value);
}
```

Not found!

Unsafe dynamic access

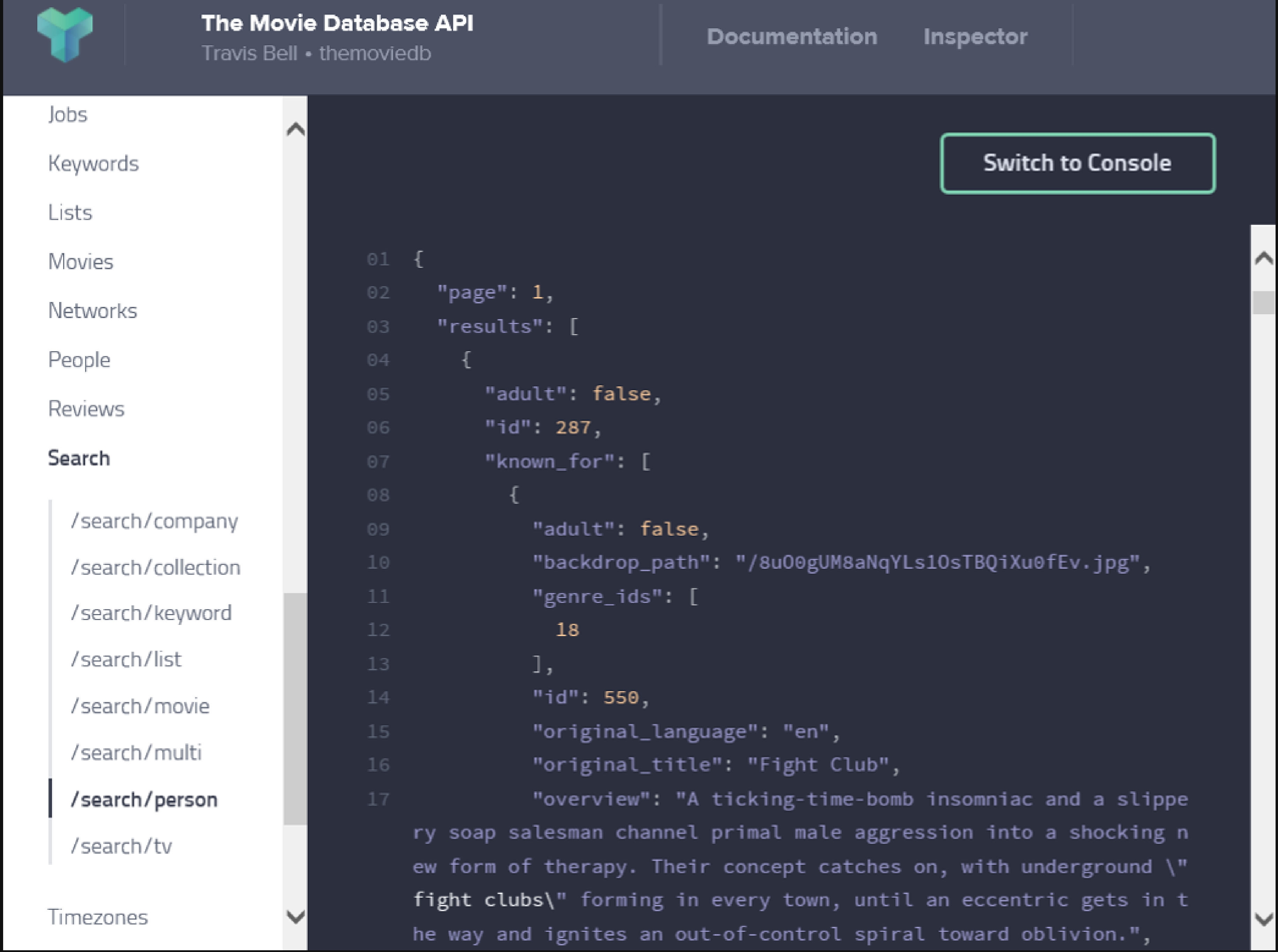
```
var url = "http://dvd.netflix.com/Top100RSS";
var rss = XDocument.Load(topRssFeed);
var channel = rss.Element("rss").Element("channel");

foreach(var item in channel.Elements("item")) {
  Console.WriteLine(item.Element("title").Value);
}
```



/search/person

Search for people by name.

Required Parameters	
api_key	
query	CGI escaped string
Optional Parameters	
page	Minimum 1, maximum 1000.
include_adult	Toggle the inclusion of adult titles. Expected value is: true or false
search_type	By default, the search type is 'phrase'. This is almost guaranteed the option you will want. It's a great all purpose search type and by far the most tuned for every day querying. For those wanting more of an "autocomplete" type search, set this option to 'ngram'.



Jobs
Keywords
Lists
Movies
Networks
People
Reviews
Search
- /search/company
- /search/collection
- /search/keyword
- /search/list
- /search/movie
- /search/multi
- **/search/person**
- /search/tv

Timezones

```
{
  "page": 1,
  "results": [
    {
      "adult": false,
      "id": 287,
      "known_for": [
        {
          "adult": false,
          "backdrop_path": "/8uO0gUM8aNqYLs1OsTBQiXu0fEv.jpg",
          "genre_ids": [
            18
          ],
          "id": 550,
          "original_language": "en",
          "original_title": "Fight Club",
          "overview": "A ticking-time-bomb insomniac and a slippery soap salesman channel primal male aggression into a shocking new form of therapy. Their concept catches on, with underground \"fight clubs\" forming in every town, until an eccentric gets in the way and ignites an out-of-control spiral toward oblivion.",
```

Tedious handwritten mappings

```
public class SearchResults
{
  public int Page { get; set; }
  public List<Result> Results { get; set; }
  public int TotalPages { get; set; }
  public int TotalResults { get; set; }
}
public class Result
{
  public int ID { get; set; }
  public List<KnownFor> KnownFor { get; set; }
  public string Name { get; set; }
  public double Popularity { get; set; }
  public string ProfilePath { get; set; }
}
```

And more...

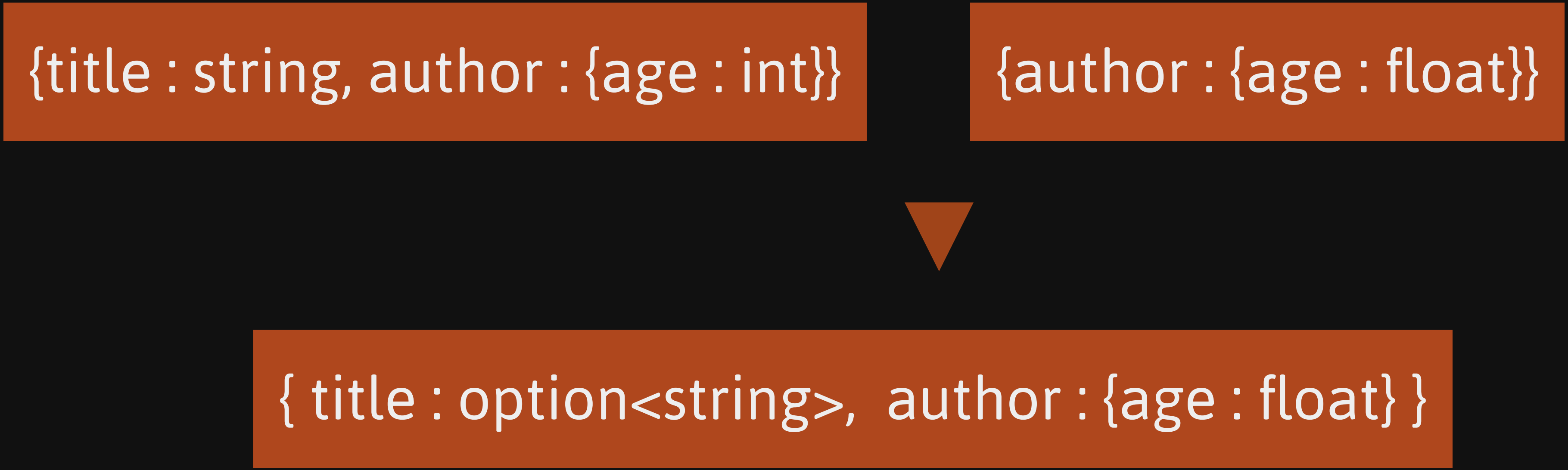
Tedious handwritten mappings

```
public class KnownFor
{
  public bool Adult { get; set; }
  public string BackdropPath { get; set; }
  public List<int> GenreIDs { get; set; }
  public int ID { get; set; }
  public string Overview { get; set; }
  public string ReleaseDate { get; set; }
  public string PosterPath { get; set; }
  public double Popularity { get; set; }
  public string Title { get; set; }
  public double VoteAverage { get; set; }
  public int VoteCount { get; set; }
  public string Name { get; set; }
}
```

Much much more...

DEMO

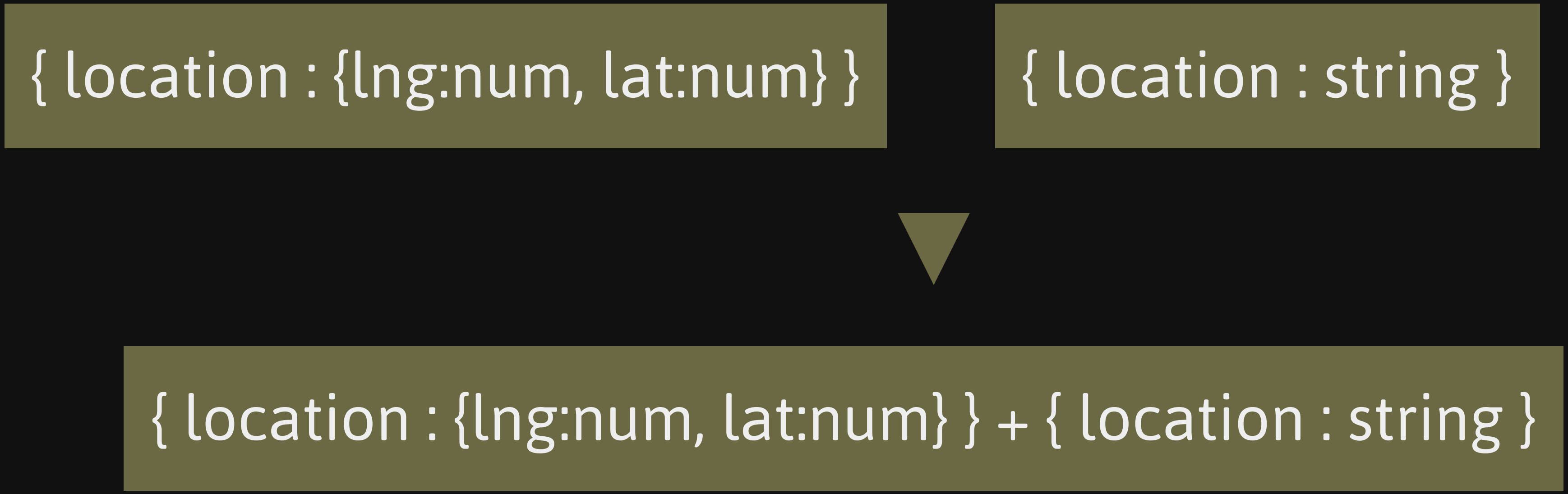
Reading news from RSS feed



{title : string, author : {age : int}}

{author : {age : float}}

{ title : option<string>, author : {age : float} }



{ location : {lng:num, lat:num} }

{ location : string }

{ location : {lng:num, lat:num} } + { location : string }

DEMO

Getting weather via REST service

Behind the scenes

Structural shape inference

Language integration via type providers

Relative type safety

Relative type safety

Given representative samples and an input value

$$S(d') \sqsubset S(d_1, \ldots, d_n)$$

Any program written user type provider reduces

$$e_{user}[x \leftarrow e_{provided}(d')] \leadsto^{*} v$$

DEMO

Error handling and schema change

Schema change and stability

Inferred type can change only in limited ways

$$C[e] \to C[e.M]$$

$$C[e] \to C[\textsf{match}\ e\ \textsf{with}\ \ldots]$$

$$C[e] \to C[int(e)]$$

Addressing practical concerns

- Prefer records for tooling support
- Predictable and stable shape inference
- Open world assumption about top shapes

Summary

Works in practice

(148k downloads, 1900 commits, 47 contributors)

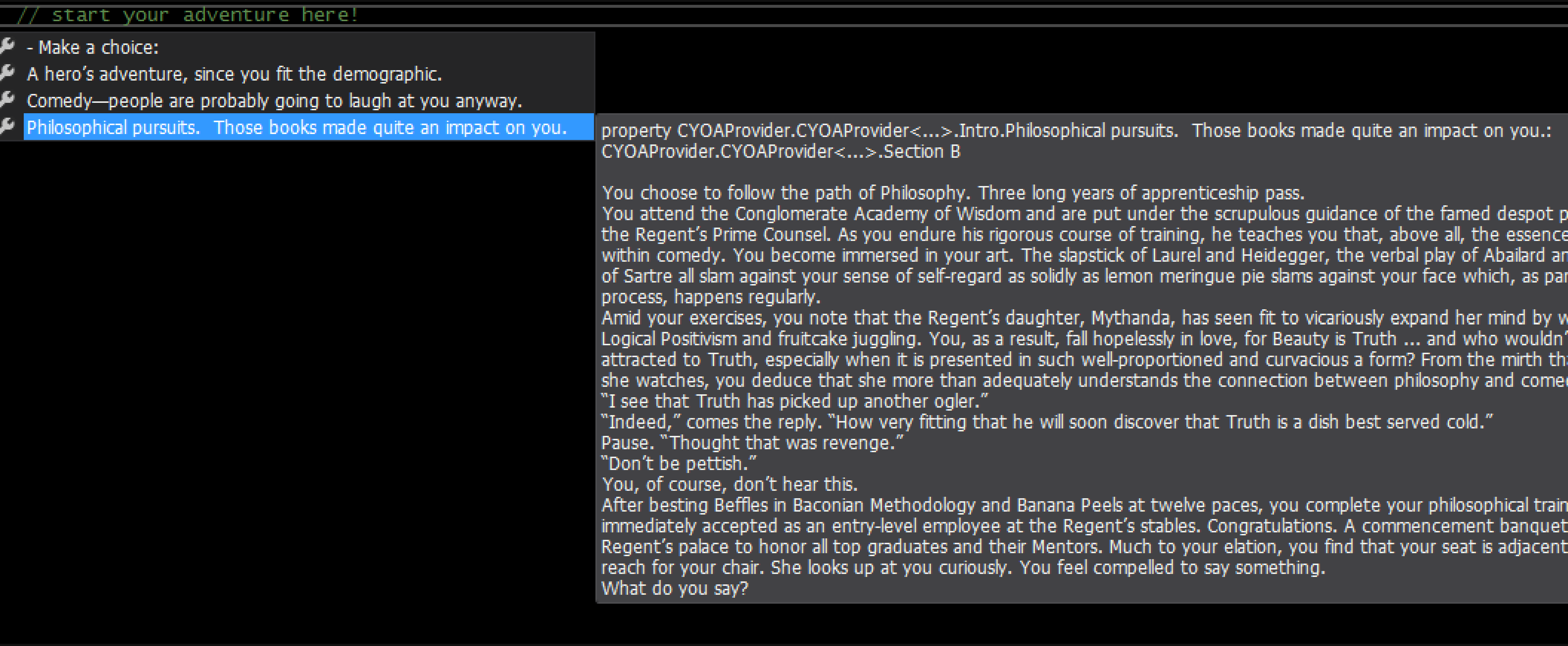
Type safety revisited

Relative safety necessary for modern programs

Tomas Petricek, University of Cambridge

tomas@tomasp.net | @tomaspetricek

Bonus slides

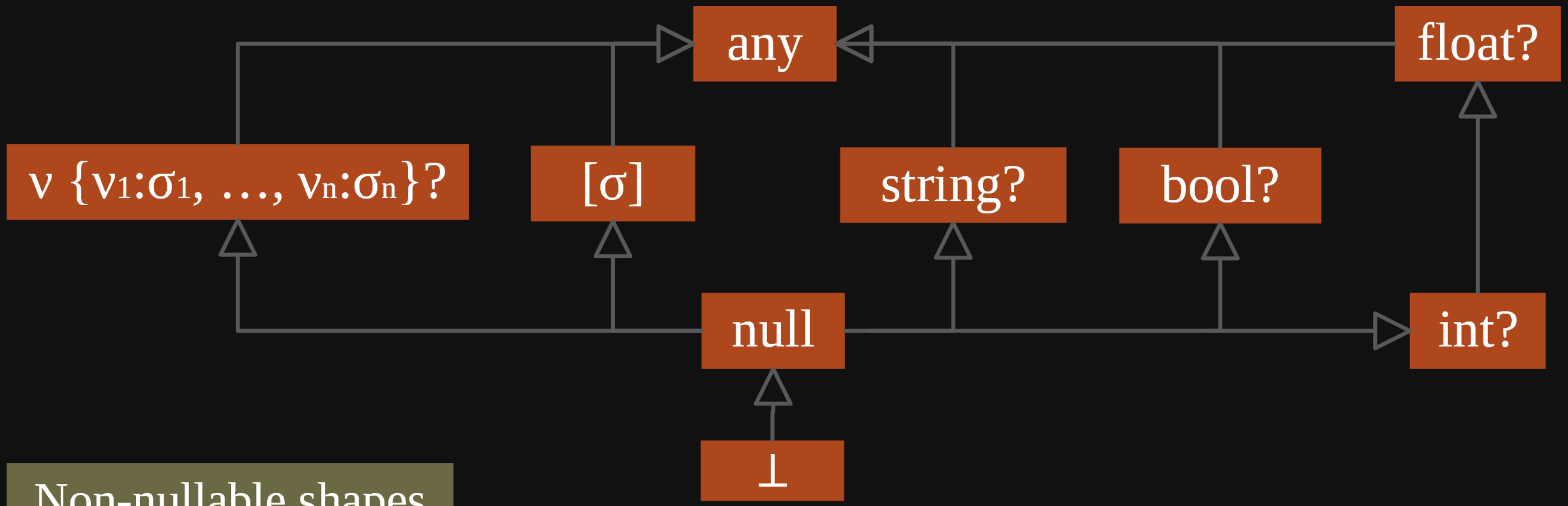


```
// start your adventure here!
```

```
- Make a choice:
A hero’s adventure, since you fit the demographic.
Comedy—people are probably going to laugh at you anyway.
Philosophical pursuits.  Those books made quite an impact on you.
```

```
property CYOAProvider.CYOAProvider<...>.Intro.Philosophical pursuits.  Those books made quite an impact on you.:
CYOAProvider.CYOAProvider<...>.Section B

You choose to follow the path of Philosophy. Three long years of apprenticeship pass.
You attend the Conglomerate Academy of Wisdom and are put under the scrupulous guidance of the famed despot p
the Regent’s Prime Counsel. As you endure his rigorous course of training, he teaches you that, above all, the essence
within comedy. You become immersed in your art. The slapstick of Laurel and Heidegger, the verbal play of Abailard an
of Sartre all slam against your sense of self-regard as solidly as lemon meringue pie slams against your face which, as par
process, happens regularly.
Amid your exercises, you note that the Regent’s daughter, Mythanda, has seen fit to vicariously expand her mind by w
Logical Positivism and fruitcake juggling. You, as a result, fall hopelessly in love, for Beauty is Truth ... and who wouldn’
attracted to Truth, especially when it is presented in such well-proportioned and curvacious a form? From the mirth tha
she watches, you deduce that she more than adequately understands the connection between philosophy and come
“I see that Truth has picked up another ogler.”
“Indeed,” comes the reply. “How very fitting that he will soon discover that Truth is a dish best served cold.”
Pause. “Thought that was revenge.”
“Don’t be pettish.”
You, of course, don’t hear this.
After besting Beffles in Baconian Methodology and Banana Peels at twelve paces, you complete your philosophical train
immediately accepted as an entry-level employee at the Regent’s stables. Congratulations. A commencement banquet
Regent’s palace to honor all top graduates and their Mentors. Much to your elation, you find that your seat is adjacent
reach for your chair. She looks up at you curiously. You feel compelled to say something.
What do you say?
```

Nullable shapes



any

float?

$\nu \{\nu_1{:}\sigma_1, \ldots, \nu_n{:}\sigma_n\}?$

$[\sigma]$

string?

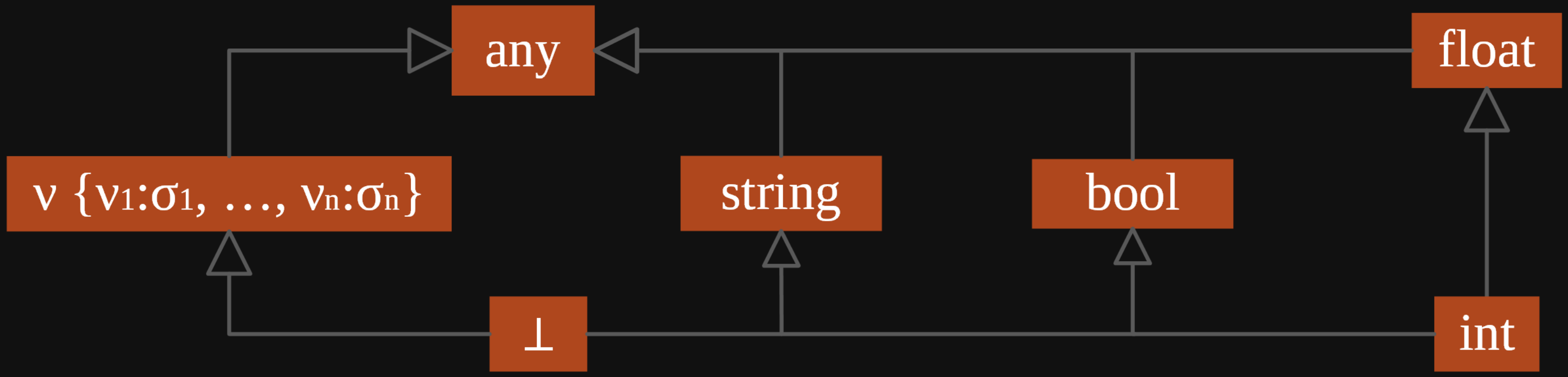
bool?

null

int?

$\bot$

Non-nullable shapes



any

float

$\nu \{\nu_1{:}\sigma_1, \ldots, \nu_n{:}\sigma_n\}$

string

bool

$\bot$

int

Labelled top types

Fundamental open world assumption

```xml
<doc>
  <heading>Working with JSON</heading>
  <p>Type providers make this easy.</p>
  <heading>Working with XML</heading>
  <p>Processing XML is as easy as JSON.</p>
  <image source="xml.png" />
</doc>
```

Labelled top types

Top type annotated with possible cases

$$\mathsf{any}\langle \sigma_1, \ldots, \sigma_n \rangle$$

Provides access to $\sigma_1, \ldots, \sigma_n$

Requires handling of unknown case