

Meeting of the LF AI & Data Technical Advisory Council (TAC)

October 20, 2022

 LF AI & DATA

Antitrust Policy

- › Linux Foundation meetings involve participation by industry competitors, and it is the intention of the Linux Foundation to conduct all of its activities in accordance with applicable antitrust and competition laws. It is therefore extremely important that attendees adhere to meeting agendas, and be aware of, and not participate in, any activities that are prohibited under applicable US state, federal or foreign antitrust and competition laws.
- › Examples of types of actions that are prohibited at Linux Foundation meetings and in connection with Linux Foundation activities are described in the Linux Foundation Antitrust Policy available at <http://www.linuxfoundation.org/antitrust-policy>. If you have questions about these matters, please contact your company counsel, or if you are a member of the Linux Foundation, feel free to contact Andrew Updegrove of the firm of Gesmer Undergone LLP, which provides legal counsel to the Linux Foundation.

Recording of Calls

Reminder:

TAC calls are recorded and available for viewing on the [TAC Wiki](#)

Reminder: LF AI & Data Useful Links

- › Web site: lfaidata.foundation
- › Wiki: wiki.lfaidata.foundation
- › GitHub: github.com/lfaidata
- › Landscape: <https://landscape.lfaidata.foundation> or <https://l.lfaidata.foundation>
- › Mail Lists: <https://lists.lfaidata.foundation>
- › Slack: <https://slack.lfaidata.foundation>
- › Youtube: <https://www.youtube.com/channel/UCfasaeqXJBCAJMNO9HcHfbA>
- › LF AI Logos: <https://github.com/lfaidata/artwork/tree/master/lfaidata>
- › LF AI Presentation Template: https://drive.google.com/file/d/1eiDNJvXCqSZHT4Zk_-czASlz2GTBRZk2/view?usp=sharing

- › Events Page on LF AI Website: <https://lfaidata.foundation/events/>
- › Events Calendar on LF AI Wiki (subscribe available): <https://wiki.lfaidata.foundation/pages/viewpage.action?pageId=12091544>
- › Event Wiki Pages: <https://wiki.lfaidata.foundation/display/DL/LF+AI+Data+Foundation+Events>

Agenda

- › Roll Call (2 mins)
- › Approval of Minutes from previous meetings (2 mins)
- › ONNX – Project review (20 minutes)
- › Adlik – Project review (20 minutes)
- › LF AI General Updates (2 min)
- › Open Discussion (2 min)

TAC Voting Members - Please note

Please ensure that you do the following to facilitate smooth procedural quorum and voting processes:

- Change your Zoom display name to include your First/Last Name, Company/Project Represented
 - example: Nancy Rausch, SAS
- State your First/Last Name and Company/Project when submitting a motion
 - example: First motion, Nancy Rausch/SAS

TAC Voting Members

Note: we still need a few designated backups specified on [wiki](#)

Member Representatives (8 out of 16 required for quorum)

Member Company or Graduated Project	Membership Level or Project Level	Voting Eligibility	Country	TAC Representative	Designated TAC Representative Alternates
4paradigm	Premier	Voting Member	China	Zhongyi Tan	
Baidu	Premier	Voting Member	China	Ti Zhou	Daxiang Dong, Yanjun Ma
Ericsson	Premier	Voting Member	Sweden	Rani Yadav-Ranjan	
Huawei	Premier	Voting Member	China	Howard (Huang Zhipeng)	Charlotte (Xiaoman Hu) , Leon (Hui Wang)
Nokia	Premier	Voting Member	Finland	@ Michael Rooke	@ Jonne Soininen
OPPO	Premier	Voting Member	China	Jimmy (Hongmin Xu)	
SAS	Premier	Voting Member	USA	*Nancy Rausch	JP Trawinski
ZTE	Premier	Voting Member	China	Wei Meng	Liya Yuan
Adversarial Robustness Toolbox Project	Graduated Technical Project	Voting Member	USA	Beat Buesser	
Angel Project	Graduated Technical Project	Voting Member	China	Bruce Tao	Huaming Rao
Egeria Project	Graduated Technical Project	Voting Member	UK	Mandy Chessell	Nigel Jones, David Radley, Maryna Strelchuk, Ljupcho Palashevski, Chris Grote
Flyte Project	Graduated Technical Project	Voting Member	USA	Ketan Umare	
Horovod Project	Graduated Technical Project	Voting Member	USA	Travis Addair	
Milvus Project	Graduated Technical Project	Voting Member	China	Xiaofan Luan	Jun Gu
ONNX Project	Graduated Technical Project	Voting Member	USA	Alexandre Eichenberger	Prasanth Pulavarthi, Jim Spohrer
Pyro Project	Graduated Technical Project	Voting Member	USA	Fritz Obermeyer	

*Current TAC Chairperson

Minutes approval

Approval of August 25, 2022 Minutes

Draft minutes from the August 25 TAC call were previously distributed to the TAC members via the mailing list

Proposed Resolution:

- › That the minutes of the August 25 meeting of the Technical Advisory Council of the LF AI & Data Foundation are hereby approved.

Approval of September 22, 2022 Minutes

Draft minutes from the September 22 TAC call were previously distributed to the TAC members via the mailing list

Proposed Resolution:

- › That the minutes of the September 22 meeting of the Technical Advisory Council of the LF AI & Data Foundation are hereby approved.

Annual Update for ONNX

October 20, 2022

<https://onnx.ai>

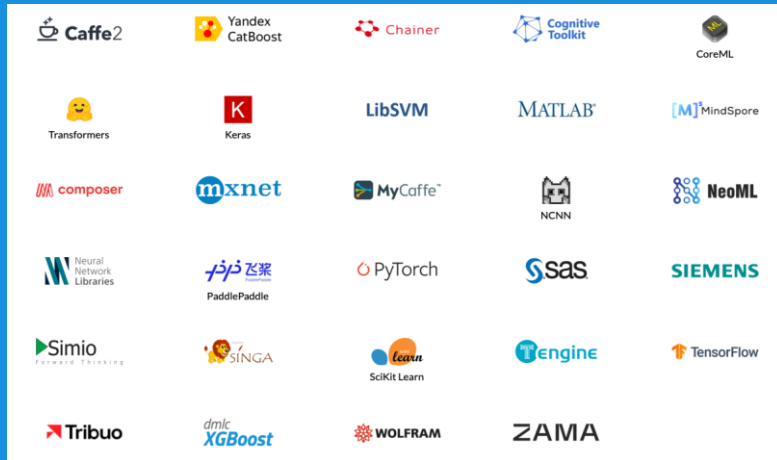
 LF AI & DATA

Open Neural Network Exchange (ONNX)

ONNX is

- › an ecosystem for vendor-neutral AI models,
- › a standard that defines an operation set and a file format.

Frameworks generating ONNX



ONNX Deployment Models



Visualize/Optimize ONNX

Key Achievements in the past year

Growing community and ecosystem

Hybrid community meetup

Regular releases and enhanced roadmap

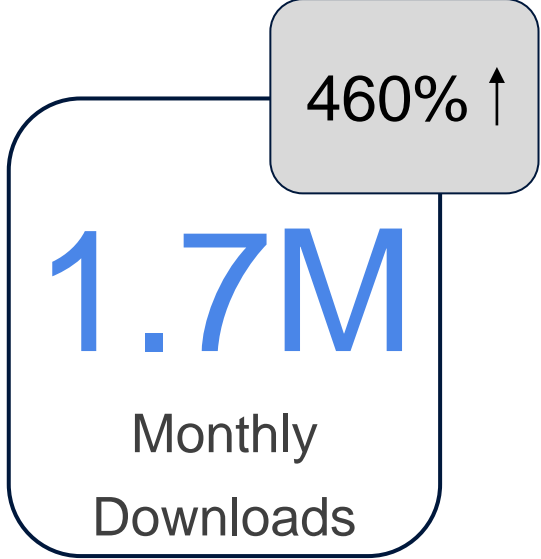
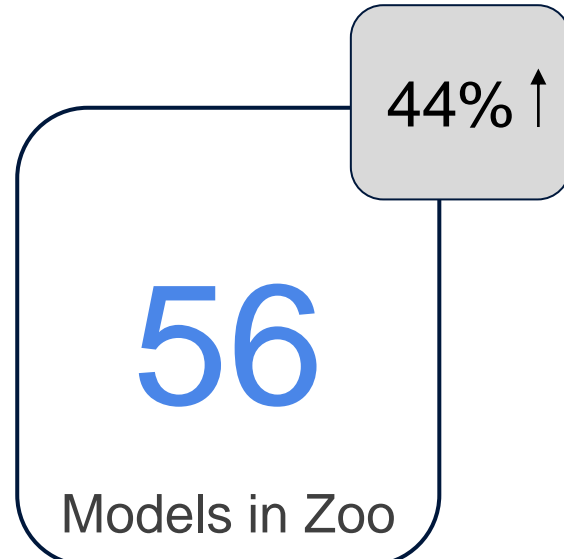
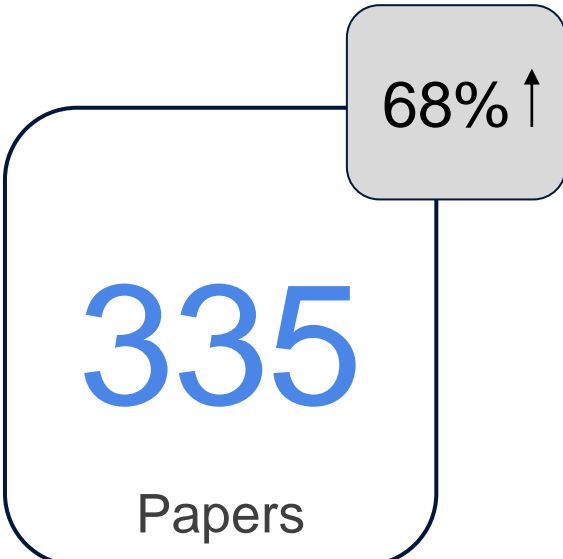
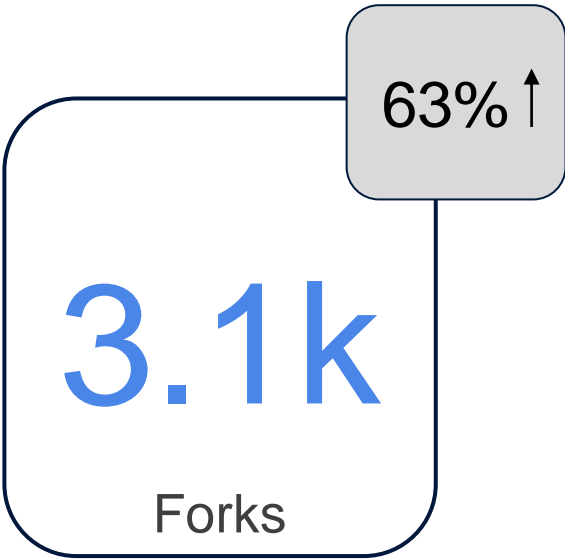
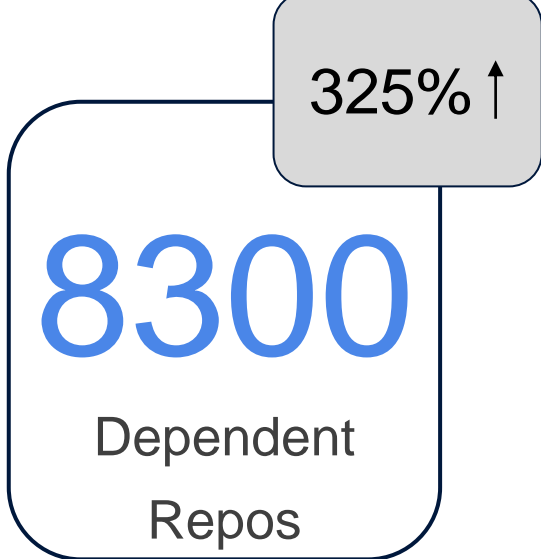
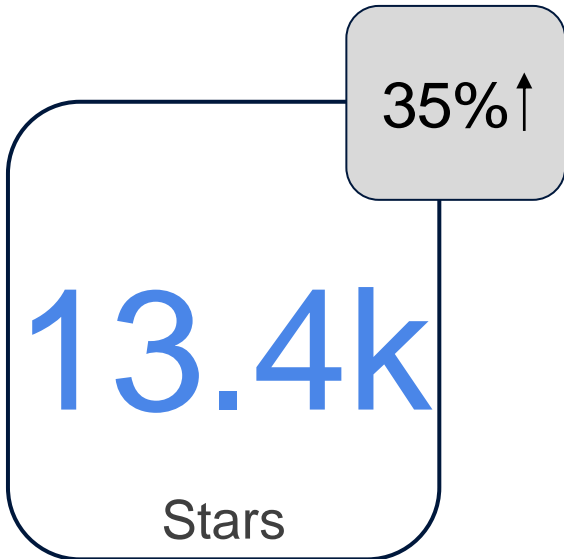
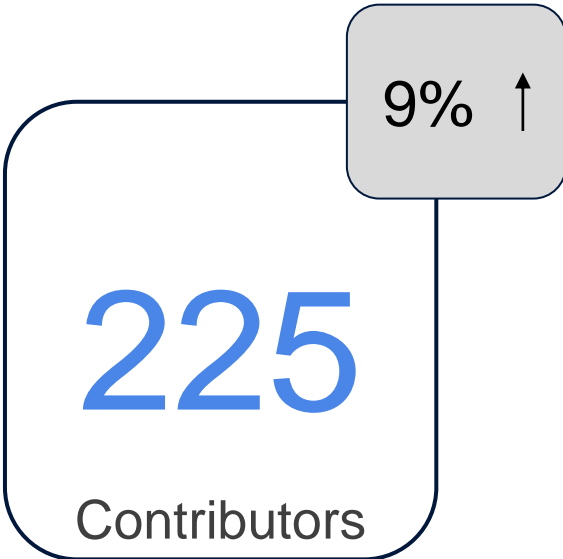
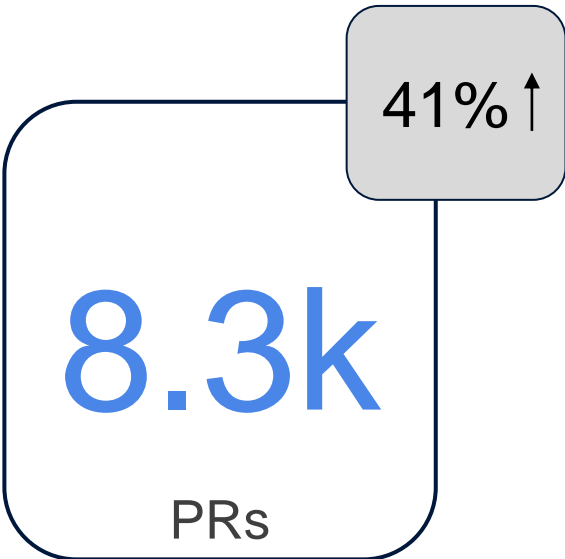
Improved documentation, YouTube channel, etc



ONNX

Community and
Ecosystem

Engagement & Usage from 3/21/21 to 10/20/2022



ONNX Community: Member Companies

ONNX Community Day: June 24, 2022

Full day in-person and online hybrid event, with YouTube live stream

Hosted by Microsoft

200+

registrations

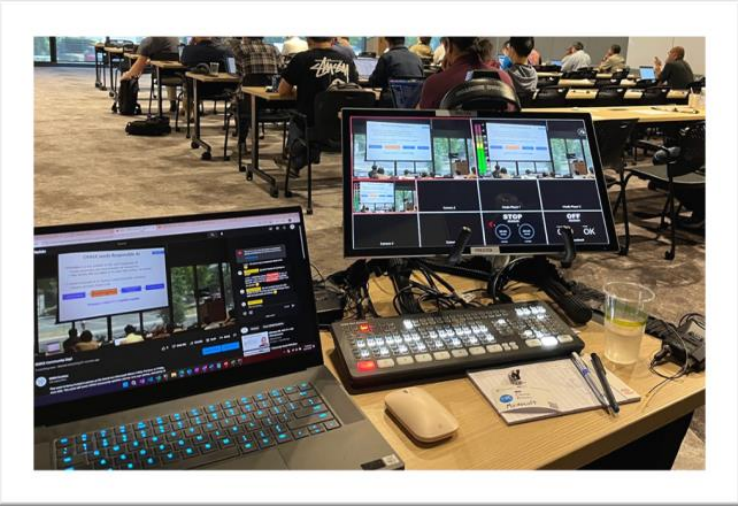
20+

talks presented

4.5 out of 5

attendee rating

Slides and recordings available at <https://onnx.ai/meetups/june2022>



Tool Support

Creation/ Manipulation



Run/ Compile



Visualization/ Test Tools

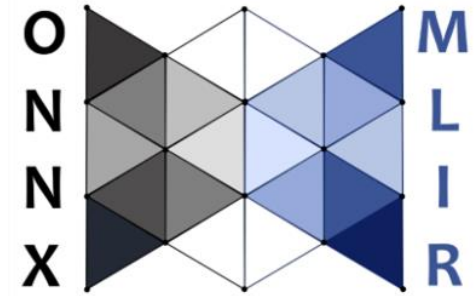


Ecosystem: ONNX Runtime

- › Cross-platform, high-performance ML engine
- › Supports all ONNX ops
- › Multiple language APIs: Python, C, C++, C#, Java, JavaScript, (Ruby, Julia, Go, Rust)
- › Broad accelerator support: TVM, NVIDIA, Intel, AMD, Qualcomm, etc
- › Over a Trillion inferences/day at Microsoft alone
- › ONNX Runtime also accelerates PyTorch training and enables on-device training.
- › Active OSS project: 400+ contributors representing many organizations and countries



Ecosystem: ONNX-MLIR Compiler



- › Compiler ecosystem based on MLIR-LLVM technology.
- › 1) Transforms ONNX models to binary code with accelerator support
 - › Tested for x86, POWER, Z under Linux, Windows, MacOS,
 - › Exploits integrated AI Accelerator in Telum (Z16),
 - › Current coverage: 70% of Model Zoo benchmarks.
- › 2) Brings ONNX to MLIR for TensorFlow and PyTorch ecosystem
 - › ONNX to Stable HLO (TensorFlow) and to PyTorch-MLIR.
- › C, C++, Java, Python bindings.
- › Supported by an active community
 - › 5 companies, 75 contributors, 650+ PRs/year, weekly public meetings.
 - › More info: <https://onnx.ai/onnx-mlir>



ONNX

Roadmap and
Releases

ONNX open governance update

Steering Committee

<https://github.com/onnx/steering-committee>

Prasanth Pulavarthi (Microsoft)
Alexandre Eichenberger (IBM)
Rajeev Nalawadi (Intel)
Mayank Kaushik (Nvidia)
Andreas Fehlner (TRUMPF Laser GmbH)

Working Groups (WGs)

<https://github.com/onnx/working-groups>

Preprocessing

- Joaquin Anton (NVIDIA)

Special Interest Groups (SIGs)

<https://github.com/onnx/sigs>

Architecture & Infra

- Liqun Fu (Microsoft)
- Ke Zhang (Ant Financial)

Operators

- Ganesan "Rama" Ramalingam (Microsoft)
- Michal Karzynski (Intel)

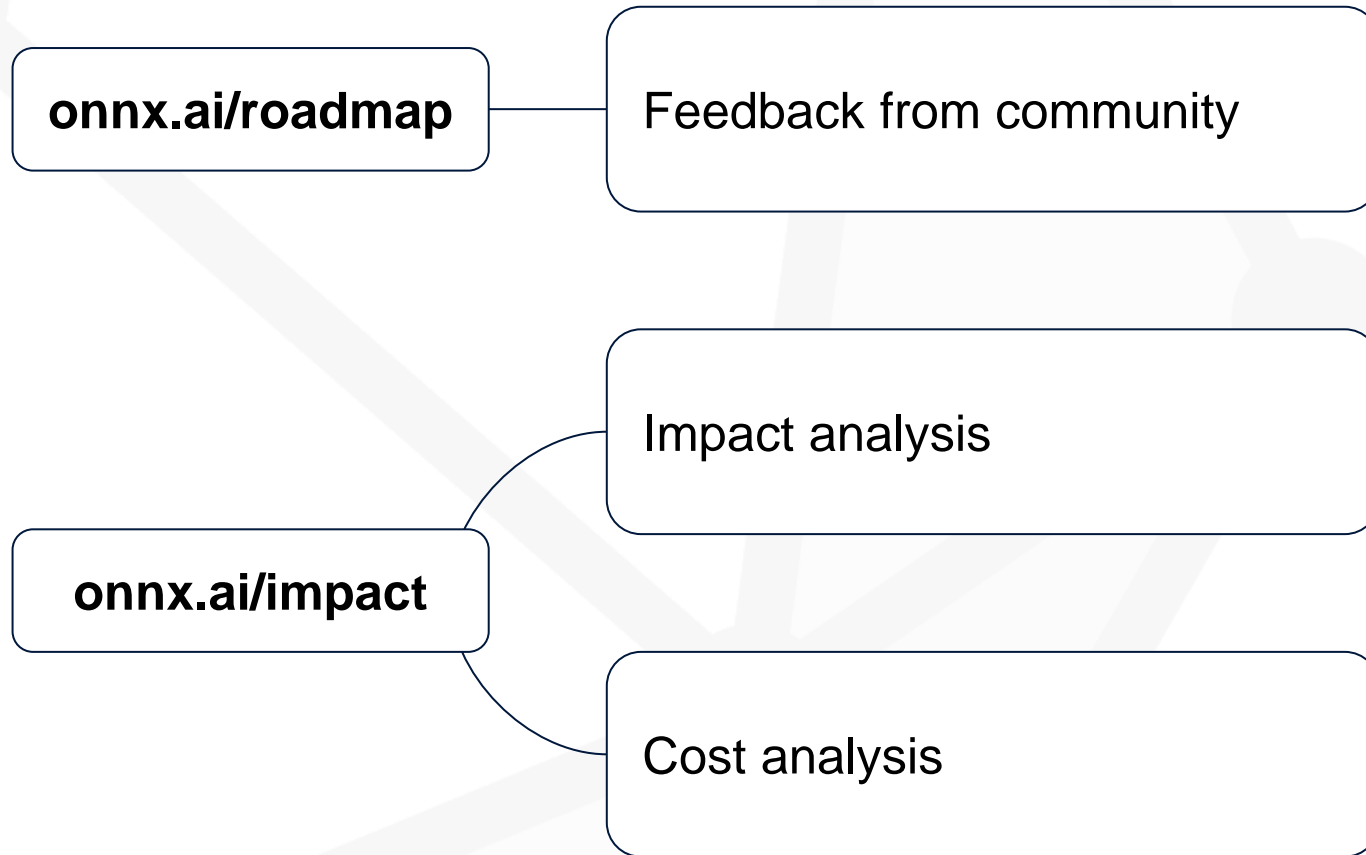
Converters

- Thiago Crepaldi (Microsoft)
- Kevin Chen (NVIDIA)

Model Zoo & Tutorials

- Jacky Chen (Microsoft)

ONNX roadmap discussions



Ongoing and Future Directions

Work on specs

- Clear and unambiguous specs: improve doc.
- Compact specs: complex ops as combination of primitive ops (ONNX functions).
- Add operators for newer models, pre/post processing

ONNX exploitation

- Compiler approach for ONNX models based on MLIR/LLVM project (onnx-mlir).
- Continued optimizations of ONNX runtimes.
- Better tools to author ONNX

Ongoing and Future Directions (cont.)

Improved conversions

- to ONNX: PyTorch 1.12 / TensorFlow 2.8 / Sklearn 1.11.2,
- from ONNX: Tensorrt 8.4 / TensorFlow 2.8.

ONNX Models

- 40 types of models (35 vision-based, 5 comprehension, quantized models),
- collaboration between ONNX model zoo & Hugging Face.

Preprocessing

- Composition & batch processing (completed)

Released: ONNX 1.11

ONNX v1.11 comes with following updates

- Opset 16 introduced with new and updated operators,
- **Model hub** (to pull pre-trained models from zoo),
- **Combine tool** to create combined model with preprocessing & inference,
- **Function builder** utility to help create function ops,
- Bugfixes, infrastructure improvements, documentation updates.

Visit the [release page on GitHub](#) for more details

Released: ONNX 1.12

ONNX v1.12 comes with following updates:

- Opset 17 introduced with new and updated operators,
- **Shape inference** enhancements,
- Bugfixes, infrastructure improvements, documentation updates,
- Added Python 3.10, dropped Python 3.6 support,
- Dropped support for x86 (32-bit) Linux due to low usage.

Visit the [release page on GitHub](#) for more details.

Coming up: ONNX 1.13

ONNX v1.13 plans to include the following updates:

- New operators: CenterCropPad, Mish, Col2Im, Bitwise ops
- Updated operators: Resize, ScatterElements, ScatterND, Split, Pad, OptionalHasElement, OptionalGetElement
- Add M1/M2 support for Mac with universal2 wheel
- Reduce primitive operator surface by defining some ONNX ops as functions

Release targeted for mid-December 2022.



ONNX

LF AI & Data
Opportunities

Areas the project could use help on

We want to explore stronger connection with LFAI projects.

- › Pytorch has strong ONNX components & support.
- › Encourage anyone to leverage ONNX.
- › Requesting help from the community to strengthen bonds between LFAI & ONNX.



ONNX

Thank you!

<https://onnx.ai>

Appendix

ONNX

[Open Neural Network Exchange \(ONNX\)](#) is an open ecosystem that empowers AI developers to choose the right tools as their project evolves. ONNX provides an open source format for AI models, both deep learning and traditional ML. It defines an extensible computation graph model, as well as definitions of built-in operators and standard data types. Currently we focus on the capabilities needed for inferencing (scoring).

ONNX is [widely supported](#) and can be found in many frameworks, tools, and hardware. Enabling interoperability between different frameworks and streamlining the path from research to production helps increase the speed of innovation in the AI community. We invite the community to join us and further evolve ONNX.

github: <https://github.com/onnx/onnx>

website: <https://onnx.ai/>



1.48M

Lines Of Code Changed



1.92K

Commits



153

Contributors



1

No Of Sub Projects



16

Repositories



Top 10 Contributors By Commits

[View All](#)

Ettore Tiotto	119.43K	244	8.04%
xiaowuhu	568.88K	131	38.31%
Alexandre Eichen...	32.29K	120	2.17%
Tung D. Le	45.83K	114	3.09%
Chun-Wei Chen	13.77K	112	0.93%
Xavier Dupré	411.38K	86	27.70%
Stella Stamenova	12.49K	69	0.84%
Prasanth Pulavar...	857	64	0.06%
Jianhao Zhang	9.08K	59	0.61%
Deyu Huang	5.37K	49	0.36%

Top 10 Organizations By Commits

[View All](#)





1.47K

Total PRs Submitted



2.5 months



Average Lead Time

105

PRs Yet To Be Merged



4.5 days



Average Time To First Review

8.1 days



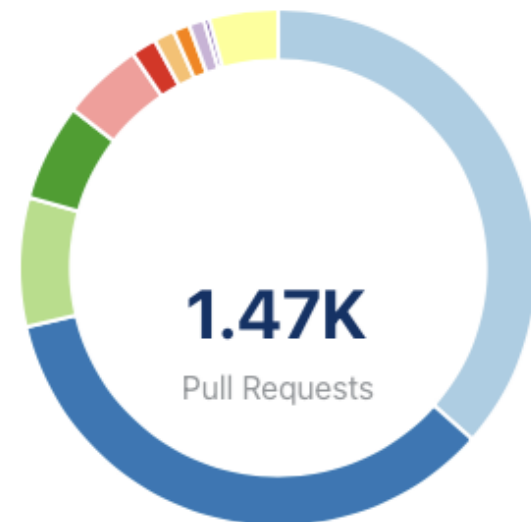
Average PR Cycle Time

Top 10 Contributors By PRs

[View All](#)

NAME	PULL REQUESTS	%
Chun-Wei Chen	123	8.40%
Tung D. Le	116	7.92%
Alexandre Eichenb...	109	7.44%
Ettore Tiotto	98	6.69%
Xavier Dupré	97	6.62%
Gong Su	48	3.28%
Deyu Huang	46	3.14%
Joaquin Anton	40	2.73%
TONG CHEN	40	2.73%

Top 10 Organizations By PRs

[View All](#)

- Microsoft Corporation
- International Business Machines Corporation
- Unknown
- Individual - No Account
- NVIDIA Corporation
- Intel Corporation
- Oneflow AB
- Groq, Inc.
- Preferred Networks, Inc.
- Hugging Face, Inc.
- Others



1.16K

Issues



719

Submitters



519

Open Issues



28 days



Time To Close

1.1 months



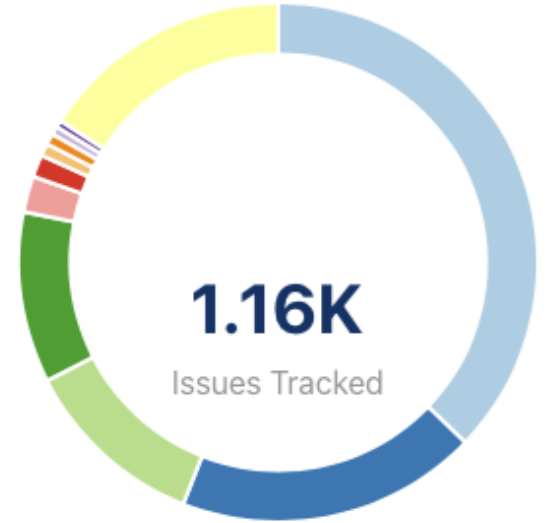
Time In Open

Top 10 Contributors By Issues Submitted

[View All](#)

NAME	ISSUE	%
Alexandre Eichenb...	33	2.85%
Chun-Wei Chen	33	2.85%
Ettore Tiotto	28	2.42%
Gary Miguel	19	1.64%
Tung D. Le	18	1.56%
TONG CHEN	17	1.47%
Charles Volzka	11	0.95%
Jay Zhang	11	0.95%
JinTian	9	0.78%

Top 10 Organizations By Issues Submitted

[View All](#)

- Individual - No Account
- Unknown
- Microsoft Corporation
- International Business Machines Corporation
- NVIDIA Corporation
- Intel Corporation
- Google LLC
- Preferred Networks, Inc.
- Huawei Technologies Co., Ltd
- Groq, Inc.
- Others

Annual Update for Adlik

October 20, 2022

Liya Yuan

<https://github.com/Adlik/Adlik>

Incubation/Graduation Project review criteria

To be accepted into the Graduation stage, a project must meet the Incubation stage requirements plus:

- Have a healthy number of code contributions coming from [at least five organizations](#).
- Have reached a [minimum of 1000 stars on GitHub](#).
- Have achieved and maintained a Core Infrastructure Initiative Best Practices [Gold Badge](#).
- Have demonstrated a substantial ongoing flow of commits and merged contributions for the past 12 months*.
- Receive the affirmative vote of two-thirds of the TAC and the affirmative vote of the Governing Board.
- Have completed at least one collaboration with another LF AI & Data hosted project
- Have a technical lead appointed for representation of the project on the LF AI & Data Technical Advisory Council.

Adlik

Brief Description:

Adlik is a toolkit for accelerating deep learning inference. The goal of Adlik is to accelerate deep learning inference process both on cloud and embedded environments. Adlik consists of two sub projects: model compiler and serving platform. Model compiler supports several optimizing technologies like pruning, quantization and structural compression to optimize models developed in major frameworks like Tensorflow, Keras, and Caffe, so that they can run with lower latency and higher computing efficiency. Serving platform provides deep learning models with optimized runtime based on the deployment environment such as CPU, GPU, and FPGA. Based on a deep learning model, the users of Adlik can optimize it with model compiler and then deploy it to a certain platform with serving platform.

With Adlik, different deep learning models can be deployed to different platforms with high performance and much flexibility.

Contributed by:

ZTE corporation in October 2019 as Incubation Project



Key Links:

Github: <https://github.com/Adlik/Adlik>

Website: <https://adlik.ai/>

Artwork:

<https://artwork.lfaidata.foundation/projects/adlik/>

Mailing lists:

<https://lists.lfaidata.foundation/g/adlik-announce>

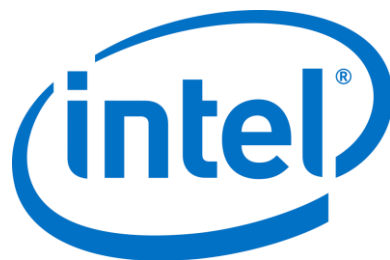
<https://lists.lfaidata.foundation/g/adlik-technical-discuss>

<https://lists.lfaidata.foundation/g/adlik-tsc>



Organizations contributing

ZTE中兴



北京大学
PEKING UNIVERSITY



Contributions



34.07K

Lines Of Code Changed



54

Commits



20

Contributors



1

No Of Sub Projects



4

Repositories



Top 10 Contributors By Commits

[View All](#)

NAME	LINES OF CODE	COMMITTS	%
靳伟昭 10307459	1.02K	9	3.01%
kellyZhang	19.90K	8	58.40%
王佳越 10335419	614	6	1.80%
nengyinyibeiwu	261	5	0.77%
Xiaoyang Zhang	2.89K	5	8.48%
wangjiayue	4.20K	3	12.34%
chehuier	4	2	0.01%
che.hui	10	2	0.03%
Shi mei	38	2	0.11%

Top 10 Organizations By Commits

[View All](#)



● ZTE Corporation ● Unknown

Key Achievements in the past year

- 2 Releases: V0.4, V0.5
- Some Key features
 - Optimizer: Auto pruning, combined optimization, int4 quantization framework
 - Compiler: Support compilation models from Paddle and OneFlow, support chips like Cambricon, Enflame
 - Inference Engine: Support Torch runtime; optimized implementation of some operators
 - New added repo: Model zoo
- Closer cooperation with more AI frameworks, e.g., Paddle, OneFlow and OpenVINO, and some hardware vendors
- A white book written together with Intel
- More new users
- Continuous blog update on WeChat and Zhihu channel

Areas the project could use help on

- › More contributors and users to join the community;
- › More promotion opportunities;

Feedback on working with LF AI & Data

- Benefit from LF AI & Data resources including website, Zoom, mailing lists, wiki space, etc.
- Great support for Adlik Community Meetings
- Great support for Adlik Release Blogs and Social Media
- Great opportunities for Adlik to reach out

TAC Open Discussion

Upcoming TAC Meetings

 **DLF** AI & DATA

Upcoming TAC Meetings

- › November 3, 2022 – Docarray, new incubation proposal
- › November 17, 2022 – Claimed, new sandbox proposal

Please note we are always open to special topics as well.

If you have a topic idea or agenda item, please send agenda topic requests to tac-general@lists.lfaidata.foundation

Open Discussion

 **OLF** AI & DATA

TAC Meeting Details

- › To subscribe to the TAC Group Calendar, visit the wiki:
<https://wiki.lfaidata.foundation/x/cQB2> _____
- › Join from PC, Mac, Linux, iOS or Android: <https://zoom.us/j/430697670>
- › Or iPhone one-tap:
 - › US: +16465588656,,430697670# or +16699006833,,430697670#
- › Or Telephone:
 - › Dial(for higher quality, dial a number based on your current location):
 - › US: +1 646 558 8656 or +1 669 900 6833 or +1 855 880 1246 (Toll Free) or +1 877 369 0926 (Toll Free)
- › Meeting ID: 430 697 670
- › International numbers available: <https://zoom.us/j/430697670>

Legal Notice

- › The Linux Foundation, The Linux Foundation logos, and other marks that may be used herein are owned by The Linux Foundation or its affiliated entities, and are subject to The Linux Foundation's Trademark Usage Policy at <https://www.linuxfoundation.org/trademark-usage>, as may be modified from time to time.
- › Linux is a registered trademark of Linus Torvalds. Please see the Linux Mark Institute's trademark usage page at <https://lmi.linuxfoundation.org> for details regarding use of this trademark.
- › Some marks that may be used herein are owned by projects operating as separately incorporated entities managed by The Linux Foundation, and have their own trademarks, policies and usage guidelines.
- › TWITTER, TWEET, RETWEET and the Twitter logo are trademarks of Twitter, Inc. or its affiliates.
- › Facebook and the "f" logo are trademarks of Facebook or its affiliates.
- › LinkedIn, the LinkedIn logo, the IN logo and InMail are registered trademarks or trademarks of LinkedIn Corporation and its affiliates in the United States and/or other countries.
- › YouTube and the YouTube icon are trademarks of YouTube or its affiliates.
- › All other trademarks are the property of their respective owners. Use of such marks herein does not represent affiliation with or authorization, sponsorship or approval by such owners unless otherwise expressly specified.
- › The Linux Foundation is subject to other policies, including without limitation its Privacy Policy at <https://www.linuxfoundation.org/privacy> and its Antitrust Policy at <https://www.linuxfoundation.org/antitrust-policy>. each as may be modified from time to time. More information about The Linux Foundation's policies is available at <https://www.linuxfoundation.org>.
- › Please email legal@linuxfoundation.org with any questions about The Linux Foundation's policies or the notices set forth on this slide.