

# AIMS Philanthropy Project:

Studying AI, Machine Learning & Data Science Technology for Good

Data has a better idea



Image Source: Franki Chamaki on Unsplash <https://unsplash.com/photos/1K6lQsQbiz1>



**Patricia Snell Herzog, PhD**

with Harshal Naik and Haseeb Khan

February 2021

## **1. Introduction**

This project investigates philanthropic activities related to Artificial Intelligence, Machine Learning, and Data Science technology (AIMS). Advances in AIMS technology are impacting the field of philanthropy in three primary ways:

1. Substantial amounts of philanthropic resources are being invested in the development, advancement, and application of AIMS technology.
2. AIMS technology is changing philanthropic mission delivery and process through automation, data mining, predictive analytics, and impact insights.
3. Workforce needs for AIMS skills are altering the future of work in the field of philanthropy by requiring a nexus of domain, data, and technical expertise.

This AIMS Philanthropy project addresses these changes through three long-term aims: analyze and visualize philanthropic funding for AIMS technology; partner with philanthropic practitioners to innovate ways that AIMS technology can improve philanthropy by harnessing its power and limiting its risks; and address the changing needs of workforce preparation by developing modernized AIMS-informed curriculum and programming that joins domain knowledge of philanthropy with data science.

This report focuses on methods employed in targeting the first of these aims. The goal was to develop an accessible website platform that engaged human-centered UX user experience design techniques to present information about AIMS Philanthropy. The website experience was designed toward the following goal, for the following set of intended users, and with the following data analysis methods.

### **A. UX-Goals**

In order to provide website users with an experience designed to identify and describe AIMS philanthropy, the goal of the project is to visualize data on the size, scope, organizations, and projects funded by AIMS philanthropy. First, existing data sources were scraped. Second, amassed data were wrangled and analyzed. This included coding funds that were specifically directed at AIMS technology and its applications. Third, visuals were created to illustrate the amount, purposes, and recipients grant awards and other philanthropic investments.

### **B. UX-Users**

The website and interactive data dashboards were designed with three specific and one broad set of users. The primary intended users are practitioners in nonprofit and philanthropic organizations. This user set includes workers in grantmaking foundations who are currently funding or considering investing in AIMS technology and its applications. The primary user set also includes workers, volunteers, and board members in nonprofit organizations who are currently or considering implementing AIMS technology to enhance mission delivery, organizational effectiveness, or program social impacts.

A second intended user set includes students. Students in data science disciplines, social sciences, and philanthropic studies can investigate the data presented here to further develop their research agendas and consider data visualization techniques.

Another intended user set focuses on researchers who are interested in methods for creating and analyzing data, within or at the intersection of social sciences, informatics, and philanthropy. The details of the data methodology below are especially of benefit to this set of users.

A fourth general user is the public, communities, workers in media organizations, and everyday people who are curious about the potential of AIMS technology and its impacts, as well as those who are skeptical or who have ethical concerns about the large amounts of money and organizational resources invested in these activities.

### **C. UX-Process**

To create a user design experience website, with stimulating data experiences, and intentionally designed for multiple user sets, the project focused on these priorities. The first priority was to collect and synthesize information about AIMS funding. Second, a priority was anticipating user curiosities in relation to AIMS funding, visualizing data to respond to questions, and providing users with initial and stimulating answers. The third priority was engaging specifics of AIMS philanthropy as a way to foster a broader understanding of what philanthropy is generally, including the size and scope of the AIMS subsector. Fourth, bridging gaps between technical and domain knowledge was a priority in order to aid users in better understanding AIMS impacts, by visualizing AIMS and philanthropy intersections.

## **2. Methods & Preliminary Analysis**

The goal of this project was to identify and visualize the grant awards that involve Artificial Intelligence, Machine Learning, or Data Science technology (AIMS). Each dataset was analyzed for a set of general questions that could be answered visually. The visuals aim to provide answers to these two primary questions: (1) How much funding was invested in AIMS? and (2) What focus areas, applications, discovery, or other purposes was AIMS-funded directed toward? To answer these questions, a total of five datasets were investigated: two structured and open-source grant databases, one structured paywalled nonprofit database, and two unstructured social media text-based analyses. Each dataset was cleaned and parsed for AIMS-related data. Specific methods for parsing and coding each dataset are described below.

### **A. Open Philanthropy**

The data for this first analysis are from an open-source grants database on the [Open Philanthropy website](#). Open Philanthropy is an organization founded by Cari Tuna and Dustin Moskovitz, who was a co-founder of Facebook and Asana. This project originated from [GiveWell](#), a nonprofit organization founded by two former hedge fund managers: Elie Hassenfeld and Holden Karnofsky. An offshoot program, [GiveWell Labs](#) was formed and then rebranded as the Open Philanthropy Project (OPP).

OPP is composed of several affiliated organizations, including a 501(c)(3) called Open Philanthropy Project, a philanthropic foundation called [Good Ventures](#), a 501(c)(4) called [Open Philanthropy Action Fund](#). Grants are typically made through a donor-advised fund at the [Silicon Valley Community Foundation](#). OPP provides an open-source database of all grants awarded since 2012, totaling to more than 850.

## Preliminary Analysis

The Open Philanthropy [grants database](#) (n=879) specified 13 grantmaking focus areas: Biosecurity and Pandemic Preparedness; Criminal Justice Reform; Farm Animal Welfare; Global Catastrophic Risks; Global Health & Development; History of Philanthropy; Immigration Policy; Land Use Reform; Macroeconomic Stabilization Policy; U.S. Policy; Scientific Research; Potential Risks from Advanced Artificial Intelligence; and Other Areas. Figure 1 visualizes grant amounts for each focus area.

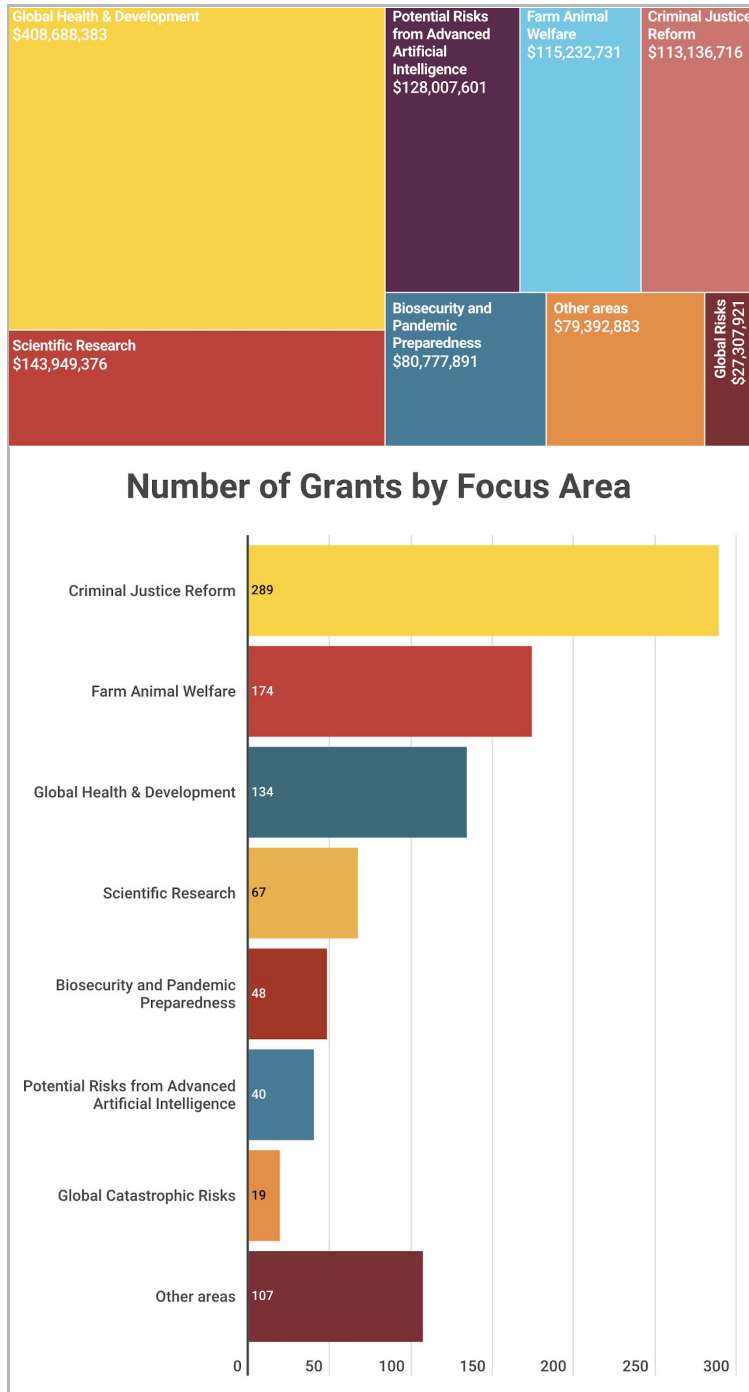


Figure 1. Grants by Focus Area (Top-Amounts; Bottom-Number of Grant Awards)

The most evident AIMS-related focus area for Open Philanthropy is: Potential Risks from Advanced Artificial Intelligence, which was awarded \$128.01 million in grants as the 3rd-highest-funded focus area receiving 12% of Open Philanthropy grants. From

these preliminary analyses, it is evident that AIMS topics are a priority for Open Philanthropy. Additionally, several additional analyses were undertaken to further parse funding initiatives that are related to AIMS technology.

### **Parsing In-Depth Analysis**

While the grants database was structured, the information relevant for parsing AIMS activity was unstructured and needed to be coded for analysis. Toward developing a relatively automated process that was not as time-intensive as individually searching each grant manually, a keyword parsing process was employed based on the fields containing names of organizations and grant project descriptions. If these fields contained any terminology related to AIMS, they were scraped for further analysis.

For example, if a grant description contained keywords such as artificial intelligence, machine learning, or data science, the grant data was scraped for further analysis. Advancing beyond only the focus area, this more in-depth analysis revealed that many additional grants involved AIMS-related efforts. These grant activities were categorized within additional focus areas and would have been missed had an automation algorithm scraped data only from the focus area field.

Moreover, there were several grants that could not be identified from the database fields alone. For these, additional investigations of organization websites resulted in manually coding several additional projects based on social media descriptions. Figure 2 displays the results of this parsed and in-depth analysis for AIMS projects.

After further parsing, the in-depth analysis revealed that an additional \$13.85 million in AIMS-related funding was awarded within the focus area coded Global Catastrophic Risk, \$5.97 million in Scientific Research, \$4.55 million in Biosecurity & Pandemic Preparedness, \$2.85 in Land Use Reform, \$0.77 million in Farm Animal Welfare, and \$10.87 million in Other Areas.

Of these additional AIMS-related activities funded in other grant focus areas, 36% were for applications of AIMS technology in Global Catastrophic Risks applications, along with an additional 12% in Biosecurity and Preparedness and Global Catastrophic Risks, 7% for Land Use Reform applications, and 28% were for applications of AIMS technology in other focus areas.

In summary, a preliminary analysis of focus areas identified \$128.01 million dedicated to AI activities, and the in-depth analysis revealed an additional \$38.96 million in Open Philanthropy funding for AIMS-related discovery and applications. ***Cumulatively, Open Philanthropy invested \$166.97 million in AIMS since 2012.***



Figure 2. Additional AIMS-Related Grants (Top-Amts; Middle-#.; Bottom-Total)

## B. Rockefeller Foundation

The data for this second analysis are from an open-source grants database on the [Rockefeller Foundation website](#). The Rockefeller Foundation was founded in 1910 by John D. Rockefeller, co-founder of Standard Oil. The foundation began its work

by funding the [National Research Council](#), since renamed the National Academies of Sciences, Engineering, and Medicine. This attention to science continues in the foundation, and in 2020 the Rockefeller Foundation partnered with [The Mastercard Center for Inclusive Growth](#) to launch [Data.org](#) to fund data science for social impact. The Rockefeller Foundation provides an open-source grants database of all the grants awarded since 2014, totaling to more than 600 awarded grants.

### **Preliminary Analysis**

The Rockefeller Foundation provides an open-source [grants database](#) of all the grants that their organization awarded since 2014 (n=674). The foundation awards a wide range of [initiatives](#), and the following set of 10 focus areas combine the grants database categories with the website initiative objective titles: Climate and Resilience (End Energy Poverty); Health (Achieve Health for All); Food (Nourish People and Planet); Jobs (Expand Equity and Economic Opportunity); Emerging Frontiers (Seize upon Emerging Frontiers); Power (Solving Global Challenges with Lasting Impact); Co-Impact; Innovation; Innovative Finance; and Data & Technology. Figure 3 visualizes grant amounts by focus area, highlighting Data & Technology.

The most evident AIMS-related focus area for Rockefeller Foundation is: Data & Technology, which was awarded \$20.28 million in grants as the 9th-highest focus area receiving 4% of the total grant awards. In less than a decade, Rockefeller Foundation made 10 grant awards to fund Data & Technology. From these preliminary analyses, it is evident that AIMS topics are a priority for the foundation.

Indeed, the newly launched Rockefeller Foundation funded project, Data.org [reports](#) that it is “using the power of data analytics and insights to unlock solutions to society’s most pressing challenges.” The project continues by stating, “Data science is being used to bring products and services to markets that are profoundly affecting our world. Yet too many mission-driven organizations lack the know-how and tools to harness the power of data.” Identifying the problem as a dearth of data science tools in the field of philanthropy, the foundation positions Data.org as the solution: investing in use cases with high demonstration value, building capacity of nonprofits leaders and organizations to employ and learn from data science applications, and lessening pressing challenges by strengthening the public goods ecosystem.

Toward these ends, the foundation committed more than \$50 million dollars in data science investments over the next five years, including \$20 million to fund [DataKind](#) as a global epicenter for data science and \$7.5 million to [Benefits Data Trust](#). These initiatives were launched in partnership with [Mastercard Center for Inclusive Growth](#).





Figure 3. Grants by Focus Area (Top--Amts.; Middle/Bottom-Prop.)

### Parsing In-Depth Analysis

Similar to the Open Philanthropy information, the grants database was structured, but the information relevant for parsing AIMS activity was structured and needed to be coded for analysis. Adding to the [Open Philanthropy parsing details](#) above, this analysis focused on relevant keywords, specifically each combination of these words: artificial, intelligence, machine, learning, data, technology, analysis, analytics.



Boolean logic was employed to conduct *and* searches (such as data and technology). Alternatively, *or* searches were conducted to identify either instance of the keywords, in which case more manual parsing is needed afterwards to assess whether the selection is indeed a hit. The results of each Boolean logic set of searches and keyword color schemes are included in Table 1.

*Table 1. Results for Each Boolean Logic Keyword Search Terms and Color Scheme*

Keyword 1	Keyword	Boolean Logic	Number of Identified Grants
Artificial	Intelligence	<i>and</i>	2
Machine	Learning	<i>and</i>	3
Data	Technology	<i>and</i>	19
Artificial	Intelligence	<i>or</i>	5
Machine	Learning	<i>or</i>	20
Data *	Technology	<i>or</i>	99
Analysis **	Analytics	<i>or</i>	21

\*Data or Technology also identified combinations such as data analytics, data analysis.

\*\*To thoroughly parse relevance, additional searches employed the terms: analysis or analytics.

After further parsing, the in-depth analysis revealed that an additional \$36.48 million in AIMS-related funding was awarded within nearly all the other grant focus areas. The top focus areas for spending on AIMS-related activities included Food Initiatives for \$68.99 million, Co-Impact for \$66.00 million, Health Initiatives for \$55.39 million, Partnerships & Opportunities for \$47.79 million, Innovative Finance for \$42.17 million, Climate and Resilience for \$34.95 million, Equity & Economic Opportunity for \$25.71 million, Global Resilience Partnership for \$20.32 million, Legacy for \$19.97 million, and New Frontiers for \$16.69 million.

There are additional several applications areas that engage AIMS technology, including Digital Jobs Africa for \$8.30 million, Fresh Water for \$5.19 million, Inclusive Economies for \$2.16 million, and various other foundation initiatives. Figure 4 displays the results. ***The \$19.93 million for Data & Technology grants combines with another \$36.48 million in additional AIMS-related grant awards, resulting in a total of \$56.42 million in AIMS investments by the Rockefeller Foundation.***

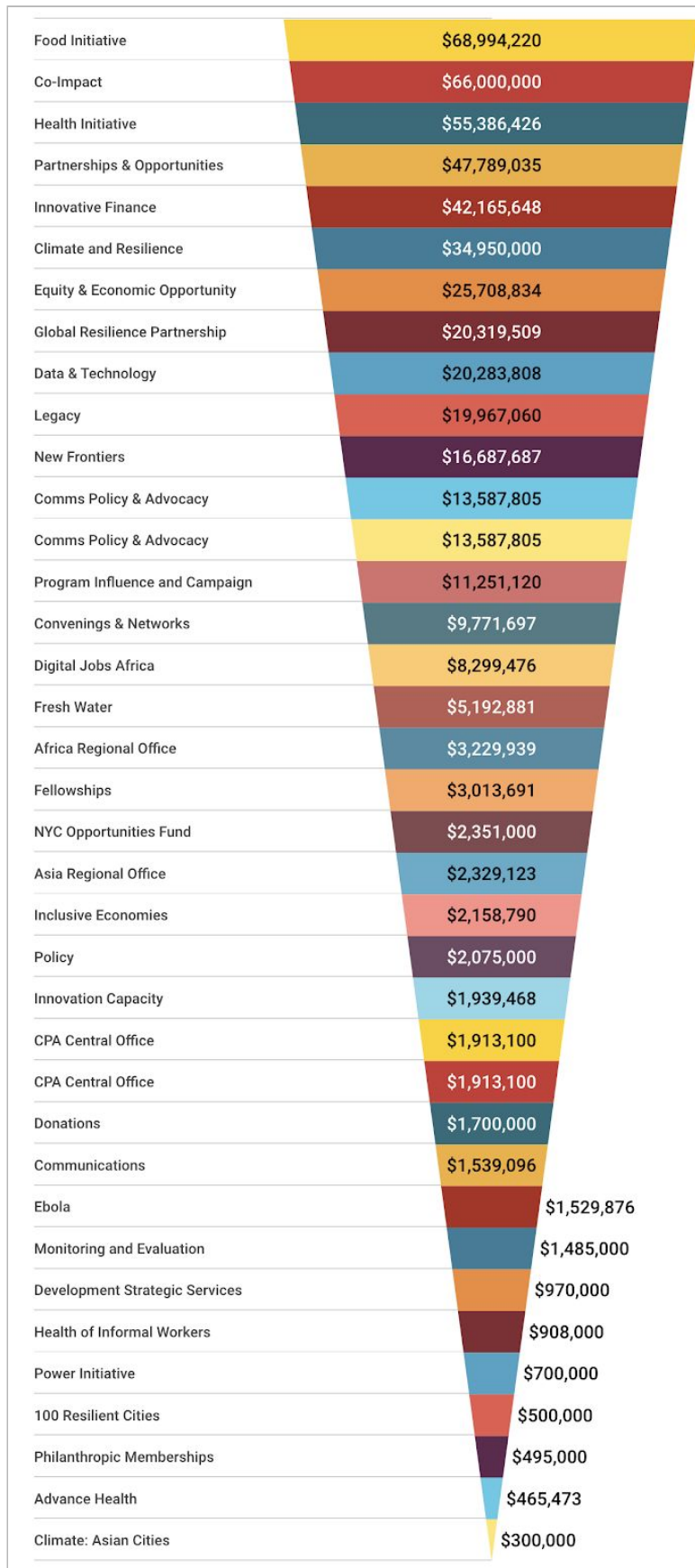


Figure 4. Additional AIMS-Related Rockefeller Foundation Grants

### C. Chronicle of Philanthropy

The data for this third analysis are from an unstructured and paywalled news media source, the [Chronicle of Philanthropy](#). The chronicle began as a magazine and now publishes content online. It is a subsidiary of the [Chronicle of Higher Education](#). In addition to publishing original research on the 400 largest nonprofit organizations, the Chronicle of Philanthropy also regularly reports on large donations from philanthropists. Their online article database of more than 15,000 articles facilitates keyword searches by term and was employed to search for AIMS-related donations.

#### Parsing In-Depth Analysis

Unlike the prior two sources, there was not a preliminary analysis performed. This section describes the parsing process to conduct an in-depth text analysis. Similar procedures were used for the keyword search and Boolean logic processes described above in the [Rockefeller Foundation section](#). By using these keywords, the news website displayed relevant articles pertaining to those keywords. The next step was to create a database in which information received from the article regarding the grant was filled in. The database contained information such as: who donated, for what purpose, to which organizations, and in what amounts. Table 2 displays the number of articles returned for each keyword and Boolean logic set.

*Table 2. Results for Each Set of Boolean Logic Keyword Search Terms*

<b>Keyword 1</b>	<b>Keyword 2</b>	<b>Boolean Logic</b>	<b>Number of Articles</b>
Artificial	Intelligence	<i>and</i>	146
Machine	Learning	<i>and</i>	107
Data	Science	<i>and</i>	428
Artificial	Intelligence	<i>or</i>	462
Machine	Learning	<i>or</i>	3646
Data	Science	<i>or</i>	4737
Analysis	Analytics	<i>or</i>	982

Relevant articles were then coded for the following primary focus areas: Artificial Intelligence, Machine Learning, and Data Science. Several articles reported on intersections of these topic areas, coded in the following sets: AI & Data Science; AI & ML; AI, ML, and Cybersecurity; AI and Supercomputing; Scientific Research, Advanced Technologies, and Scholarship Funds. For each article, the philanthropic contribution amount was recorded, along with the name of the donor, the organizational recipient (if listed), and a description of the contribution.

Figure 5 displays the results, sorted hierarchically with the greatest contribution amounts on the top of the figure, and the least contribution on the bottom. The results indicate that \$812. million has been contributed for AI, \$369 million for AI and Supercomputing, \$306 million for Machine Learning, \$115 million for Advanced Technologies, \$106 for AI, ML, and Cybersecurity, \$65 million for Data Science, \$50 million for AI and ML, \$34 million for AIMS-related scholarship, \$30 million for AIMS-applications in neuroscience, \$15 million for AI and Data Science, and \$10 million for Advanced Technologies.

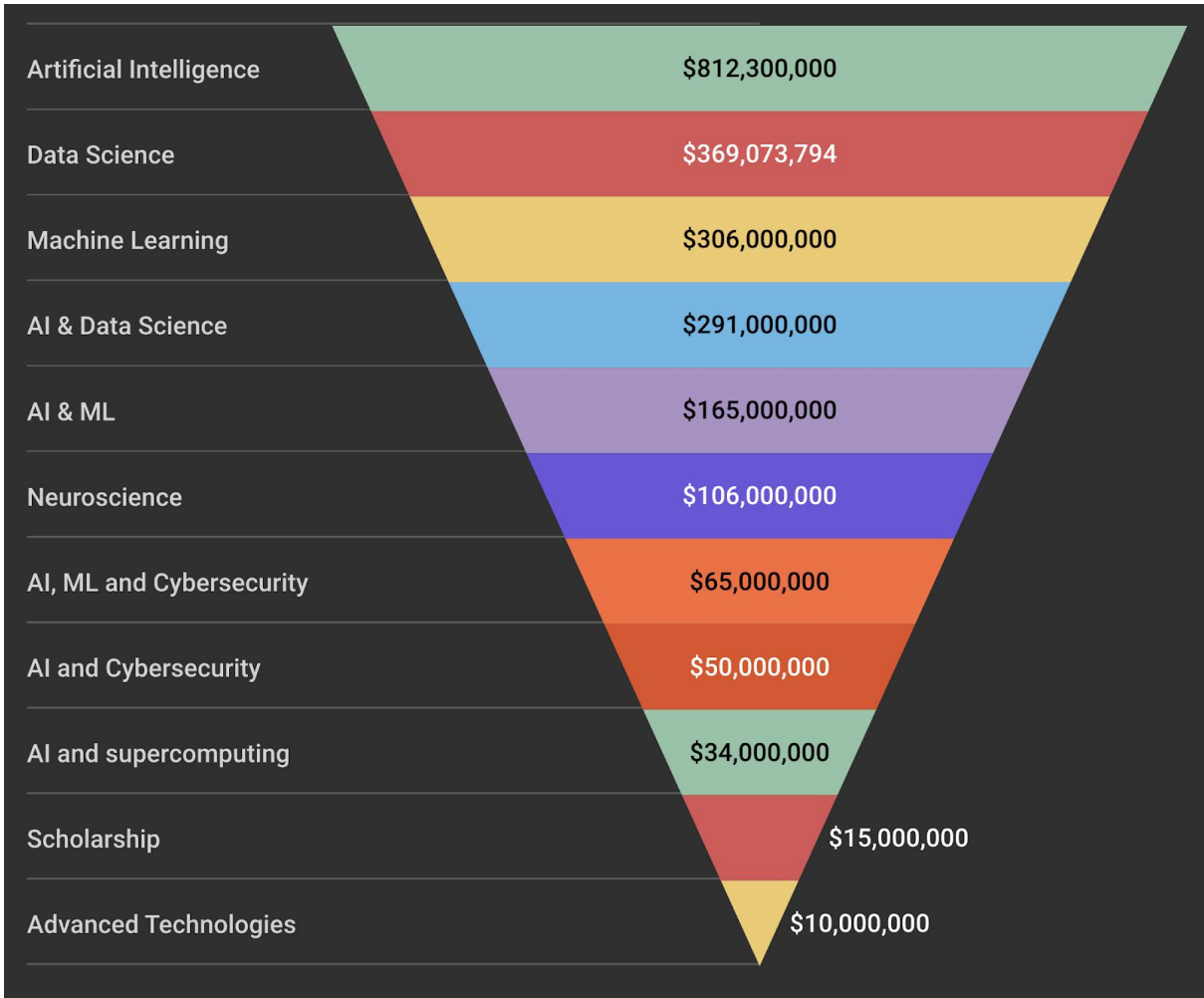


Figure 5. AIMS-Related Contributions Reported in the Chronicle of Philanthropy

Cumulatively, the Chronicle of Philanthropy reports that 47 philanthropists and philanthropic organizations have contributed to AIMS technology. The top ten contributors are Stephen Schwarzman (\$355 million), The Knight Foundation (\$300 million), Tom Lord (\$261 million), Paul Allen (\$125 million), Jaffray Woodruff (\$120 million), Sanford and Joan Weill (\$106 million), David and Barbara Roux (\$100 million), Amin and Julie Khoury (\$100 million), and Microsoft (\$93 million).

These funds were reported as received by 53 unique organizations, with the top ten recipients including MIT (\$350 million), Support Local Journalism (\$300 million), University of Southern California (\$261 million), the Allen Institute for Artificial Intelligence (\$125 million), University of Virginia (\$120 million), Weill Neurohub (\$106 million), Roux Institute (\$100 million), University of California at San Diego (\$75 million), and the Indiana University Luddy Center for Artificial Intelligence (\$60 million). **Cumulatively, this totals to \$2.22 billion in AIMS-related philanthropy.**

#### D. GuideStar Nonprofit Database

The data for this fourth analysis are from a structured paywalled nonprofit database on the [GuideStar website](#). GuideStar was formed in the mid-1990s as an organization designed to extract information from publicly available tax information

on organizations that are legally incorporated to be tax-exempt. These organizations are typically incorporated with a legal status designated as [501\(c\)3](#) and are required by the IRS to file [990 forms](#) that contain organizational data, including legal name, intended social purposes, and annual operating budget amounts. People began to realize that these public data could be leveraged to better understand the purposes, goals, and effectiveness of organizations that are partially subsidized by public funds. The tax breaks the organizations receive are intended to contribute toward their missions to provide social goods to communities beyond only the individuals engaged directly in the organization. Over time, donors, public policymakers, media, watchdog organizations, and the general public have come to use the GuideStar database to track which organizations to support for desired social impacts.

In [2019](#), GuideStar merged with the [Foundation Center](#) to form a new organization called [Candid](#). The Foundation Center emerged in the mid-1950s for similar purposes: to promote transparency in the charitable process. The Foundation Center could be understood as having focused on the funders: philanthropic foundations, whereas GuideStar could be understood as focused on the recipients: charitable organizations that are funded by philanthropic foundations, as well as by individual donations and sometimes through government grant funds. The artificial boundary between these two groups was never a literal distinction, and the lines between the two sets of organizations have further blurred, thus a merger was a logical step. Candid describes this hybrid organization as composed of “data scientists and designers, sociologists and strategists, programmers and policy wonks, veterans and upstarts” ([Candid 2021](#)). As both idealists and realists, Candid promotes transparency on the inputs and outputs of the philanthropic and charitable sectors.

### **Parsing In-Depth Analysis**

The GuideStar database minimally contains data scraped from 990 forms, and it can upwards contain additional data entered by representatives of nonprofits who subscribe to GuideStar services in order to convey more information to the public regarding their annual reports, organizational leadership composition, and outcomes. The website contains information of more than 2.7 million organizations. Similar to the [Chronicle of Philanthropy](#) data scraping process described above, there was no preliminary analysis performed for this data source. This section describes the parsing process to conduct an in-depth analysis. Similar procedures were used for the search process described in prior sections, and Table 3 displays the number of organizations returned for each keyword and Boolean logic set.

*Table 3. Results for Each Set of Boolean Logic Keyword Search Terms*

<b>Keyword 1</b>	<b>Keyword 2</b>	<b>Boolean Logic</b>	<b>Number of Organizations</b>
Artificial	Intelligence	<i>and</i>	315
Machine	Learning	<i>and</i>	9,094
Data	Science	<i>and</i>	5,907
Artificial	Intelligence	<i>or</i>	1,699
Machine	Learning	<i>or</i>	272,385
Data	Science	<i>or</i>	107,653
Analysis	Analytics	<i>or</i>	10,580

Relevant organizations were scraped for title, annual operating budget (if available), and mission statement (if available). Organizations were coded for three focus areas: artificial intelligence, machine learning, and data science. The resulting database consists of 188 organizations. Figure 6 visualizes AIMS-relevant nonprofits by area.

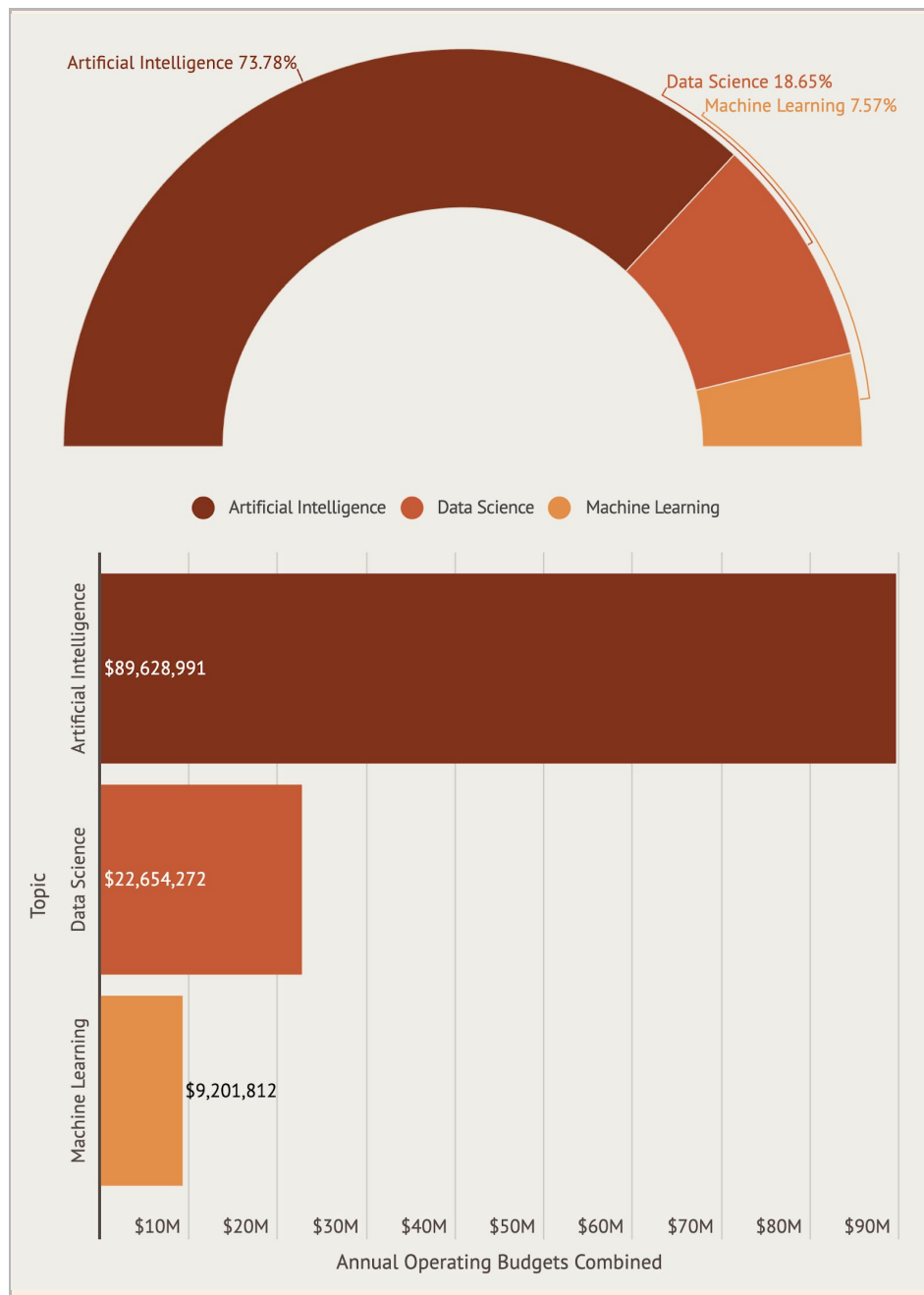


Figure 6. Nonprofit Organizations by Focus Area (Top--Proportion.; Bottom--Amounts)

Of the 188 AIMS-relevant organizations, 74% are focused on artificial intelligence, 19% on data science, and 8% on machine learning. Many of these organizations did not report their annual operating budget in the GuideStar database. Even among those that included financial information, the totals are large, with \$89.63 million dedicated annually to artificial intelligence, another \$22.65 million for data science, and an additional \$9.20 million for machine learning. ***In total, GuideStar reports at least \$121.46 million spent annually by AIMS-related nonprofit organizations.***

## E. Google AI for Social Good

The data for this fifth analysis are derived from an unstructured social media source: the [Google AI for Social Good website](#). [Google AI](#) describes its mission as “applying AI to products and to new domains” ([Google AI 2021](#)). A set of initiatives within [Google.org](#) targets social impacts, with the objective of “applying AI to address some of the world’s biggest societal challenges” ([AI for Social Good 2021](#)). The list of cross-sectoral targeted problems includes humanitarian and environmental, public health, nature and society, climate and energy, accessibility and crisis responses challenges. Figure 7 visualizes Google AI’s [Social Impact Challenge](#) grant initiative.

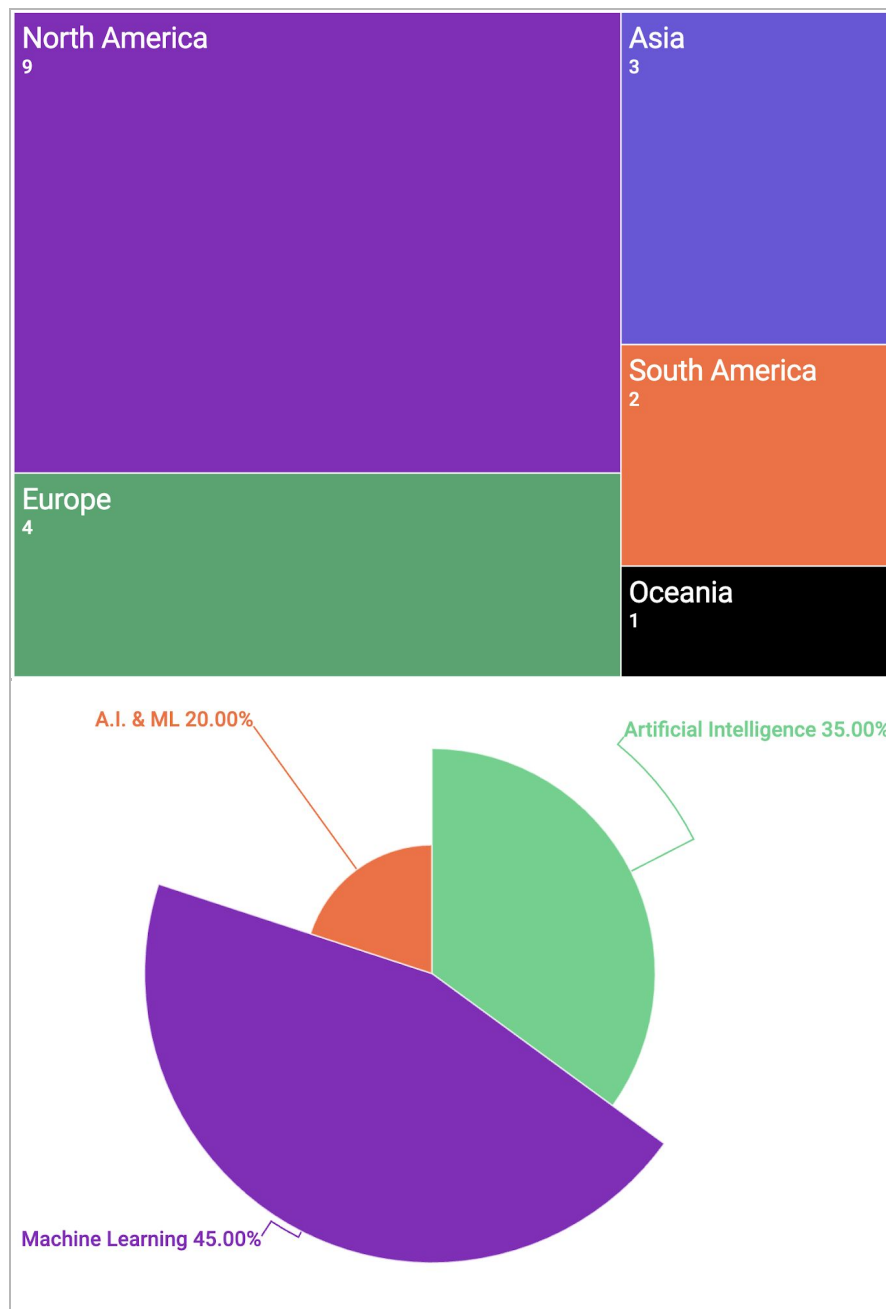


Figure 7. Grants by Focus Area (Top--Amts.; Middle/Bottom-Prop.)

Google.org reports that a global call invited organizations to submit ideas for how AI could address societal challenges and received 119 applications from 6 continents.



From these submissions, 20 awardees were granted a total of \$25 million to pursue their project ideas. Data on award winners was scraped, including the organization name, geographic focus, and project description. Awardees were then coded for focus areas: artificial intelligence, machine learning, and applications of both AI and ML technology. While all the projects involved data of some type, none of the projects were explicitly focused nor coded as data science.

Nine of the awardees are based in the United States, and the others include two from South America, one from Africa, one from the Middle East, two from Asia, one from Oceania, and four from Europe. The specific non-U.S. countries are: Brazil, Columbia, Uganda, Lebanon, India, Indonesia, Australia, France, Switzerland, the Netherlands, and the United Kingdom.

The projects include seven that are focused on AI, nine on ML, and 4 on AI+ML. There were five awards distributed to universities, including: New York University Marron Institute of Urban Management, Pennsylvania State University Department of Civil & Environmental Engineering, American University of Beirut Department of Agricultural and Food Sciences, Makerere University Department of Computer Science, and Colegio Mayor de Nuestra Señora del Rosario Department of Economics. Seven awards were to U.S.-based nonprofit organizations that are included in the GuideStar database: Quill.org, WattTime, The Trevor Project, Nexleaf Analytics, Crisis Text Line Inc., Rainforest Connection, and TalkingPoints. The remaining eight awardees are to non-U.S. non-governmental organizations: Full Fact (UK), HURIDOCS (Switzerland), Skillab BV (the Netherlands), La Fondation Médecins Sans Frontières (France), Turning Point Eastern Health (Australia), Hand Talk (Brazil), Wadhvani AI (India), and Gringgo Indonesia Foundation (Indonesia).

### **3. Conclusion & Cumulative Results**

In summary, this project analyzed five data sources that report information about philanthropic grantmaking and nonprofit organizations. The data were scraped and parsed for identification of philanthropic activities directed at Artificial Intelligence, Machine Learning, and Data Science technology (AIMS). Cumulatively, a combined total of **325 unique organizations** were identified across the five sources: Open Philanthropy, Rockefeller Foundation, Chronicle of Philanthropy, Guidestar Nonprofit Database, and Google AI for Social Good. Contributions to and expenses of these organizations tallied to a total of **\$2,578,107,211 in funds** for AIMS philanthropy.

Figure 8 visualizes funding and organizational totals by focus area. Specifically, a total of \$944,678,991 in philanthropic funds were contributed for artificial intelligence through 142 unique organizations, another \$326,451,812 in funds were contributed for machine learning through 30 unique organizations, plus an additional \$407,642,874 contributed for data science and data technology through 81 unique organizations. Additionally, \$830,705,112 in funding supported integrations in AI, ML, and DS through 46 unique organizations, as well as \$68,628,422 in funding that supported the application of AIMS technology for social impacts through 39 unique organizations. In summary, AIMS philanthropy is at least a 2.6-billion-dollar industry.

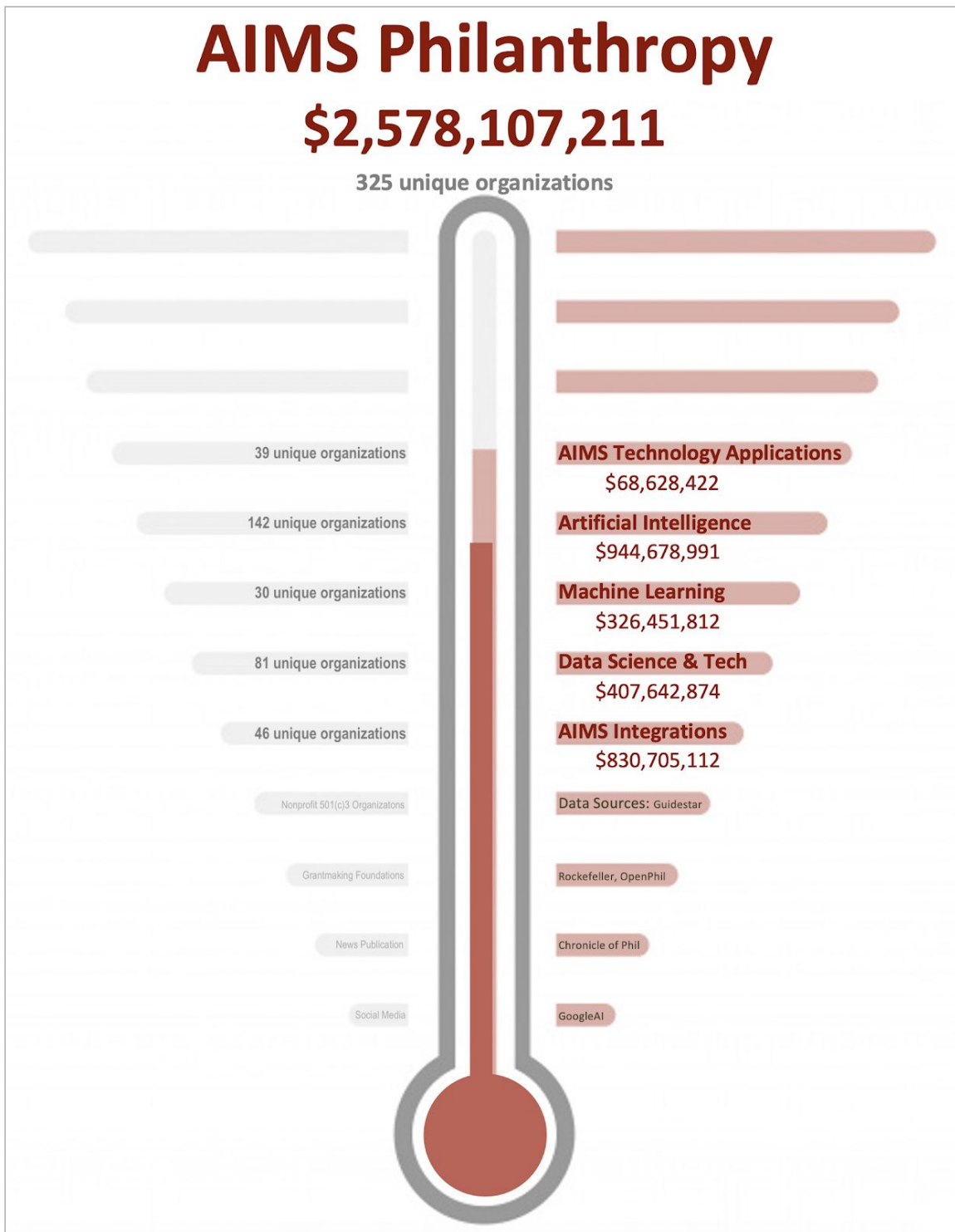


Figure 8. Total Amounts and Number of Organizations for AIMS-Related Philanthropic Contributions

#### 4. Credits & Acknowledgements

This report was written by Patricia Snell Herzog, Melvin Simon Chair and Associate Professor in the Indiana University Lilly Family School of Philanthropy and Affiliate Professor in the Indiana University School of Informatics and Computing, with the assistance of two graduate students in the IU School of Informatics and Computing: Harshal R. Naik, pursuing a master's of science in User Experience Design, and Haseeb A. Khan, pursuing a master's of science in Applied Data Science. The authors benefited from the resources of the IUPUI University Library in accessing the GuideStar and the Chronicle of Philanthropy resources, as well as graduate assistant support from the Lilly Family School of Philanthropy for the collecting and visualizing of these data.

