

“Lili”

Liberated Library Online Platform

Contents

Mission.....	3
Justification and Objectives.....	3
Developing and Distributing Literature.....	4
Producing E-Books and Printed Copy.....	8
Teaching To Improve Literacy.....	9
Children.....	9
Parents.....	10
Content Creators and AI.....	11
A Note on Copyright Free.....	12
Reconfigure Global Literacy	12
Solving Token Problem and Data Ownership.....	13
Summary.....	14

Mission

The Liberated Library, known as “Lili”, is an education technology nonprofit that uses an innovative approach to improve the rate of literacy. One of the keys to improved literacy is making literature more accessible to children across all strata of society.¹ Lili aims to accomplish its mission through two primary focuses:

- Developing and distributing “copyright-free” literature (e.g., books, e-books, video) using micro-volunteerism supported by artificial intelligence. We refer to this as the “product” branch.
- Educating children by teaching research-based reading strategies to parents/caregivers and providing children access to digital books that model research-based strategies. We refer to this as the “education” branch.

Lili is a digital platform in which the product and the education exist in a symbiotic relationship to achieve the primary goal of improving child and adult literacy.

Justification and Objectives

Literacy in both children and adults in the United States is markedly low. *Sixty-three percent* of graduating high school students are non-proficient readers², and *forty-three percent* of adults are either illiterate or face significant challenges when engaging with literary text.³ Illiteracy disproportionately impacts minorities and low socioeconomic individuals and families.^{4 5} These deficiencies cause a host of downstream issues that are individual, societal, and perpetual.

At Lili, we believe that by teaching parents and caregivers research-based, at-home reading strategies and providing free reading material, we can accomplish our primary objectives of improving child and, subsequently, adult literacy. By improving child literacy, we will inevitably, over time, improve adult literacy.

Developing an education platform that provides equal access and equal voice to all people regardless of race, gender, background, and socioeconomic status is not just a fundamental objective of Lili but also is foundational for the mass development and distribution of free books. Because minorities and low socioeconomic individuals make up a significant portion of those low in literacy, developing an inclusive and free-to-use program will promote engagement within

¹ <https://www.sciencedirect.com/science/article/abs/pii/S0049089X18300607>

² <https://eric.ed.gov/?id=ED542022>

³ https://nces.ed.gov/naal/kf_demographics.asp

⁴ <https://www.apa.org/pi/sep/resources/publications/minorities>

⁵ <https://www.oecd.org/skills/piaac/Country%20note%20-%20United%20States.pdf>

these demographics, thereby achieving our goal of literacy improvement across all socioeconomic strata.

To learn more about literacy rates in the U.S., cross-generational education, or how socioeconomic status impacts literacy, please read our in-depth review of these topics in our academic synthesis, *Why It's Necessary*.

Developing and Distributing Literature

At its core, Lili is a moderated online platform that brings together individuals in a collaborative capacity who are willing to invest their skills and time to produce “copyright-free” literature. Lili’s user framework is designed around three primary roles:

- **Content Creators** include authors who write stories and artists who illustrate stories.
- **Content Editors** who edit and approve submitted text and image data.
- **Content Consumers** interact with six types of content: plain text, image samples, final books, final videos, educational text, and educational videos.

Lili uses crowdsourced text and image data uploaded from its community of users. Authors and artists (content creators) provide the data hosted on Lili’s platform and, if desired, are supported by advanced AI tools. Content creators and editors operate in the context of micro-volunteerism. There are no restrictions on who can provide content; instead, we believe that a diversity of users improves the quality of the uploaded data, resulting in a spectrum of engaging content. It is important to note that we believe a potential user base in the millions is possible. Justifications for this include user metrics from analogous platforms like:

WRITING

- Goodreads - 90m users⁶
- Reddit r/writingprompts - 13.7m subscribers⁷
- Facebook CreativeWriting Group - 247k writers⁸
- Instagram writing.prompt.s 780k followers⁹

IMAGE

- Deviantart - 61m users¹⁰
- Reddit r/painting - 1.1m members¹¹
- Reddit r/digitalpainting 1.1m members¹²
- Reddit r/illustration 777k members¹³
- Pixiv - 50m users¹⁴

⁶ <https://www.statista.com/statistics/252986/number-of-registered-members-on-goodreadscom/>

⁷ <https://www.reddit.com/r/WritingPrompts/>

⁸ <https://www.facebook.com/groups/whercreativitylives>

⁹ <https://www.instagram.com/writing.prompt.s/?hl=en>

¹⁰ <https://www.deviantart.com/about/#:~:text=We%20have%20over%2061%20million,art%2C%20films%2C%20and%20anime.>

¹¹ <https://www.reddit.com/r/illustration/>

¹² <https://www.reddit.com/r/DigitalPainting/>

¹³ <https://www.reddit.com/r/illustration/>

¹⁴ <https://en.wikipedia.org/wiki/Pixiv>

The production process works as follows:

1. Authors write stories on Lili or upload them to the platform where the works are screened, utilizing automated software that identifies key terms (swear words, sexually explicit language, etc.). AI language software is integrated into this system as an optional support tool, broadening the potential content creator base.
2. The uploaded pre-screened work is then made available to volunteer editors (through our partnerships with Universities and colleges) who read and approve the written work.
3. The approved document or file is uploaded to Lili's database and displayed for general user/community interaction.
4. Content consumers and artists can access and engage with text-based stories.
5. The stories are subjected to community approval (up/down voting) and artists can select stories that have high user engagement or that effectuate personal interest.
6. Artists then collaborate with authors and editors to produce the final product, a "copyright-free" book. Image generation AI software is available for either authors or artists, thus broadening the scope of potential participants and alleviating possible image content production limitations.
7. The new book is hosted on Lili's platform (e-book, video, and audio), other hosting sites, and in some cases, produced into traditional print copies to be distributed through our partnerships.

This process includes an initial automated digital screening and a review procedure so the content remains age-appropriate. All literary works can be searched and accessed by children's age (i.e., 0 to 18 years old) and type of literature (e.g., fiction, nonfiction, prose, fantasy, sci-fi, etc.). Most authors are expected to write material for pre-k and elementary-age children (i.e., 4- to 11-year-olds).

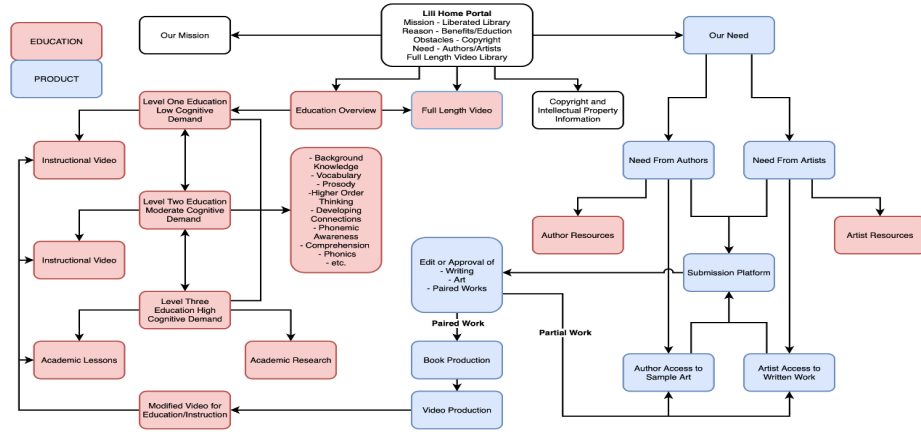
Once the text data is online, all users can vote on the perceived value or quality of the content, similar to an upvote/downvote system used by many social media platforms. A significant benefit of community engagement is the leveling of the authoring playing field. In this new paradigm, conventions like grammar may not be prioritized over the social contextual

value. Authors without formal training will be able to share meaningful stories not moderated by corporate publishing interests, thereby providing a platform where all voices (regardless of education, background, or socioeconomic status) can be heard.

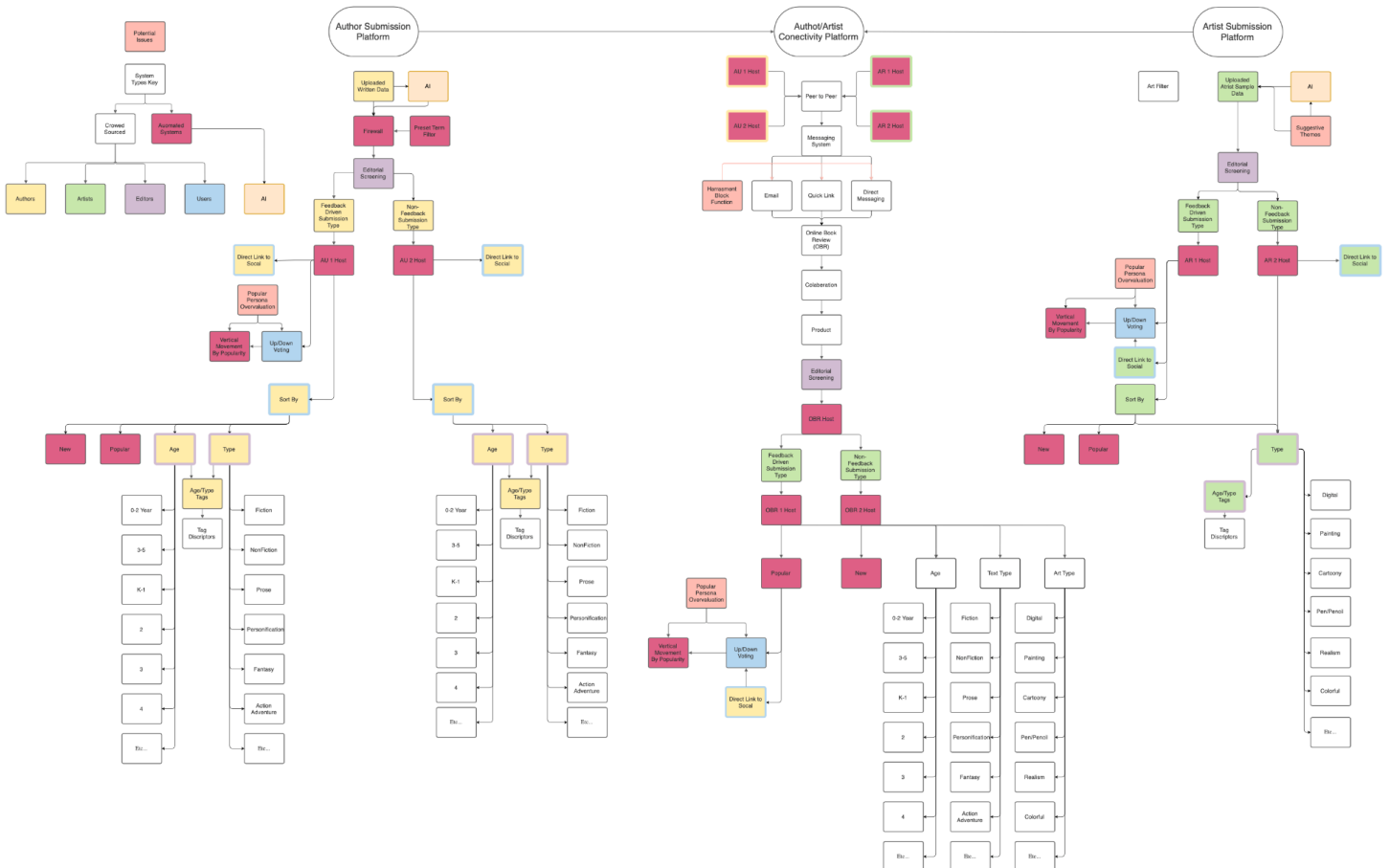
Additionally, the use of artificial intelligence in both the writing and image production of books allows for a reduction in the level of skill necessary to participate in the development of literary works. AI can lower the barrier to entry, providing an opportunity for untrained users to produce meaningful stories without having pre-developed writing or illustrating schemas. By integrating artificial intelligence into the production of literature, Lili is broadening the creator base of potential authors by reducing limitations.

The process for submitting image data is similar to that of text data except for minor differences, such as automated screening and a separate system for categorizing submitted works (i.e., digital, pencil, cartoon, etc.). Artists upload their content to the platform, where it is stored in each artist's portfolio as either free available work to be used by authors or as samples of artistic style. Artists and authors are encouraged to engage with one another to develop literary works together, and on-site communication tools (e.g., email and direct messaging) are available to help them collaborate on projects.

A generalized framework (both product and education) of Lili can be seen in Fig. 1.



Lili's workflow can be seen in Fig. 2.



Producing E-books and Printed Copy

In the initial stages of development, volunteer editors and staff compile the text and image data into completed works using third-party software. Moving forward, to scale this process, new software will be designed, purchased, or leased and built into the platform's user interface to accommodate authors and artists. Eventually, this will allow users to create literary works without relying on editors to complete the process. Hosting the software needed to format books on Lili's platform will streamline the development process and allow the platform to scale production.

Once the uploaded data moves through the initial digital production process (written work paired with art, edited, and compiled), the new "copyright-free" book will be published in a readable online format. Partnering with a software company such as Flippingbook, which provides e-book readers, will eliminate the cost of developing software separately and allow us to integrate third-party APIs directly into the ecosystem. Producing digital copies of the finished texts is inexpensive and provides access to the written works at a very low cost. The digital e-book files will be downloadable in multiple formats free of charge.

Due to the vast quantities of uploaded data, hard copies of the texts will require a more limited production approach to mitigate costs. Lili relies on user metrics to identify texts that have high community approval. These high-approval works are selected for limited print runs, video production, educational instruction, and promotional activities. Partnerships in producing and distributing highly rated hard copy texts could include Dolly Parton's Imagination Library or the Scholastic Book Program. It is important to note that while these partnerships are still in development, we feel the benefits that Lili offers these organizations provide substantial incentives. Additionally, users (e.g., parents, caregivers, children, authors, artists, etc.) can select any digital work by request and pay for a copy through our potential partnering publisher (e.g., Scholastic, Blurb).

A significant component of Lili's production branch is developing and distributing free educational video content. Books that receive high user ratings are created into hard copy format, read, recorded, and uploaded to digital content hosts (e.g., Lili, YouTube, Odyssey). Eventually, an extensive catalog of videos modeling the research-based strategies contained in Lili's education branch will be available. We will seek partnerships with various readers, including well-known personalities such as LaVar Burton who recently expressed the need for a system like Lili.



Once the full-length videos have been produced and uploaded to the content host(s), they will be copied and modified (i.e., shortened, dubbed) to be used as instructional material associated with Lili’s education branch. Each video will provide visual examples of a literacy component within the education framework and model researched-based instructional methods (i.e., strategies for teaching phonics, vocabulary, prosody, phonemic awareness, etc.).

Teaching to Improve Literacy

Education is fundamental to Lili’s mission of improving literacy. Within the education branch, we have three primary objectives:

- Teach children literacy.
- Teach parents and caregivers how to teach children literacy.
- Teach authors, artists, and editors the essential skills necessary to produce literature for children.

All three of these objectives are deeply interconnected. Authors, artists, and editors create the literature given to parents and children. Parents learn how to best use the literature to teach their children, and Lili will be the stand-in for children who do not have anyone to teach them.

Children

To improve childhood literacy rates, we provide children's books for free to all children. Lili’s books can be accessed for free through the platform using an onsite digital e-book reader. Additionally, books that generate high user approval will be produced as hard copies. These books will be read, video/audio recorded, and uploaded to video-hosting and audiobook-hosting sites. In some cases, these books will be printed in quantity and distributed using third-party partnerships.

Eventually, an extensive video catalog of children's literature will be available and categorized by literature types and age groups. Video content/readings will follow and model research-based shared book reading strategies—the same strategies taught to parents through Lili’s education branch. Examples of lessons in the videos will include using cues such as: “When reading, follow the words with your finger to teach the left-to-right convention” or “When encountering an unfamiliar word, stop and explain the meaning of the word.” Over time, Lili will partner with influential philanthropic individuals (e.g. John Cena, LaVar Burton, Danny Trajio) to create and read their own books or read user-created literature—all while following and teaching shared book-reading strategies. Potential revenue streams are associated with this type of content creation.

Additionally, by partnering with philanthropic organizations such as the Barbara Bush Foundation and Dolly Parton's Imagination Library, we can provide children with hard copies of our highest-rated books. By relying on volunteer authors, illustrators, and editors and eliminating traditional publisher fees, the production cost of these texts will be substantially lower than traditionally manufactured literature. We expect this to motivate philanthropic partnership development because significantly more books can be produced and distributed for the same cost.

Parents and Caregivers

Using research and scaffolded instruction, we teach parents sixteen fundamental components of literacy (i.e., vocabulary, prosody, comprehension, fluency, phonemic awareness, etc.). Each of the sixteen components is scaffolded into three groups of descending rigor, starting with an academic synthesis encompassing the component definition/characterization, correlation/relevance, and research-based instructional methods. For example, an explanation of vocabulary as it relates to literacy will be included in the vocabulary component, along with academic research that clearly outlines how vocabulary directly correlates to overall literacy performance in children. This information is associated with research that shows the best methods to teach vocabulary while reading to a child. These academic overviews are the foundation of the education branch for both adults and children.

While the structure of the parent education program is built in descending rigor, allowing for homogeneity of the content, parent instruction is ordered in ascending rigor. Research shows that parents of low education or socioeconomic status are less likely to engage with their children in literacy-building activities¹⁵; therefore, this educational gap is a major target of Lili's education strategy. Starting with low cognitive demand instruction, parents who may have low literacy levels are engaged via simple text outlines and video content that model the instructional strategies derived from the academic outlines. It is important to note that all instructional videos will be produced by pulling examples from the video catalog used to engage with children. Considering that this demographic may have little to no formal education, the cognitive demand and overall rigor must be low, with a heavy focus on the use of instructional videos.

Increasing in rigor, parents can easily navigate to more in-depth information relating to each of the sixteen literacy components if desired. Targeting parents with basic high school or secondary degrees, these component outlines will be reasonably accessible regarding cognitive demand and will expand on the sixteen literacy components. Again, parents will engage with the literacy component of their choosing in both text and video formats.

¹⁵ https://www.aft.org/ae/spring2003/hart_risley

Further increasing in rigor, parents can access foundational outlines (academic synthesis) of the literacy components that are more academic. Each part is designed using a significant amount of research focused on the correlation between literacy development and the application of research-based strategies, specifically for shared book readings. To accomplish our mission of promoting literacy for children, the most up-to-date and relevant literature is used in the development of the education framework.

Content Creators and Artificial Intelligence

Although encouraging creativity and providing a platform on which people can write in nontraditional ways is a major objective, it is also important to provide aspiring writers and artists with the resources needed to learn traditional writing conventions. Critical information about effective writing and art strategies used in literature will be available for interested content creators. These instructional overviews will address traditional writing conventions.

As for artists, a diverse portfolio of children's book art styles will be provided. Other useful information, such as how to develop wordless books, will be included in the education material for content creators. In time, we anticipate scaling up to include short, informative write-ups by well-known illustrators in this framework.

The use of artificial intelligence (AI) as scaffolded support can play a significant role in the development of literary works. At Lili, we intend to use AI to support both writers and artists, thereby lowering the barrier to entry. We recognize that writing is a developed skill that only some users may possess and while a user may not have the necessary skills to write while holding to traditional conventions, they may have an important story to tell. AI can be used as a support tool to provide a scaffolded writing experience.

AI language models can assist individuals in writing stories by providing a range of supports, not just the production of text. These include but are not limited to:

1. Brainstorming: AI can help individuals generate ideas for stories by asking questions and suggesting prompts.
2. Plot development: AI can help individuals develop their story's plot by providing suggestions for how to advance the plot, create tension, and build conflict.
3. Character development: AI can help individuals flesh out their characters by suggesting personality traits, motivations, and backgrounds.
4. Writing techniques: AI can provide tips and techniques for improving writing style, such as how to create descriptive settings, write dialogue, and incorporate sensory details.
5. Feedback: AI can offer constructive criticism on elements such as pacing, characterization, and dialogue to help individuals refine their stories.

6. Editing: AI can assist in the editing process by correcting grammar, spelling, and punctuation errors, and suggesting ways to improve sentence structure and flow.
7. Research: AI can provide information on various topics to help individuals add depth and authenticity to their stories.¹⁶

A Note On Copyright-Free

While copyright-free is the best way to conceptualize the intellectual property (IP) designation of the books Lili’s authors and artists create, certain issues present themselves under this licensing format—specifically, the problem of third-party monetization. Technically, “copyright-free” simply means that a work is in the public domain and can be used by the public (including corporate publishers) for any reason, including monetization.

At Lili, we are determined to create books that start free and stay free, forever. This is why we use Copyleft and a modified GPL licensing system rather than the public domain. Using the GPL licensing system, each element of intellectual property created on Lili is granted protections. The license allows for the free use, modification, and reproduction of the IP; it also requires that any additional use of the intellectual property must remain free in perpetuity. This means that the IP can be used in any way imaginable, except monetization, thereby perpetuating open access to high-quality literature for all future generations.

Reconfiguring Global Literacy

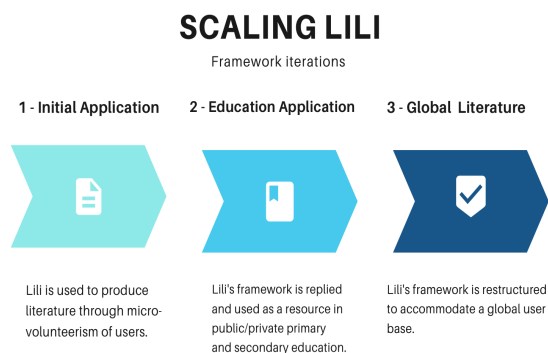
At Lili, we are a forward-looking company. The framework we have designed to improve child (and subsequently adult) literacy will be extracted and reimplemented into larger applications. Our roadmap for the future use cases for the Lili framework includes modifying the software into public and private education programs (both primary and secondary) and further developing into a global literature production and repository resource.

As Lili is scaled, our framework allows for the reapplication into multiple sectors. This reapplication process opens several significant data conduits. In Lili's initial iteration, we can significantly improve child and adult literacy through the mass proliferation and distribution of free literature. Our use of LLMs and other generative pre-trained transformers will facilitate this process by reducing the overall requirement of human input and the reduction of the required base knowledge of users.

In the second iteration of Lili, our framework will be directly integrated into the public school system allowing for the moderated connectivity of classrooms worldwide. In this format,

¹⁶ ChatGPT, <https://chat.openai.com/chat>, 2023

students will collaborate on literary works while supported by AI. For primary education, we see a future where a high school ELL language arts classroom in San Francisco actively collaborates with an art class in New Your City to produce books that students in Dallas read. The integration of AI educational supports into the software will afford students opportunities to write and express themselves in a way that, up until now, has not yet been possible.



Cross-community interaction on Lili will be further scaled into a global student workshop and supported by machine translation advancements. This collaborative academic format will be applied to higher education models where University students work together on the development of literary works, including academic papers as scientific research.

As the final iteration, Lili will be scaled to offer a place for the worldwide community of authors to collaborate, produce and review literary works together. The future is a place where publishing companies and academic journals are no longer the arbiters of literature, but rather, the library of humanity is liberated, and the global community or readers determines the value of written works and has free access to *all* literary works.

Token Problem and Data Ownership

Two significant barriers to the continued growth of Large Language Models (LLM) are the future availability of high-quality training tokens and the associated question of data ownership.¹⁷ At scale, Lili provides a solution to both of these issues. Data inputs into Lili offer LLMs access to new large training data sets, and as our system scales, these data sets will grow in both size and quality. This mutually beneficial relationship could be critical for the future development of LLMs as some evidence shows that data, not parameter count, will be the primary constraint on further development.¹⁸ Considering the current stock of high-quality training tokens is roughly $4.6e12$ and $1.7e13$ words, and potential exhaustion of the supply is likely, additional expansion of the data pool is necessary.¹⁹ Lili can provide the needed training tokens, especially considering the token type produced on Lili's platform aligns with definitions of high-quality data.

¹⁷ <https://arxiv.org/pdf/2211.04325.pdf>

¹⁸ <https://arxiv.org/pdf/2203.15556.pdf>

¹⁹ <https://arxiv.org/pdf/2211.04325.pdf>

Further, intellectual property rights associated with other data sets do not apply to the data generated on Lili. As mentioned, because our primary objective is to produce literature that stays free-to-use in perpetuity, we use a modified GPL licensing system. Using this licensing system comes with the supplemental benefit that any LLM using the tokens generated on the platform would not be subject to traditional intellectual property constraints. Lili is creating a new paradigm in which data becomes open access and is no longer controlled by central authorities.

Summary

Lili's highest priority is the development of literacy in children. This is accomplished with two strategies. First, our global user volunteer base produces copyright-free children's books at a very low cost. High-quality literature is produced via the integration of AI into the Lili ecosystem. These books are available in multiple formats, including e-book, video, audio, and traditional hard copy. Second, parents and caregivers are taught researched-based book reading strategies that develop literacy. Additionally, Lili provides an array of educational supports, including artificial intelligence for volunteers in developing books to directly impact the quality of literature provided to the children. Future iterations of Lili's framework will impact both primary and secondary education as well as eventually offer global access to significant amounts of free high-quality literature.