This is a preliminary list of resources for those looking to get into the field of AI for Math, put together as part of the AI to Assist Mathematical Reasoning workshop (the official workshop proceedings can be found in the link here). It is not complete, but hopefully once distributed, it can continue to grow.

Please feel free to add comments or suggestions to this document. To add comments or suggestions (or to correct any existing text) you may either use the "comment" feature of Google Docs, or you may type directly into the document in a suggestion mode. Everyone who views this should have comment/suggestion access. We will review and approve suggestions.

Thank you to everyone who has contributed so far!

Education

Here are some educational materials that may be useful for folks interested in this space.

Textbooks and Survey Papers

- **Formal Proof**
  - Software Foundations series
  - Certified Programming with Dependent Types (Coq, 2013)
  - QED at Large: A Survey of Engineering of Formally Verified Software (2020)
  - AMS Special Issue on Formal Proof (2008)
  - Concrete Semantics (Isabelle 2014)
  - Mathematics in Lean (Lean, Jeremy Avigad and Patrick Massot, in progress)
  - Logic and Mechanized Reasoning (Lean, Jeremy Avigad, Marijn Heule and Wojciech Nawrocki, in progress)
  - Theorem Proving in Lean 4 (Lean, Jeremy Avigad, Leonardo de Moura, Soonho Kong and Sebastian Ullrich)
  - Logic and Mechanized Reasoning (Lean, CaDiCal, Z3, Vampire; Avigad, Heule, Nawrocki, 2021)
  - The Mechanics of Proof (Lean, Heather Macbeth, 2023)
  - Mathematical Components book (Coq, Mahboubi & Tassi, 2022).
  - Handbook of Practical Logic and Automated Reasoning (OCaml, John Harrison, 2009, code and resources)
- **Machine Learning / Deep Learning**
  - Grokking Deep Learning (2019)
  - Neurosymbolic Programming (2021)
  - Automatically Correcting Large Language Models (2023)
  - Deep Learning (2016)
  - Probabilistic Machine Learning: An Introduction (reference) (2022)
  - Understanding Deep Learning (2024)
- **Machine Learning for Formal Proof**
  - A Survey on Deep Learning for Theorem Proving (2024)

Wikis and Glossaries

- **Programming Languages**
  - 1lab for homotopy type theory (https://homotopytypetheory.org/)
- **Mathematics**
  - MathGloss (https://mathgloss.github.io/MathGloss/database)
  - nLab (https://ncatlab.org/nlab/show/HomePage)
  - Error Correction Zoo (https://errorcorrectionzoo.org/)
  - Parmesan (http://www.jacobcollard.com/parmesan2/) concepts in Cat Theory
  - Mathworld (https://mathworld.wolfram.com/)

Tutorials

- **Formal Proof**
  - Natural Number Game (Lean)
  - Homotopy Type Theory Game (Cubical Agda)
  - Isabelle/HOL tutorial
  - HOL4 tutorials and guidebooks
  - Lean for the Curious Mathematician workshop materials and recordings: 2020, 2022, 2023, 2024
  - Learning Lean community webpage
  - Terry Tao's Lean phrasebook
  - A publicly editable list of formalized concepts in Lean (additions greatly welcome)
  - List of Lean tactics
- **Machine Learning for Formal Proof**
  - A tutorial on neural theorem proving by Sean Welleck
  - NeurIPS Tutorial on Machine Learning for Theorem Proving by Emily First, Albert Jiang, and Kaiyu Yang
- **Machine Learning**
  - Numerical Methods for Deep Learning (lecture slides updated 2021).

Course Materials

- Proof Automation by Talia Ringer
- Formalising Mathematics by Kevin Buzzard
- Machine Learning by Andrew Ng
- Machine Learning for the Working Mathematician seminar in Sydney
- Software Foundations course at Penn
- Programming language foundations in Agda by Philip Wadler
- Lean for teaching stream on the Lean Zulip and courses webpage
- Real Analysis (in French) by Patrick Massot
- The Hitchhiker's Guide to Logical Verification by Anne Baanen et al
- The Concrete Semantics by Nipkow and Klein
- Introduction to Homotopy Type Theory by Egbert Rijke
- HoTTEST summer school - covers homotopy type theory and formalization in Agda

Blogs, Opinion Pieces, and Personal Experiences

- Mathematics and the Formal Turn (survey paper by Avigad)
- Why formalize mathematics? (survey paper by Massot)
- QED Manifesto (manifesto, 1994) (2014 follow-up)
- Is deep learning a useful tool for the pure mathematician?, Williamson
- A volume of the Bulletin of the American Mathematical Society dedicated to AI, formal proof, and mathematics (first of a two-volume series 2024).
- Embracing change and resetting expectations, Tao

- Machine assisted proof, Tao
- Formalising Mathematics – in Praxis; A Mathematician's First Experiences with Isabelle/HOL and the Why and How of Getting Started, Koutsoukou-Argyraki
- Formalization and Automated Reasoning: A Personal and Historical Perspective, John Harrison, 2023

Collaboration

This is a highly collaborative intersection of fields, so it is very helpful to know where to connect with people who may have complementary expertise, experience, or interests.

Forums

- Lean Zulip: A discussion medium for the community around the Lean proof assistant (including an active stream on "Machine Learning for Theorem Proving").
- Coq Zulip: A discussion medium for the community around the Coq proof assistant.
- EleutherAI Discord: A discussion medium for an open source AI community centered around the open source AI nonprofit EleutherAI. There is a channel specifically for #ai-for-math. Fantastic for forming collaborations. In dire need of more active involvement from academics.
- Neurosymbolic AI Slack for discussing the intersection of neural and symbolic methods for machine learning (theorem proving and math are domains of interest)
- Univalent Agda Discord for the community around Cubical Agda
- Isabelle Zulip for the community around the Isabelle proof assistant
- HoTT Zulip for the community around Homotopy Type Theory
- Proof Assistant Stack Exchange
- Code4Math

Tools and Repositories

A non-exhaustive list of tools that may be useful for getting started in this space.

ML Frameworks

- PyTorch
- Tensorflow
- JAX
- MindSpore
- PaddlePaddle
- HuggingFace 🤗 — Tokenizers, Transformers and other frameworks
- Weights & Biases for experiment tracking

Proof Assistants

One promising direction of research in AI for Math is combining AI-based automation with machine-checkable proof. Here is a non-exhaustive list of tools that can be used for writing machine-checkable proofs, in no particular order:

- Lean
  - Active community of mathematicians
  - Has an especially rich mathematical library: Mathlib
  - Strong automation with mathematicians in mind
  - Foundations:
    - Dependent type theory (Calculus of Inductive Constructions)
    - Constructive, but common to use classical axioms (and automation does by default these days)
    - Common to assume functional extensionality and get proper quotients
    - Proof irrelevant
    - Intensional
    - Explicit proof objects
  - Search engines: Loogle, Moogle, LeanSearch, Search Mathlib, LeanSearchClient
  - Package registry: Reservoir
  - Lean playground
  - Lean blueprint (a software for coordinating Lean formalization projects)
    - Pietro Monticone: Getting Started with Blueprint-Driven Formalization Projects in Lean (video)
      - LeanProject template repository
      - Zulip discussion
- Coq
  - Most math development in Coq happens in MathComp (most famously, the formal proof of the Four Color Theorem and the formal proof of the Odd Order Theorem)
  - Coq-platform, a very large maintained library of formal proofs
  - There is also UniMath and (https://github.com/HoTT/Coq-HoTT), an outgrowth of Voevodsky's foundations library that is still maintained by an active community
  - Strong automation
  - Foundations:
    - Dependent type theory (Calculus of Inductive Constructions)
    - Constructive, but compatible with classical axioms
    - Possible to assume functional extensionality and get proper quotients
    - Agnostic about proof relevance
    - Intensional
    - Explicit proof objects
  - jsCoq: try Coq in your web browser
- Isabelle
  - Rich, massive archive: Archive of Formal Proofs

    - Still best-in-class support for structured proofs with many features to aid readability and reduce the gap to “paper proofs”
    - Strong automation
        - Sledgehammer is still the most powerful “hammer” interfacing a very wide range of automatic first-order provers and using ML algorithms internally
    - Foundations:
        - Simply-Typed Lambda Calculus-based
        - Several logical systems to choose from
        - Isabelle/HOL (the most widely used) is a classical, higher-order logic
            - Isabelle/HOL’s developers and users are somewhat proud it **doesn’t** have Dependent Types. Some researchers conjectured that Dependent Types are necessary to formalize modern mathematics as long as most mathematical structures are indexed by values. Bordg, Paulson and Li set out to disprove this conjecture through the formalization of Grothendieck's Schemes without Dependent Types in Isabelle/HOL (https://www.tandfonline.com/doi/full/10.1080/10586458.2022.2062073).

              The paper also features a discussion of Isabelle/HOL schemes (no relation to Grothendieck) which are kind of more powerful Type Classes and could be very handy for structuring formal math libraries.
        - Isabelle/ZF implements ZF set theory
        - Ephemeral proof objects
    - Search engine: SErAPIS Isabelle Search, FindFacts
- HOL4
    - Foundations:
        - Logic-based
        - Classical, higher-order logic
        - LCF-style
- HOL Light
    - Foundations:
        - Logic-based
        - Classical, higher-order logic
        - LCF-style
- Agda
    - There is a rich univalent mathematical library Unimath here
    - Automation is minimal
    - Foundations:
        - Dependent type theory (Calculus of Inductive Constructions)
        - Constructive, but compatible with classical axioms
        - Agnostic about proof relevance
        - Intensional

        - Explicit proof objects
- Cubical Agda
    - Good for topology and homotopy theory, and for reasoning about how different proofs relate to each other
    - Has an absolutely beautiful and powerful notion of equality corresponding exactly to topological paths
    - Automation is extremely minimal
    - Can take a ton of expertise to get comfortable using at all
    - Foundations:
        - Dependent type theory (Homotopy Type Theory, specifically Cubical)
        - Univalent
        - Constructive
        - Proof relevant (at higher h-levels)
        - Has native functional extensionality and quotient types
        - Intensional
        - Explicit proof objects
- Other proof assistants or languages for verification
    - Nuprl
    - red* family of proof assistants
    - F*
    - Twelf
    - SasyLF
    - Metamath
    - Mizar
    - PVS
    - ACL2
    - ProofPower
    - Dafny
    - Boogie

Constraint Solvers & Automatic Theorem Provers

- CVC5 SMT Solver
- Vampire Theorem Prover
- E Theorem prover
- Z3 SMT Solver
- Rosette language for building tools that use constraint solvers
- Egg library for reasoning about equalities

Computational Mathematics Tools

- CoCalc (https://cocalc.com/, with access to ChatGPT assistance)
- CoCoa (https://cocoa.dima.unige.it/cocoa/)
- CGSuite (https://www.cgsuite.org/)

- Fermat (https://home.bway.net/lewis/)
- FLINT (https://flintlib.org/) (includes Arb, Antic, and Calcium)
- GAP (https://www.gap-system.org/)
- Julia (https://www.julialang.org/)
- Macaulay2 (https://macaulay2.com/)
- Magma (https://magma.maths.usyd.edu.au/)
- Mathematica (https://www.wolfram.com/mathematica/)
- Maple (https://www.maplesoft.com)
- MATLAB (https://www.mathworks.com/)
- Maxima (https://maxima.sourceforge.io/)
- Mathics (https://mathics.org/)
- GNU Octave (https://octave.org/)
- OSCAR (https://www.oscar-system.org/about/)
- PARI/GP (https://pari.math.u-bordeaux.fr/)
- Polymake (https://polymake.org/doku.php/start)
- REDUCE (https://reduce-algebra.sourceforge.io/)
- SageMath (https://sagemath.org/)
- SymPy (https://www.sympy.org)
- SciLab (https://www.scilab.org/)
- Singular (https://www.singular.uni-kl.de/index.php.html)
- The CompuTop.org Software Archive (https://nmd.web.illinois.edu/computop/)
- WolframAlpha (https://www.wolframalpha.com/)
- Wolfram System Modeler (https://www.wolfram.com/system-modeler/)
- Desmos (https://www.desmos.com/calculator)
- Geogebra (https://www.geogebra.org/)
- Yacas (https://github.com/grzegorzmazur/yacas)

Mathematics Databases

- MathBases – index of mathematical databases (https://mathbases.org/)
- PolyDB (https://db.polymake.org/)
- L-functions and Modular Forms Database (https://www.lmfdb.org/)
- ATLAS of Finite Group Representations, Atlas of Lie Groups and Representations, Groupprops
- Projective Planes of Small Order
- The online database of Vertex Operator Algebras and Modular Categories
- Catalogue of Lattices, Database of sphere packings, record tables, Packomania
- Error Correction Zoo
- OEIS, Inverse Symbolic Calculator
- DLMF
- Fungrim
- Mathematical Equations - EqWorld
- factordb.com
- elliptic multiple zeta values datamine

- Polynomials with Small Mahler Measure
- House of Graphs
- π-Base, Topospaces
- KnotInfo / LinkInfo, Knot Atlas
- Fanography, Calabi Yau data, Graded Ring Database
- Cantor's Attic, Googology Wiki, Reverse Mathematics Zoo, Forking and Dividing, Consequences of the Axiom of Choice
- Complexity Zoo
- LifeWiki (conwaylife.com) / Catagolue
- The Mathematics Genealogy Project (nodak.edu)
- Catalogue of Mathematical Datasets
- KnotFolio
- Graph curvature
- Erich's Packing Center
- Auslander-Reiten quiver for representation-finite gentle algebra
- Cayley Graph
- some other interactive databases

Mathematical Search Engines

- https://www.searchonmath.com/
- https://approach0.xyz/search/

Integrated AI for Math Tools

- LeanDojo (and also LeanCopilot) for Lean
- Proofster web tool for Coq, and corresponding plugin
- llmstep: LLM proofstep suggestions in Lean
- PISA (Portal for Isabelle)
- AlphaGeometry
- FunSearch

Datasets and Benchmarks

Note that many of the below can be used as training data or as evaluation benchmarks. Some come with a standard training/test split, and some do not. Regardless, care should be taken to ensure that test data does not pollute the training dataset whenever you are building any kind of tool, otherwise your results will not be valid. It is also worth noting that HuggingFace stores a lot of public datasets and benchmark suites, and so it is always good to look there as well.

- **Training Datasets**
  - **OpenWebMath**

    - The Pile (not entirely math but contains a large math subset that can be independently downloaded). This dataset has an associated paper and datasheet.
    - ProofPile
    - MetaMath
    - MPTP
    - REPLICA small atomic edit dataset for Coq
    - PRISM Proof Repair dataset for Coq (larger releases coming)
    - Isabelle Parallel Corpus (aligned natural and formal proof data)
    - MLFMF: Data Sets for Machine Learning for Mathematical Formalization
    - List of parts of the French curriculum that have been formalized in Lean's mathlib
    - Math Olympiad problem data sets
        - AIMO list of resources (see also this discussion)
        - External problem set of 25K problems
        - A Dataset for The Global Artificial Intelligence Championship Math 2024 (AGI Odyssey)
        - Compfiles: Catalog Of Math Problems Formalized In Lean

- **Libraries of formalized proofs**
    - Mizar Mathematical Library
    - Isabelle Archive of Formal Proofs
    - IMO solutions in Isabelle
    - IMO solutions in Lean
    - PutnamBench
- **Evaluation Benchmarks**
    - MATH
    - GSM8K
    - Some components of MMLU
    - CoqGym for Coq
    - LeanDojo for Lean
    - ProofNet autoformalization benchmark
    - Math Olympiad benchmarks: miniF2F, OlympiadBench
    - ARC Abstraction & Reasoning Corpus
    - ARB Advanced Reasoning Benchmark for Large Language Models
    - GHOSTS
    - "Evaluating Language Models for Mathematics through Interactions", Collins et al.

Language Models and Chatbots

Note: it is common for AI tools to be labeled as "open" or "open source" when they are downloadable by people outside the hosting institution. These often have substantial use-based or distribution restrictions that violate the norms of open source software. In this section we use the words “free and open source” in accordance with the OSI definition. For models labeled as “publicly available,” make sure to read the licenses carefully before making assumptions about permissible use.

General Purpose Models

- **Free and Open Source**
    - **Open Data**
        - GPT-J-6B
        - GPT-NeoX-20B
        - Pythia models (8 models ranging from 70M through 13B)
        - Scispace
    - **Limited or Non-Open Data**
        - Falcon
        - Mistral 7B
- **Publicly Available**
    - LLaMA 1
    - LLaMA 2 (7B, 14B, and 70B)
    - LLaMA 3
    - Code LLaMA
    - Falcon-180B
    - BLOOM (multilingual)
    - Command R+
    - DBRX
    - More can be found on the HuggingFace leaderboard
- **Gated by Public API**
    - PaLM
    - GPT-3.5
    - GPT-4
    - Claude Haiku
- **Private**
    - Chinchilla
    - Gopher

Mathematics Models

- **Free and Open Source**
    - **Open Data**
        - Llema
- **Publicly Available**
    - WizardMath
    - MetaMath
    - MathGLM
    - Abel

    - InternLM-Math
    - NuminaMath-7B-TIR
        - Math olympiad problem solver demo
    - Skywork-Math
- **Gated by Public API**
    - GPT-3.5 and GPT-4 with Wolfram Alpha Plug-In
- **Private**
    - Minerva (fine-tuned for math specifically)

ML for Formal Proof Models

- **Free and Open Source**
    - Proverbot9001
    - CoqGym/ASTactic
    - LeanDojo/ReProver
    - Passport
    - Diva
    - TacTok
- **Uncategorized**
    - TacticToe
    - Tactician
- **Private**
    - Baldur

Chatbots

- **Free and Open Source**
    - Falcon-Instruct-7B
    - Falcon-Instruct-40B
- **Publicly Available**
    - LLaMA 2/LLaMa 3 Chat
    - Stable Beluga 2
    - Falcon-Chat-180B
- **Gated by Public API**
    - Gemini (formerly Bard)
    - ChatGPT

Math OCR

- https://mathpix.com/
- https://simpletex.cn/ai/latex_ocr

- Nougat (facebookresearch.github.io) the META library OCR for extracting formulas and tables from pdf papers the library is licensed with MIT license and the code can be used and the paper is quite explanatory.
- Pix2Text (https://github.com/breezedeus/pix2text)

Research

Research that is published in this space and where to find it.

Meta-Bibliography

Some bibliographies listing relevant work in this area:

- Paper list - Deep Learning for Theorem Proving (DL4TP)
- Paper list - Deep Learning for Mathematical Reasoning (DL4MATH)
- Advancing mathematics by guiding human intuition with AI.
- NeurIPS 2023 Tutorial on Machine Learning for Theorem Proving (TBA)
- Paper list - Machine Learning for Theorem Proving
- Talia's list of ml-for-proofs papers
- Mathlib's references.bib
- Is deep learning a useful tool for the pure mathematician?, Geordie Williamson
- Large Language Models for Mathematicians, Frieder et al.
- Proof engineering bibliography from QED at Large (2019)
- Haocheng Jiu's AI4Math bibliography

Venues

These are venues where work in AI for Math or related areas may be found.

- **AI for Math**
    - AITP
    - NeurIPS AI for Math Workshop (3rd Workshop, Dec 15, 2023)
    - 2022 Fields Medal Symposium: Akshay Venkatesh (Oct 17-19, 2022)
    - IPAM Machine Assisted Proofs (Feb 13-17, 2023)
    - AI to Assist Mathematical Reasoning: A Workshop (Jun 12-14, 2023)
    - Dagstuhl 23401 on Automated Mathematics (Oct 01 – Oct 06, 2023)
    - Hausdorff Trimester "Prospects of formal mathematics" (May 06, 2024 - August 16, 2024)
    - National Academies AI To Assist Mathematical Reasoning Workshop (June 12th - June 14th, 2023)
        - Followup webinar
        - Second followup webinar (on formal proofs)
    - CICM

    - ICML 2024 AI for Math Workshop (Jul 26, 2024)
    -
- **Formal Proof and Constraint Solving**
    - CPP
    - ITP
    - CAV
    - Lean-related conferences and events
    - Journals
        - Journal of Automated Reasoning
        - Formalized Mathematics
        - Journal of Formalized Reasoning
        - Annals of Formalized Mathematics
- **Programming Languages and Software Engineering**
    - PLDI
    - POPL
    - ICFP
    - SPLASH/OOPSLA
    - TOPLAS
    - ICSE
    - ESEC/FSE
    - OPLSS (Summer School)
- **Artificial Intelligence and Machine Learning**
    - NeurIPS
    - ICML
    - ICLR
    - TMLR
    - JMLR
    - AAAI
    - JAIR
    - IJCAI
    - IEEE Transactions on Pattern Analysis and Machine Intelligence
- **Mathematics**
    - Joint Mathematics Meetings (JMM) [every year in January]
    - MathSciNet, zbMATH
    - SIAM meetings
    - Mathematics Institutes, including those supported by NSF DMS, (Workshops, Schools, Long Programs)
        - SLMATH (formerly MSRI) (https://www.slmath.org)
        - ICERM (https://icerm.brown.edu/)
        - BIRS (https://www.birs.ca)
        - IPAM (https://www.ipam.ucla.edu)
        - AIM (aimath.org)
        - IMSI (https://www.imsi.institute/)
        - Fields Institute (http://www.fields.utoronto.ca/)

- IAS (https://www.ias.edu)
- MFO (https://www.mfo.de/)
- MPIM (https://www.mpim-bonn.mpg.de/)
- CMI (https://www.claymath.org/)
- CRM (https://www.crm.umontreal.ca/en/)
- Newton Institute (https://www.newton.ac.uk/)

- **General**
  - Nature
  - PNAS
  - Science

Incentive Structures

Some incentive structures that only some of these fields do, but that are very useful for this kind of work, especially when it comes to large collaborations, building practical tools, and formalizing proofs. Maybe these will be useful to other communities.

- Artifact evaluation committees
  - It is common to post artifacts on Zenodo
- ICSE tool track
- CRA Hiring & Tenure Guidelines
- Experience Reports at ICFP
- Proof Pearls at ITP
- The AI Mathematical Olympiad (AIMO) challenge
  - Progress Prize 1
- NSF Funding Opportunity: Artificial Intelligence, Formal Methods, and Mathematical Reasoning (AIMing)
- NAIRR Pilot Resource Requests to Advance AI Research