

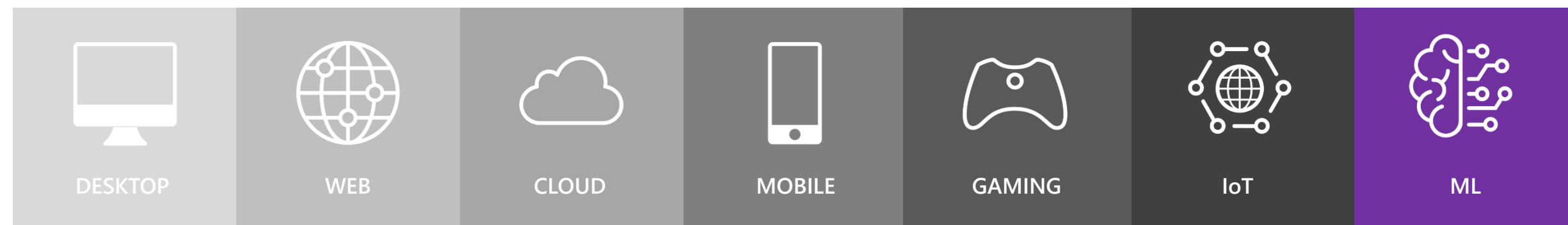
Machine Learning at Microsoft with ML.NET

Matteo Interlandi, Sergiy Matusyevych, Saeed Amizadeh, Shauheen Zahirazami, Markus Weimer



.NET

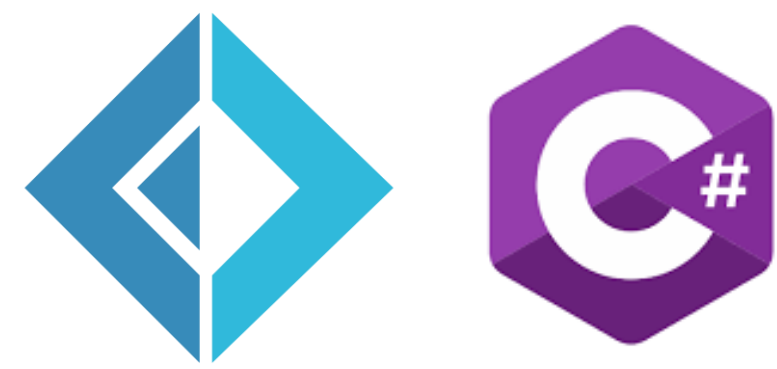
.NET is the platform for building anything



On any platform



With easy to use functional / object oriented languages



And native integration with C/C++ for best performance

ML.NET

<https://dot.net/ml>
<https://github.com/dotnet/machinelearning>

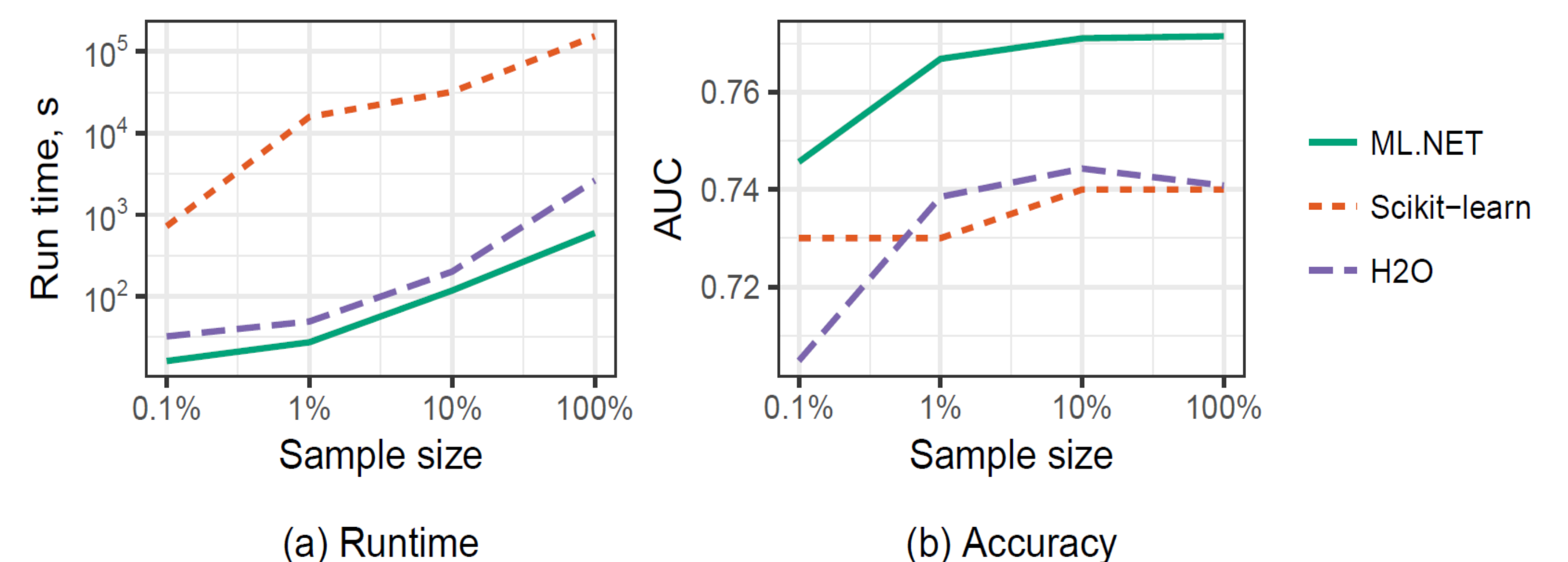
A machine learning framework made for .NET developers. Used internally at Microsoft for several years and by hundreds of services



ML.NET is a framework for building ML pipelines

```
var context = new MLContext();
var loader = context.Data.TextReader(...);
var trainingData = loader.Reader(trainDataPath);
var featurizer = context.Transforms.Text
    .FeaturizeText("Text", "Features");
var trainer = context.BinaryClassification.Trainers
    .FastTree("Label", "Features");
var pipeline = featurizer.Append(trainer);
var model = pipeline.Fit(trainingData)
```

Experiments over Criteo with a pipeline which (1) fills in the missing values in the numerical columns of the dataset; (2) encodes a categorical columns into a numeric matrix using a hash function; and (3) applies a Gradient Boosted Tree Classifier (LightGBM for ML.NET).



.NET Open Source Software

4M

Monthly Active Developers

450K

Monthly New Developers

15%

Monthly GitHub Growth

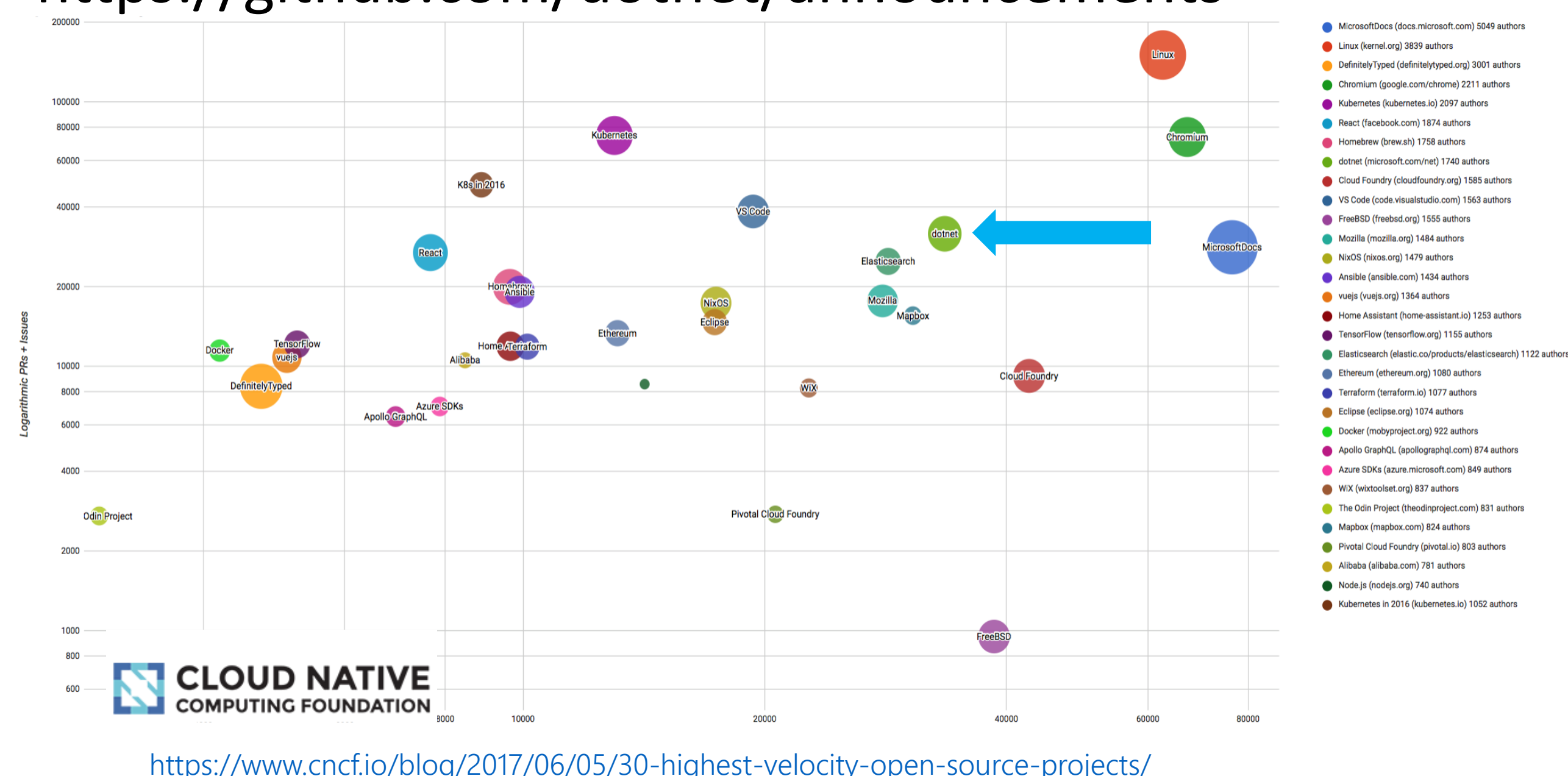


19,000+ Community Contributions from 3700+ Companies **outside Microsoft**

Notable habits in .NET OSS Projects:

- API review meetings are public on YouTube, notes and decisions managed on GitHub.
- Online Community Standup meetings
- Using GitHub for news:

<https://github.com/dotnet/announcements>



<https://www.cncf.io/blog/2017/06/05/30-highest-velocity-open-source-projects/>

NET foundation

Technical Steering Group



25,000



CONTRIBUTORS

1700



COMPANIES

55



ACTIVE PROJECTS

462



RESOURCES

.NET Core and ML.NET are available under the MIT license.