

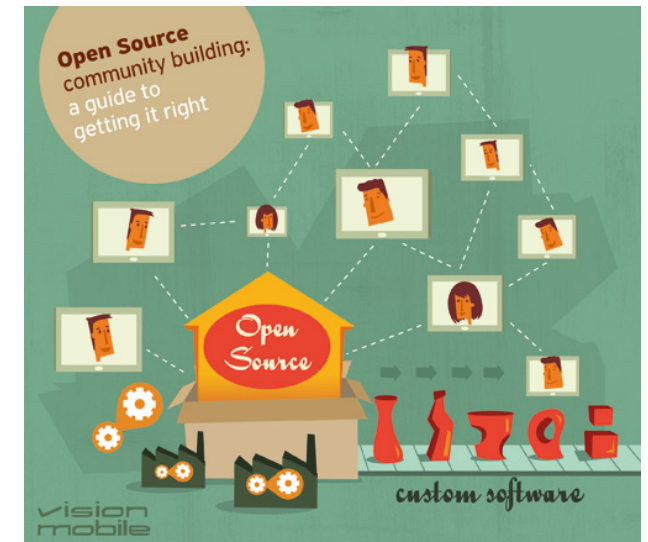
# Industry Open-Source Projects: Reasons, Benefits & Best Practices

Arun Kalyanasundaram, Carnegie Mellon University  
Mentor: Judith Bishop, Microsoft Research



# Our Goals

- Help open-source projects build stronger communities
  - Identify current motivations and expectations.
  - How are the projects benefiting from open-source?
  - Compile existing best practices across different projects.
  - Build tools to understand the Degree of External Engagement.
  - What are the dimensions on which projects differ and why?



# Our Approach

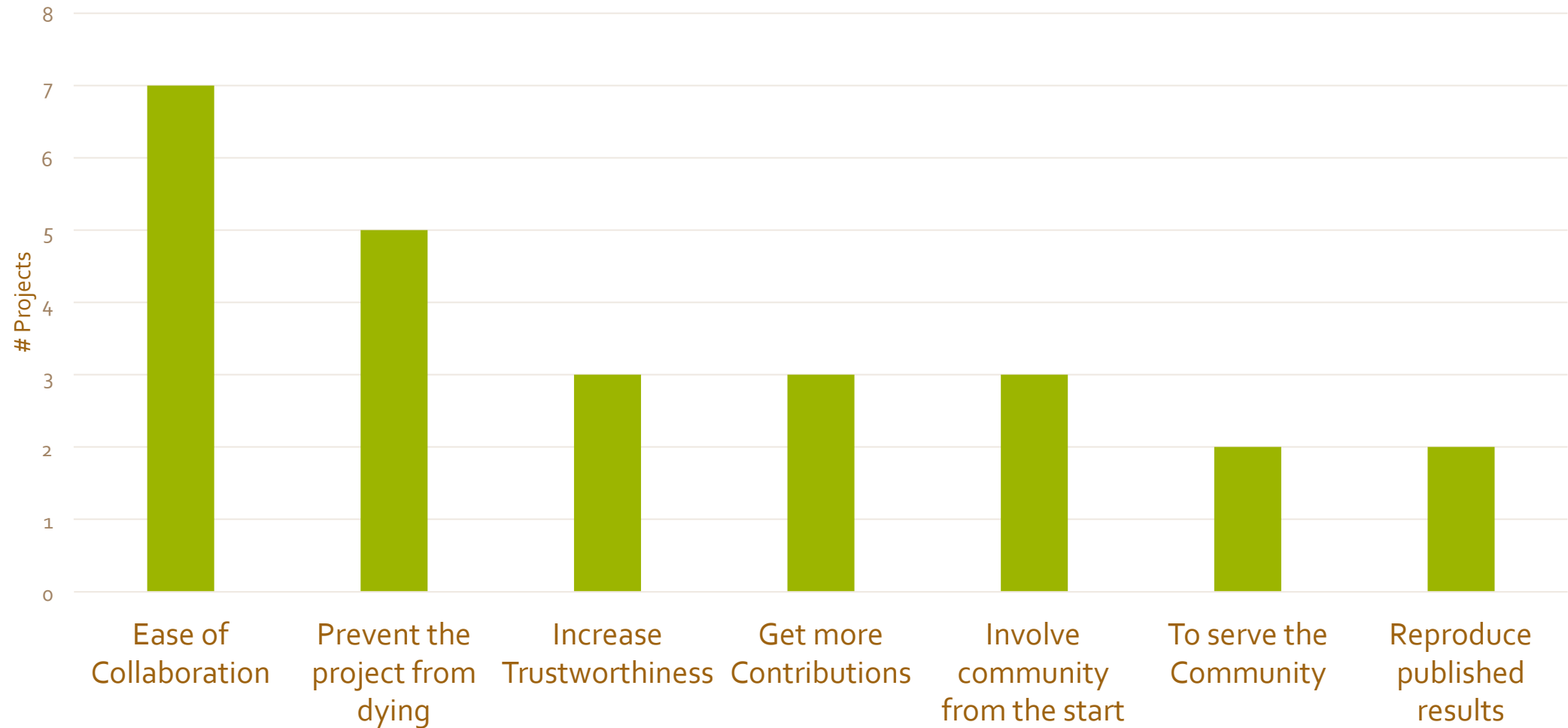
- We interviewed owners of 23 MSR and 8 product group projects.
  - **Project Usage:** target users, size of user base.
  - **Outside collaborations:** Adoption in universities, companies, etc.
  - **Open-Source practices:** Motivations, contributions.
  - **Development Process:** Managing code on Github, Addressing user requests.



- Analyze Github data:
  - Extract various metrics to look at both aggregate and temporal patterns.
  - Triangulate with other sources such as Gitter, Stack overflow in future.

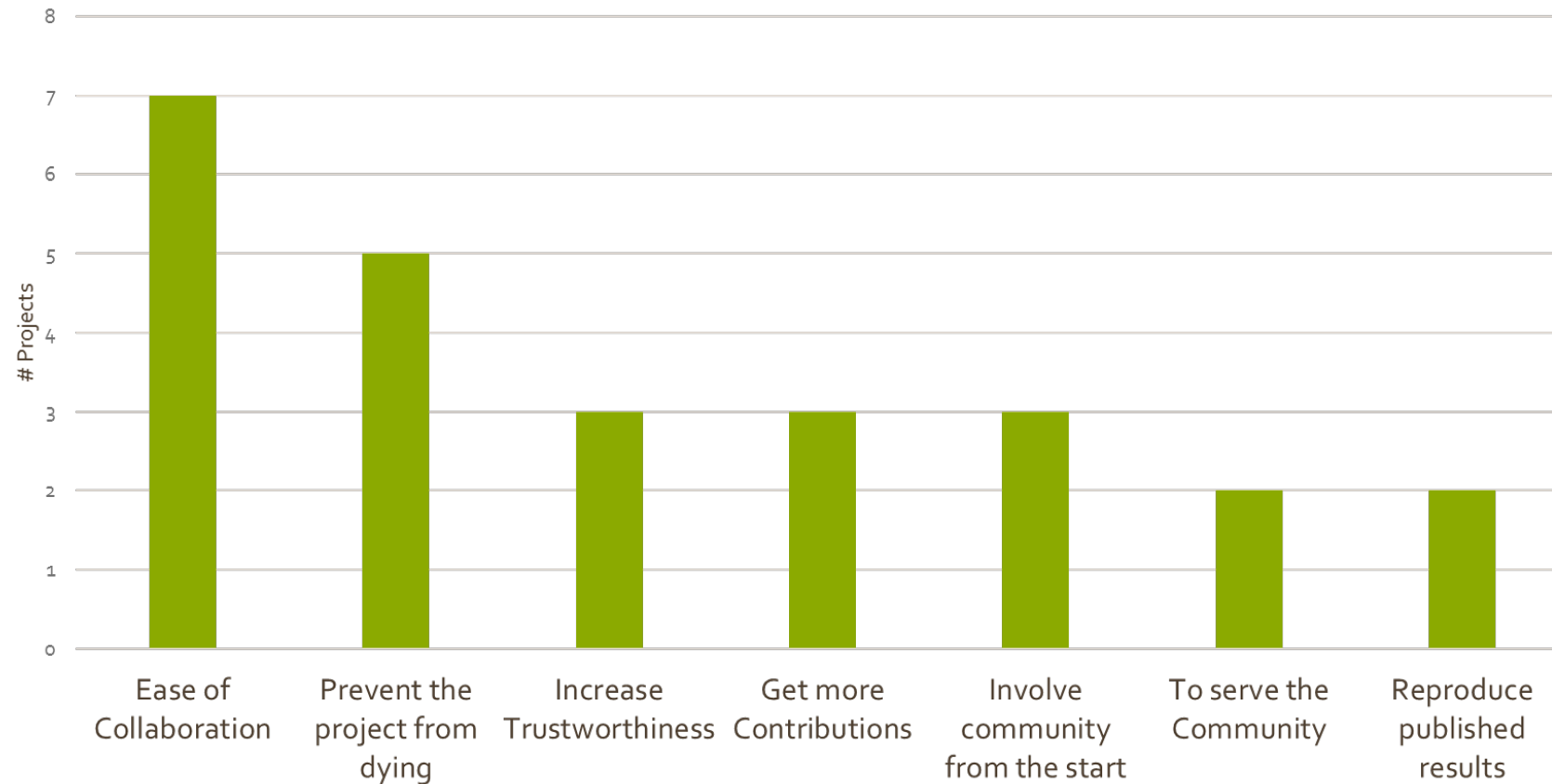


# Open Source Motivations



Others: Requests from users, Increase adoption, Improve quality

# Open Source Motivations



Others: Requests from users, Increase adoption, Improve quality

# What has Worked

# What has Worked



Improve Portability

*Orleans: Port from Azure to SQL Server*

*MSAGL: Script to automatically build a Nuget package*

*.Net Core: Ports to BSD linux and other flavors*

## Better install under non Windows systems #338

**Merged** NikolajBjorner merged 15 commits into `Z3Prover:master` from `delcypher:better_packaging` on Dec 6, 2015

Conversation 38 Commits 15 Files changed 1



delcypher commented on Nov 24, 2015

This series of patches fixes a few issues with the current build system when running `make install`.

- Adds support for `$(DESTDIR)` (see [https://www.gnu.org/prep/standards/html\\_node/DESTDIR.html](https://www.gnu.org/prep/standards/html_node/DESTDIR.html))
- Makes `libz3.so` in the Python package directory a symlink (rather than a copy) if possible, to save space when packaging.

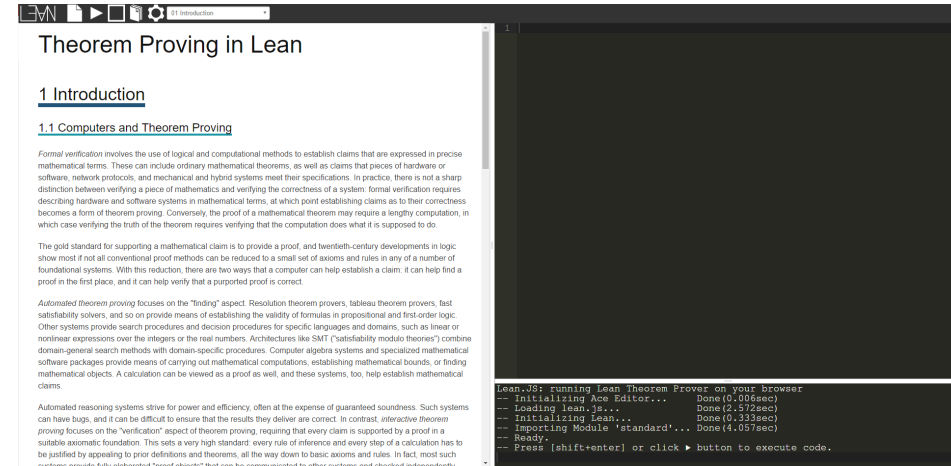
There will be other packaging issues that I'd like to fix in the future but I thought I'd drop some of my initial work so it can be reviewed.

*Z3: Fix for installing on other Operating systems*

# What has Worked



Increase Usability



*LEAN: A student implemented an Online GUI tutorial*



Self Document

## A error in user manual #29

**Closed** LionSR opened this issue 10 days ago · 1 comment



LionSR commented 10 days ago

Dear Liquid team,  
I'm trying to implement Stean7 class acorrding to your nice user manual.  
At the page of 56 and 58(Example 53 and Example 55), one line of the code was replaced by  
"Error! Hyperlink reference not valid."

*Liquid: Although not open source, Users can report bugs via Github*



# What has Worked



Entirely Community Driven



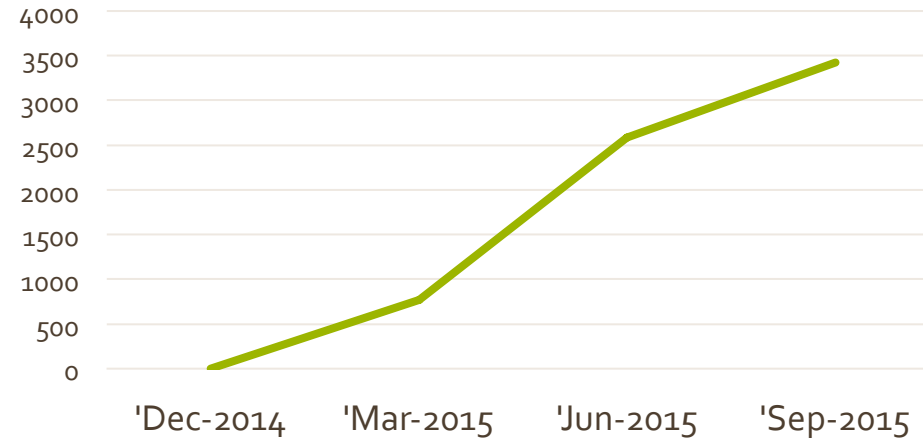
Increase Adoption

## Join the F# Software Foundation



The mission of the F# Software Foundation is to promote, protect, and advance the F# programming language, and to support and facilitate the growth of a diverse and international community of F# programmers.


*ORLEANS: No. of Downloads*



# What has Worked




Community Members  
Answer Questions



gafter commented on Nov 26, 2015


.NET Foundation member

@leppie The features flag is supported in the project file in Update 2 (i.e. imminently).



paulomorgado commented on Nov 26, 2015


What is "strict mode"?



pdelvo commented on Nov 27, 2015


The "old" compiler did have some bugs where code would compile that should, by specification, not. To not break compatibility those bugs were also implemented in roslyn. strict mode is a mode in which the compiler does produce errors in these cases.

*Roslyn Github*



In order to use **gdb** over **z3**, you first need to compile it in **debugging** mode.

0



To accomplish that you don't need to change the **makefile**, just type the following commands in z3's main directory


```
./configure -d
cd build; make
```

Then:

```
~$ gdb ./z3
(gdb) break main
(gdb) r some_file.smt2
```

share edit

answered 8 hours ago



Patrick Trentin

1,476 ●2 ●6 ●17

add a comment

*Z3 Stack Overflow*

<b>AFeldman</b> Feb 5 at 8:46 AM Edited Mar 14 at 7:13 PM	I understand what you are saying. Best of luck.
<b>publicENEMY</b> Feb 5 at 2:43 PM	@daan  haters will be haters. their judgement are clouded by hate. no need to respond or justify anything. also, i've read halfway the serialmentor post. waste of my time.  also, the discussion should be about Madoko. Not 'typical Microsoft strategy'.  Good work with Madoko daan. Keep it up.

*Madoko Codeplex*

# Popular Best Practices

# Popular Best Practices



## Acknowledge and Reward Contributors



Special thanks to all contributors to this release:

- Abubaker Bashir
- Alex Eagle
- Alexander Kuvaev
- Andrew Z Allen
- Andy Hanson

*Individually acknowledging contributors in release notes*



*Roslyn: Ship personalized swags to contributors*

## Community Contribution

This is a huge milestone and accomplishment for the entire .NET ecosystem. Nearly 10k developers contributed to .NET Core 1.0. We never imagined that many folks contributing to the product. We've also been impressed by the quality of the contributions. There are significant components that the community is driving forward. Nice work, folks!

*.Net Core: Calling out contributors on Blogs*

# Popular Best Practices



Clear Guidelines

## Contribute

---

There are many ways to contribute to Z3.

- Engage with other Z3 users and developers on [StackOverflow](#).
- Contribute tests and benchmarks to [z3test](#).
- Contribute code.
- For more information see [contribution guidelines](#).

## Contribute

---

There are many ways to [contribute](#) to TypeScript.

- [Submit bugs](#) and help us verify fixes as they are checked in.
- Review the [source code changes](#).
- Engage with other TypeScript users and developers on [StackOverflow](#).
- Join the [#typescript](#) discussion on Twitter.
- [Contribute bug fixes](#).
- Read the language specification ([docx](#), [pdf](#), [md](#)).

## Building ChakraCore

---

You can build ChakraCore on Windows 7 SP1 or above, and Windows Server 2008 R2 or above, with either Visual Studio 2013 or 2015 with C++ support installed. Once you have Visual Studio installed:

- Clone ChakraCore through `git clone https://github.com/Microsoft/ChakraCore.git`
- Open `Build\Chakra.Core.sln` in Visual Studio
- Build Solution

More details in [Building ChakraCore](#).

## Contributing Code

---

Lucian Wischik edited this page on Apr 27 · 16 revisions

### Contributing Code

Before submitting a feature or substantial code contribution please discuss it with the team and ensure it follows the product [roadmap](#). The team rigorously reviews and tests all code submissions. The submissions must meet an extremely high bar for quality, design, and roadmap appropriateness.

# Popular Best Practices



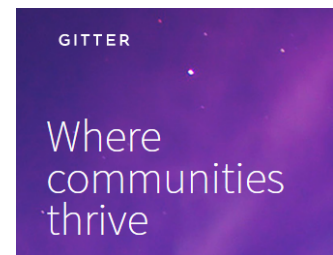
## Managing Multiple Channels



2 votes  
1 answer  
Q: Z3 : Questions About Z3 int2bv?  
I'm a little confused with the Z3 (smt2 format) operation int2bv. I wrote a such smt2 expression : (declare-const t1 Int) (assert (= ((int2bv 2) t1) #b11)) (check-sat) (get-model) when I solve ... it with Z3, it got: sat (model (define-fun t1 () Int 0) ) Is that correct? Shouldn't t1 be 3? I thought the int2bv operation just transform the int value to the equivalent bitvector value. But it seems not! Thanks! ...  
asked Jan 13 '15 by J.Huang  
z3 smt

1 vote  
1 answer  
Q: Z3 -smt2 -in: Get Z3 version  
Can I get the version of Z3 after starting it with the options -smt2 -in? Something like (get-z3-version) ; Z3 4.3.2 x64 // Desired reply ...  
asked Apr 8 '13 by Maite Schwerhoff  
version stdin z3

6 votes  
1 answer  
Q: Z3 Polarity using Z3 as SAT Solver  
I am trying to solve a SAT problem with 12000+ boolean variables using Z3. I expect that most of the variables will evaluate to false in the solution. Is there a way to guide or hint Z3 as SAT solver to try "polarity false" first? I've tried it with cryptominisat 2 and got good results. ...  
asked Dec 17 '12 by Axel Kemper  
z3



Microsoft/TypeScript TypeScript is a superset of JavaScript that compiles to clean JavaScript output.

like that?

I'm not using `tsc` though. I'm using a typescript text editor extension.

Valentin Robert @Ptival  
yes, then a second file `bar.ts` mentioning Test without importing `test.ts`

Yousef Shanawany @AskYous  
yeah so it works fine with me

# Popular Best Practices



Transparency

## C# Design Notes for May 10, 2016 #11516



**MadsTorgersen** opened this issue on May 23 · 46 comments



MadsTorgersen commented on May 23

### C# Language Design Notes for May 10

In this meeting we took a look at the possibility of adding new kinds of extension members, beyond extension methods.

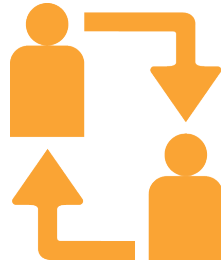
*Sharing Minutes of Meetings*

## Roadmap

For details on our planned features and future direction please refer to our [roadmap](#).



# Popular Best Practices



Responsiveness

*CNTK: Core team members take turns to respond to user requests*

*Z3: Core team responds to user queries every day on Github.  
"Community members also pitch in"*

🔔 277 Open ✓ 80 Closed		Author ▾	Labels ▾	Milestones ▾	Assignee ▾	Sort ▾
🔔	Incorrect indentation in LambdaRewriter	Area-Compilers	Bug	Up for Grabs		1
#12709 opened 11 days ago by swick 📌 2.1						
🔔	CS0123 (no overload matches delegate) should offer 'Generate method' quick action	Area-IDE	Bug	Up for Grabs		1
#12708 opened 11 days ago by miloush 📌 Unknown						
🔔	Inline Temp sometimes incorrectly positions the caret	Area-IDE	Bug	Up for Grabs		1
#12683 opened 13 days ago by dpoeschl 📌 Unknown						

*Roslyn makes it easy for newcomers by tagging easy bugs using – Up for Grabs*



# Popular Best Practices



Remain Active



*Frequent releases and Commits*



Cross Platform

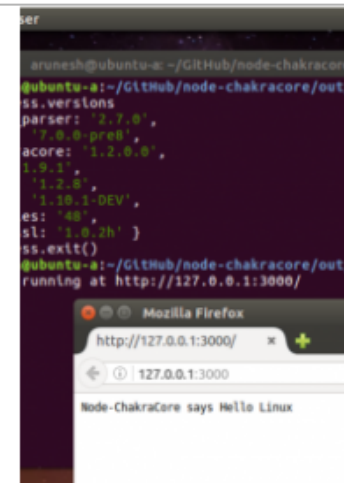
JULY 27, 2016 / PC, Tablet

## Bringing ChakraCore to Linux and OS X

By Limin Zhu and Arunesh Chandra

In January, we open-sourced ChakraCore, the core of the Chakra JavaScript engine that powers Microsoft Edge and Universal Windows We...

[Read more](#)



*Lack of Cross Platform hindered adoption in Chakracore*

# Popular Best Practices



## Archive Community Projects


### F# Community Projects

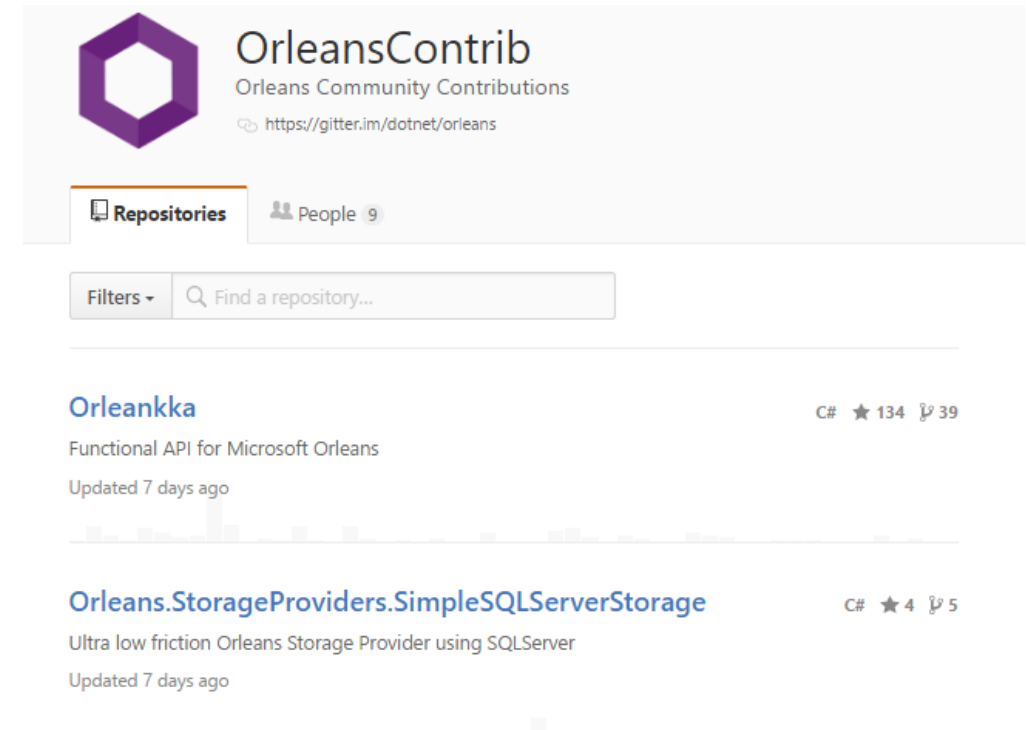
Do you want to contribute to F# community projects? We encourage all efforts which support our [mission](#). Some of the open community projects you can contribute to are below.

These projects may be at an early stage, and are not necessarily appropriate for production use. Out of small beginnings, great things come! To add a project to this list, log on to GitHub, [edit this page](#) and submit a pull request.

Tags: [Compiler](#), [Editing](#), [Dev Tools](#), [Data Access](#), [Data Science](#), [Visualization](#), [Functional Programming](#), [Text Programming](#), [Interop](#), [Parsing](#), [Testing](#), [Cloud](#), [Distribution](#), [Async/Parallel](#), [Scripting](#), [Web](#), [Financial](#), [GPU](#), [Desktop UI](#), [Math](#), [Text Search](#), [Metaprogramming](#), [Literate programming](#), [Audio Processing](#)

### Community Projects: Compiler and Core Library

-  [fsharp/fsharp](#) - The F# Compiler and Core Library. See also the [Visual F# Tools](#) and [Contributing to the](#)



**OrleansContrib**  
Orleans Community Contributions  
<https://github.com/dotnet/orleans>

Repositories People 9

Filters Find a repository...

**Orleankka** C# ★ 134 🍴 39  
Functional API for Microsoft Orleans  
Updated 7 days ago

**Orleans.StorageProviders.SimpleSqlServerStorage** C# ★ 4 🍴 5  
Ultra low friction Orleans Storage Provider using SQLServer  
Updated 7 days ago

# Degree of External Engagement

# Degree of External Engagement

- To help build communities we must understand their current degree external engagement.
  - The extent of project participation outside the organization.
- However, each project differs in several ways. E.g. Age, Target users.
- Therefore, we must take into account several factors when measuring external engagement.



# Sample Measures of External Engagement (Github)



Contributions (Commits, Lines of Code)



Number of Contributors



Number of Downloads / Users



Issues / Pull Requests (Discussions)

# Key Project Differentiators



Compilers / Prog. Language

*F#, Typescript, Roslyn*

Type of Project



Library / Engine /  
Framework / Toolkit

*CNTK, Room Alive, Chakracore*



Application

*IronClad, MSAGL, Madoko, Doc. Translator*

# Key Project Differentiators



New

## Different Stages



Inactive



Evolving



Stable Community



Winding Down

# Key Project Differentiators



End Users

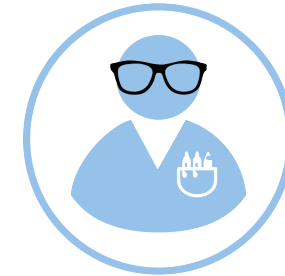
*Madoko, Doc. Translator*

Target Users /  
Niche Level



Developer

*Room Alive, Orleans*



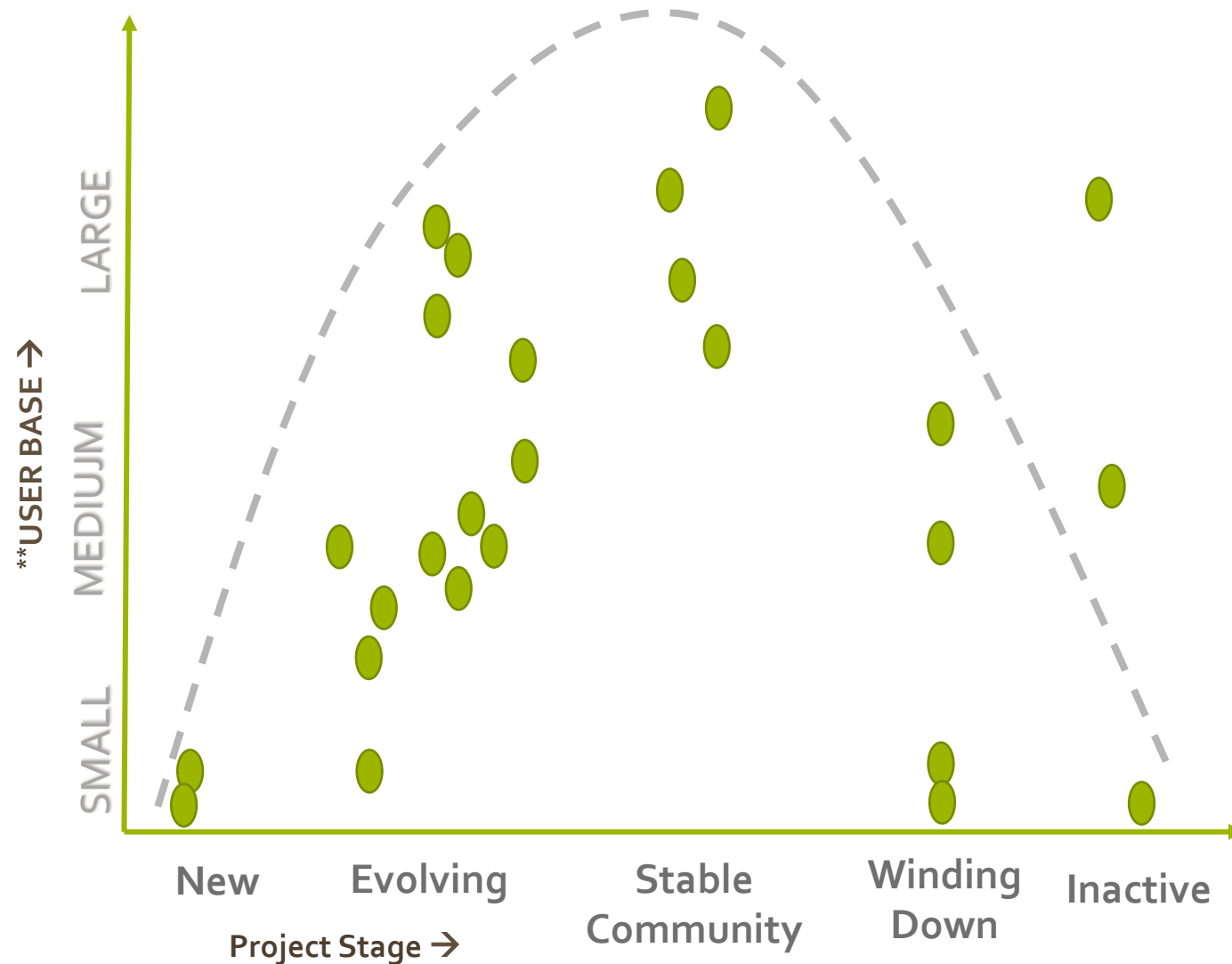
Advanced research

*Ironclad, Z3*

\*A project can be part of more than one Niche level



# Current Size of the User Base vs. Project Stage



\*\*Last 6 months

Small: 0 - 100

Medium: 100 - 1000

Large: > 1000

# Current No. of Contributors vs. Project Stage



*\*\*Last 6 months*

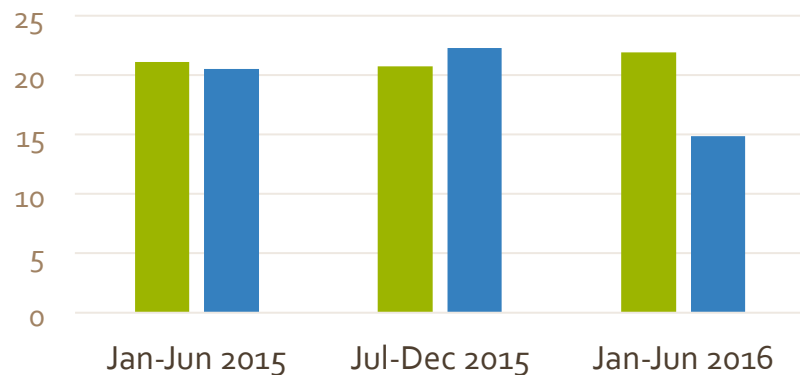
*Small: 0 - 10*

*Medium: 10 - 100*

*Large: > 100*

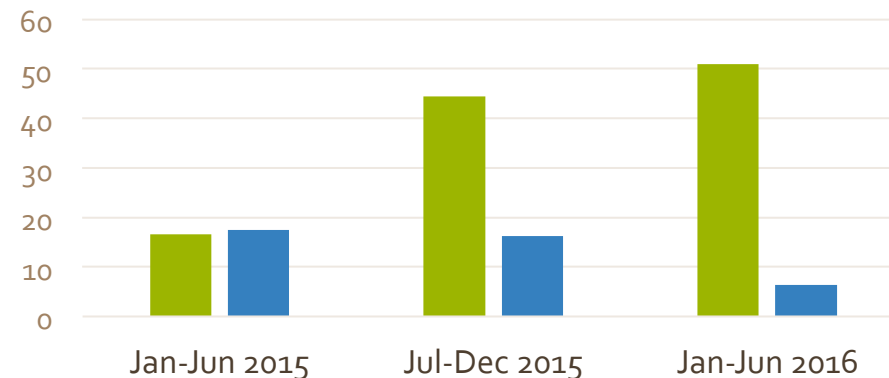
# Varying Degrees of External Engagement

ORLEANS (*Stable Community*)  
Contributions (Log)



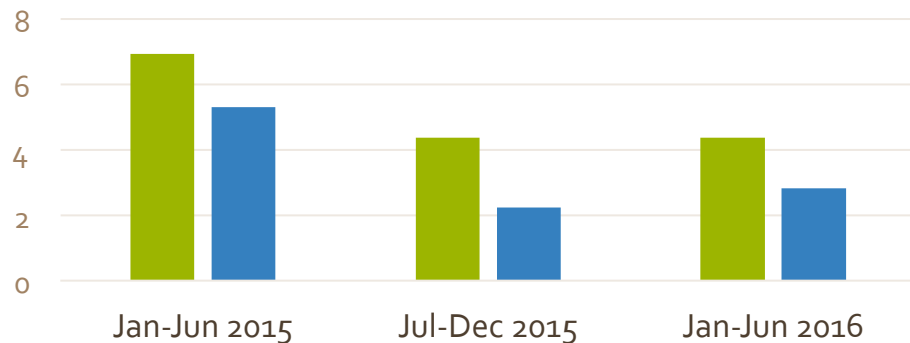
*MS and Non-MS contributions are almost equal*

CNTK (*Evolving*)  
Contributions (Log)



*Non-MS contributions are lower but MS contributors working hard to build the community.*

MSAGL (*Winding Down*)  
Contributions (Log)



*MS contributions reducing but community taking over*

■ MS Contributions (Commits)  
■ Outside Contributions (Commits)

\*Log scale helps better visualize the skewed data

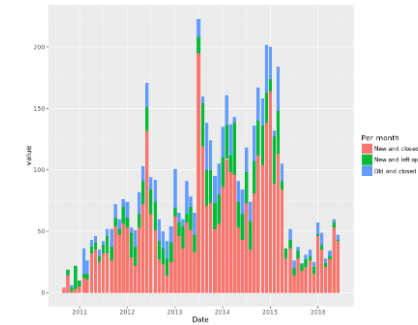
# Tools to Automate Our Analyses

- A few of tools exist to analyze Github Data:


- GHTORRENT

Pull request backlog

<http://ghtorrent.org/pullreq-perf/ipython-ipython/>



- Microsoft GHINSIGHTS (Under development)

 Microsoft / ghinsights

Watch 10Star 20Fork 3

Code

Issues 0

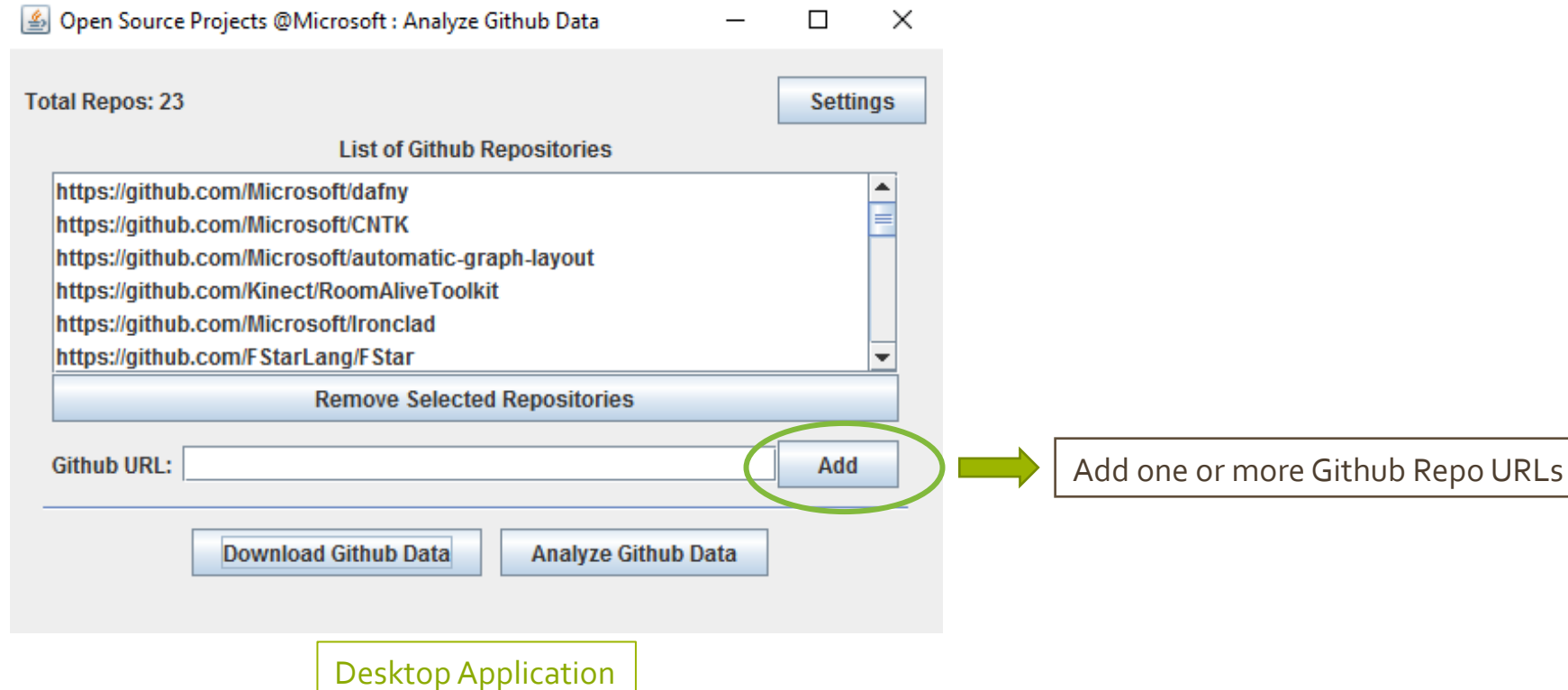
Pull requests 0

Pulse

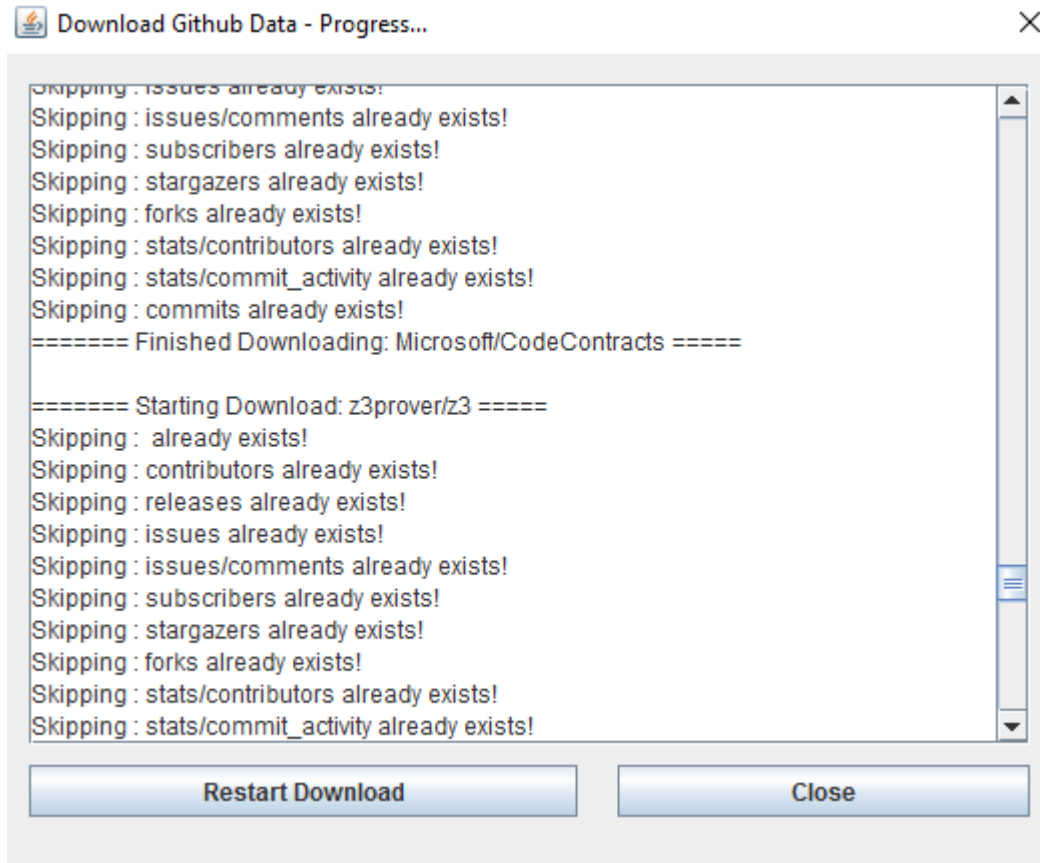
Graphs

GHInsights is a data processing pipeline using Azure Data Factory and Azure Data Lake. It processes GitHub data from the ghtorrent project. The resulting processed data is available in Azure Data Lake for users to query, generate reports, and analyze GitHub projects.

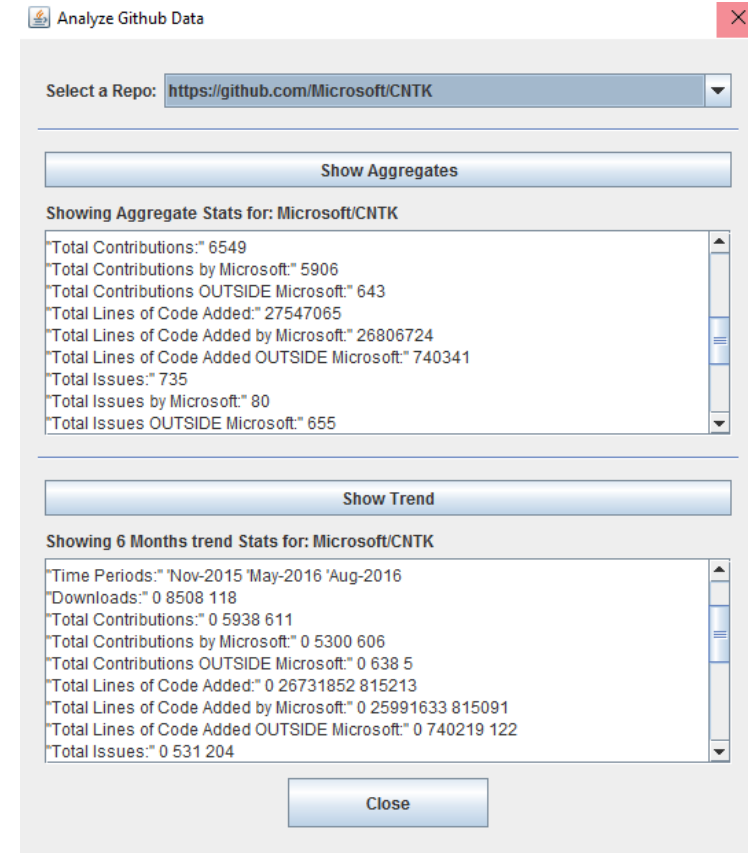
# A Simpler Tool to Automate Our Analyses



# Quick Walkthrough



Step1: Download Data



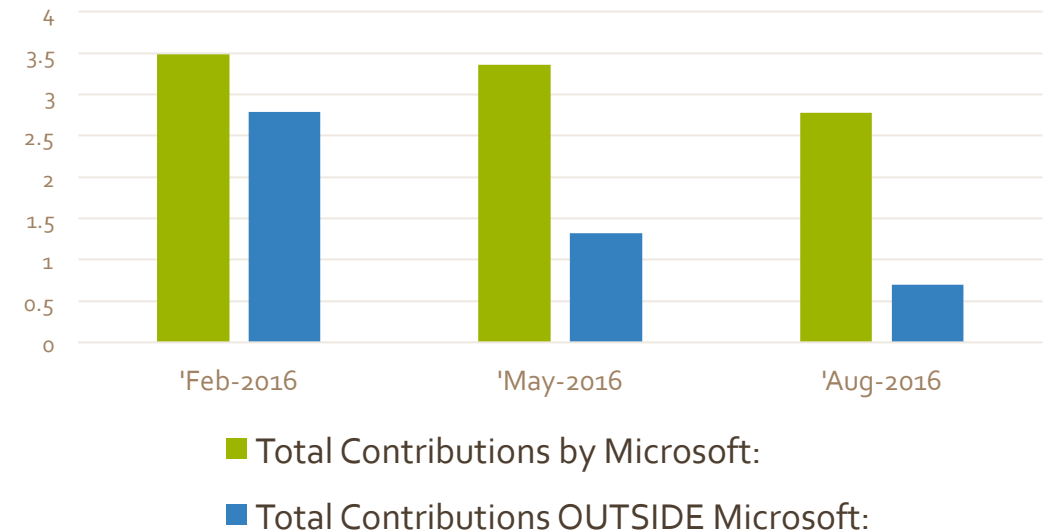
Step 2: Analyze Data

# Simply Copy and Analyze Data in Excel

Repository:	Microsoft/CNTK		
Start Date:	11/26/2015		
Current Date:	8/4/2016		
Time Periods:	'Nov-2015	May-2016	'Aug-2016
Downloads:	0	8508	118
Total Contributions:	0	5938	611
Total Contributions by Microsoft:	0	5300	606
Total Contributions OUTSIDE Microsoft:	0	638	5
Total Lines of Code Added:	0	26731852	815213
Total Lines of Code Added:	0	25991633	815091
Total Lines of Code Added:	0	740219	122
Total Issues:	0	531	204
Total Issues by Microsoft:	0	72	8
Total Issues OUTSIDE Microsoft:	0	459	196
Total Pull Requests:	0	60	4
Total Pull Requests by Microsoft:	0	12	1
Total Pull Requests OUTSIDE Microsoft:	0	48	3
Open Pull Requests:	0	7	2
Open Pull Requests by Microsoft:	0	2	0
Open Pull Requests OUTSIDE Microsoft:	0	5	2

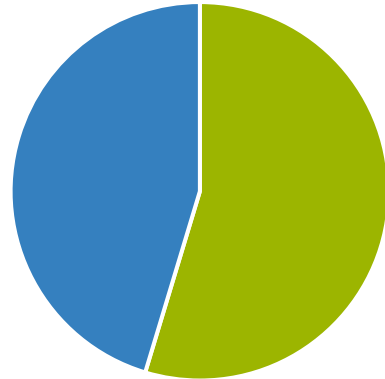


CNTK Contributions MS vs. Outside MS (Log Scale)



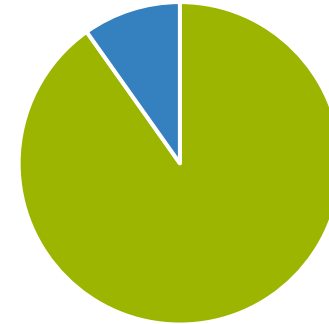
# Aggregate Data Analysis

CNTK: Compare number of Contributors



- Total Contributors in Microsoft:
- Total Contributors OUTSIDE Microsoft:

CNTK: Compare number of Contributions (Commits)

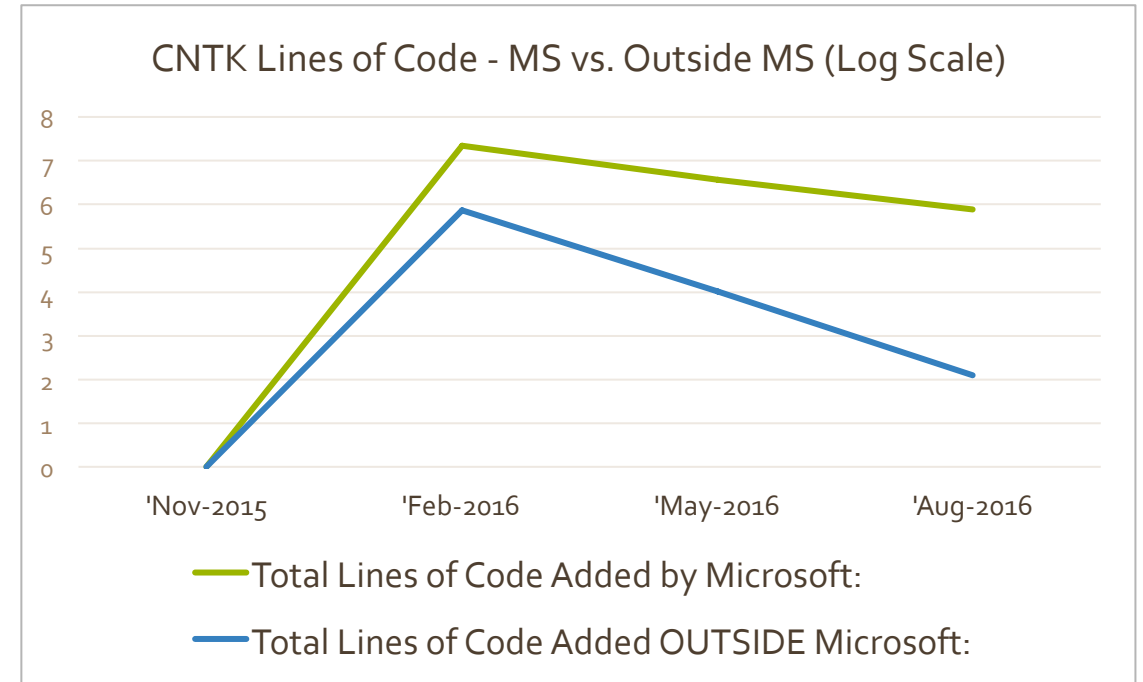
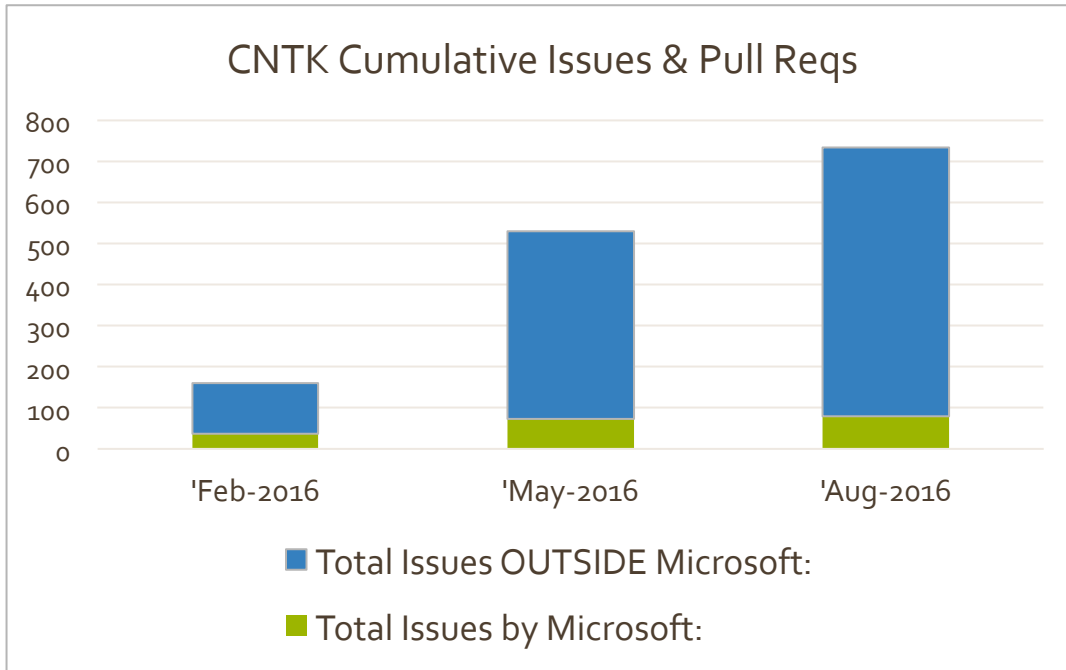


- Total Commits by Microsoft:
- Total Commits OUTSIDE Microsoft:

*Interesting that they have equal number of contributors but the contributions within MS is way higher.*



# Trend Analysis



***Signs of an Evolving Community:*** Most issues are reported from Outside MS but Code changes are done at Microsoft.

# Thank you

- Questions

# BACKUP SLIDES

# List of projects

Project	Org.
DSSM (Send2Vec)	MSR
Z3	MSR
Dafny	MSR
CNTK	MSR
MSAGL	MSR
Roomalive	MSR
Ironclad	MSR
F*	MSR
Touch Develop	MSR
PXT	MSR
Codalab	MSR
Orleans	Product Group
Madoko	MSR
Automata Toolkit	MSR
SEAL	MSR
LEAN	MSR
TPM Software Stack	MSR
Document Translator	MSR
F#	Product Group
Tabular	MSR
Scalable Joins	MSR
DDD / IDD Interactive Data Display	MSR
Scientific Dataset SDS	MSR
Typescript	Product Group
DotNet Micro Framework	Product Group
ChakraCore	Product Group
.Net Core	Product Group
Roslyn Compilers	Product Group
Code Contracts	MSR
LIQUID	MSR
WinAppDriver	Product Group

# Categorization

Compiler / Programming Language	Library / Engine	Application
Tabular F* Dafny LEAN F# TypeScript Roslyn Compilers .Net Core	DSSM - Sendzvec Automata Toolkit SEAL (NEW) Scalable Joins Interactive Data Display Scientific Dataset SDS .Net Micro Framework TPM Software Stack ChakraCore Z3 CNTK Orleans	IronClad PXT RoomAlive Codalab Document Translator MSAGL TouchDevelop Madoko

# Categorization

	Small User Base	Medium User Base	Large Users Base
Compiler / Programming Language	Tabular	F*	Dafny LEAN F# TypeScript Roslyn Compilers .Net Core
Library / Engine	DSSM - Sendzvec Automata Toolkit SEAL (NEW) Scalable Joins Interactive Data Display Scientific Dataset SDS .Net Micro Framework	TPM Software Stack ChakraCore RoomAlive	Z3 CNTK Orleans
Application	IronClad PXT (NEW)	Codalab Document Translator MSAGL	TouchDevelop Madoko

# Categorization

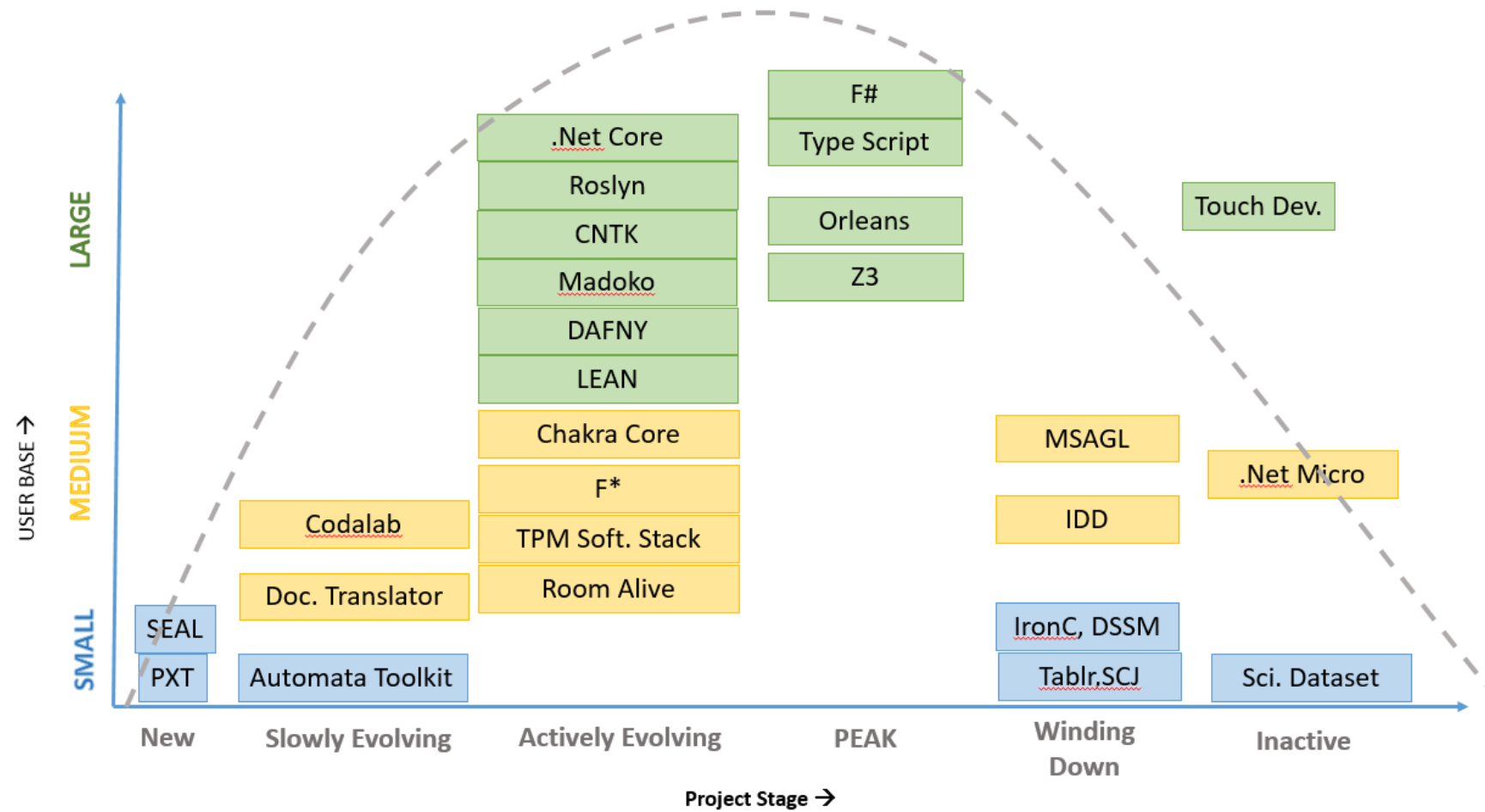
New	Evolving slowly	Evolving Actively	At Peak (or Stable community)	Winding Down	Inactive
PXT SEAL	Automata Toolkit Codalab Document Translator	RoomAlive F* TPM Software Stack ChakraCore Madoko Dafny CNTK LEAN Roslyn Compilers .Net Core	Z3 Orleans F# TypeScript	DSSM IronClad Tabular Scalable Joins MSAGL Interactive Data Display	TouchDevelop Scientific Dataset .Net MicroFramework

# Categorization

	New	Evolving slowly	Evolving Actively	At Peak	Winding Down	Inactive
Small User Base	PXT SEAL	Automata Toolkit			DSSM IronClad Tabular Scalable Joins	Scientific Dataset
Medium User Base		Codalab Document Translator	RoomAlive F* TPM Software Stack ChakraCore		MSAGL Interactive Data Display	.Net MicroFramework
Large User Base			Madoko Dafny CNTK LEAN Roslyn Compilers .Net Core	Z3 Orleans F# TypeScript		TouchDevelop



# Plot



# Individual Project Motivations

	Ease of Collaboration	Prevent project from dying	Increase trustworthiness	Get more Contributions	Involve the community from start	Serve the Community	Reproduce Results	Improve Quality	Increase adoption	Self-Document	Attract Developers	Faster turnaround	Make it more portable	Increase internal users	Request from users
1	IDD	Scientific Dataset	Madoko	CNTK	PXT	Z3	DSSM	CNTK	Orleans	Automata Toolkit	Chakra Core	.Net Core	.Net Core	WinApp Driver	MSAGL
2	Scalable Joins	.Net MF	SEAL	Document Translator	LEAN	RoomAlive	Ironclad	F#							
3	F*	Code Contracts	TPM	F#	Typescript										
4	Codalab	Scalable Joins													
5	.Net Core	Orleans													
6	Roslyn														
7	LEAN														