The Computerworld Honors Program

Honoring those who use Information Technology to benefit society

Final Copy of Case Study

Status:
Laureate

Year:
2013

Organization Name:
DataKind

Organization URL:
http://datakind.org

Project Name:
DataKind

Please select the category in which you are submitting your entry:
Philanthropy

Please provide an overview of the nominated project. Describe the problem it was intended to solve, the technology or approach used, how it was innovative and any technical or other challenges that had to be overcome for successful implementation and adoption. (In 300 words or less.)

DataKind seeks to use data in the service of humanity by bringing the latest data analysis skills, tools, and software from the tech sector to those most in need in the social sector. This initiative was designed to help the many social causes in the world make sense of the sudden flood of data now available through open data and created by technologies like digital surveys, web applications, and cellphone programs, even if they don't have the resources to pay for the services at market price. The approach entails enlisting pro bono data scientists from the private sector (e.g., Wall St., Silicon Valley) to form multi-disciplinary teams with world changers (e.g., The Red Cross, the UN, Grameen Foundation) to use data to creatively solve pressing world problems. Examples include using data from

the City of Chicago to help the Red Cross predict high-risk fire zones (Appendix 1), helping the NYC Parks Department project the future of the city's tree species using their tree survey data (Appendix 2), and improving rural healthcare through cellphone programs with Mobilizing Health. The approach is innovative because, while pro bono collectives exist for causes that social organizations understand (e.g., marketing, website design, legal), nothing exists like this for the data world. Moreover, because "big data" is suddenly a topic everyone needs to be concerned with, these groups are being thrust into a world where they could use data to improve their impact, but don't have the resources to do it. The innovation lies in DataKind's ability to bridge the gap between these two communities and bring them both into this new data-driven frontier, creating a better world for all of us.

When was this project implemented or last updated? (Please specify month and year.) Has it incorporated new technologies and/or other innovations since its initial deployment? (In 300 words or less.)

DataKind is an active project that started as a side project in October 2011 under the name "Data Without Borders". It became a full-fledged organization in June of 2012 and has been using the most cutting-edge data technologies and techniques (our volunteers come with the latest technologies on hand, including advanced machine learning algorithms, the latest ensemble methods, text regressions, R, Python, newer languages like D3, advanced data visualization in Processing and CartoDB, etc.). Moreover, DataKind is in the midst of using new technologies in-house to better match volunteer teams and track their progress.

If this is a previously submitted project that has been significantly updated and/or expanded, please describe the nature of the update here. (In 300 words or less.)

Since our last entry, the project has significantly expanded to include more projects, chapters forming around the globe, new avenues for creating learning resources about data for non-profits and data practitioners alike, and new alliances with key groups like the White House and the World Bank.

Is implementation of the project complete? If no, please describe the project's phases and which phase the project is now in. (In 300 words or less.)

The project has implemented the component of data service, provided through weekend events and long-term projects (usually 3-6 months). The services component is always being improved but does exist. In addition, arms of the company are developing to make the results of projects open source, to develop educational resources for groups, to bring this idea to more chapters and groups

around the world, and to generally spread the use of data science in the service of humanity.

Please provide at least one example of how the technology project has benefited a specific individual or organization. Feel free to include personal quotes from individuals who have directly benefited from the work. (In 300 words or less.)

We worked last spring with DC Action for Children, a group that is devoted to improving the well-being of children in the district. They attended our DC DataDive, a weekend event where we team data scientists with social organizations to see if they could turn their piles of CSVs and PDFs of data about indicators like school performance, lunch programs, income levels, and more, into something more interesting than a static report. We managed a team of volunteer data scientists to work with them over the weekend, during which they created a beautiful interactive map that DC Action could use to analyze child well-being in a more sensible way and actually dig into the data on their own (Appendix 3). The work over the weekend was so helpful that DC Action continued to work with the group for 6 months, managed by DataKind. They built the map into a usable product, which was unveiled by the Mayor of DC and heralded as a boon for child welfare in the city (http://www.dcactionforchildren.org/kids-count/dc-kids-count-data-tools). Gwen Rubinstein of DC Action had this to say: "Without our data heroes we wouldn't have been able to do this. We're four people and we can't afford that kind of data firm. That's the beauty of DataKind, and our data scientists keep pushing us, talking about making the maps even less static." The project was so successful that other non-profits are looking to use the open source code produced from the project and DC Action's funders are looking to extend this work to other child well-being groups across the country.

Would this project be considered an innovation, a best practice or other notable advancement that could be adopted by or tailored for other organizations and uses? If yes, please describe that here. (In 300 words or less.)

One of the most exciting aspects of DataKind is that the data skills it provided can be brought to any group, be they a non-profit, foundation, or for-profit company that is trying to improve the world but doesn't have the budget for analytics and data science skills. DataKind's model can be extended to foundations that are trying to provide data science skills to their grantees, or it can be extended to governments looking to understand how to use their overwhelming amount of data for the better, or it can be taken on by for-profit companies to enhance their corporate social responsibility programs. Moreover, individual chapters may use this skills transfer to expand DataKind's work in

using data for the greater good. Anywhere there are people looking to use their data skills for good, DataKind's best practices can be adopted.

If there are any other details that the judges should know about this project, please note them here. (In 300 words or less.)

Examples of our weekend events can be found here: NY DataDive: http://www.youtube.com/watch?v=3ajFogCaRbs; Chicago DataDive: http://www.shareable.net/blog/datakinds-vision-of-a-data-driven-social-change-movement. More media available upon request. TEDxMontreal talk: http://www.youtube.com/watch?v=fZ3xXXeVrIQ

We really believe we are at the beginning of a data revolution that is going to touch everyone and will require every organization to have data skills in-house to do their best work. The entire world is becoming data-driven and there is such potential for the people who wield those rare data science skills to put them to use for the greater good. We desperately need greater data ability and data literacy in the social sector, in places where we're not just writing apps to find better bars, but instead writing apps to help people find food. Through service work and general education, DataKind is taking up the torch in bringing us one step closer to a truly data-driven world. We hope to win this distinguished honor for that work and appreciate your time in reviewing our application.