



Using the Crowd for Improved Data Quality

Crowdsourcing results in 50% reduction of incomplete records.

The Company

Autodesk provides design, engineering and entertainment software to customers in architecture, manufacturing, building, and media and entertainment, including the last 17 Academy Award winners for Best Visual Effects.

Autodesk had traditionally sold through resellers and focused their sales team on individual products. However, their customers were increasingly asking for end-to-end solutions. Autodesk realized the need to understand more about its customers, so that sales could provide solutions rather than pushing products.

Autodesk needed to be able to identify its most promising sales leads and incentivize the sales team to pursue them. To do that, Autodesk needed complete data on every lead, such as company industry, size, and parent company.

The Challenge

Getting Quality Data from Unreliable Sources

Like many large companies, Autodesk's CRM data comes from multiple sources with variable data quality. Customer-entered data is often inaccurate and hard to normalize. Different sources of data have different levels of completeness, and everything needs to be integrated to produce a single, reliable data store.

Autodesk had a large budget for purchasing company firmographic data from traditional data vendors. However, these vendors could provide data enhancement for only 70% of Autodesk's database. This made it hard for Autodesk to implement their new sales strategy.

"We started incentivizing sales people based on the industry of the customer. So what that means is, the 70% yield we were getting was not good enough - 30% of sales were not being incentivized, and no one was happy with that," says Patrick Booher, Director of Enterprise Data Management for Autodesk. "So that's when we said: 70% is not good enough. "How do we get higher?"

A New Approach

Autodesk started with an in-house team, but knew they needed something more scalable and permanent. Crowdsourcing quickly became the leading candidate. "We saw that with the crowd we could mimic some of our manual processes," says Booher.

Autodesk decided to evaluate vendors based on yield (how provide data for?), quality (how accurate was the data?), scalability (could they handle large datasets quickly?), and cost.

"We evaluated other crowdsourcing vendors, and CrowdFlower had the best results. Additionally, CrowdFlower did a really good job of proving the company and service, not just the product," says Booher.



The Challenge

- Go to market and incent sales performance based on customer industry
- Find data for companies not matched by traditional data vendors (accounting for 30% of Autodesk's CRM database)
- Build a scalable, cost-effective solution

The Solution

- CrowdFlower uses the crowd to perform company research on demand
- Custom workflow collects company URL, industry, parent company, and company size for every sales lead
- API integration automatically updates records as they're added to Autodesk's database

The Results

- Improved customer industry infill rate from 70% to 85%
- Reduction in annual data licensing fees

Vendor reliability was key for Autodesk because the goal was tight integration, pushing quality data to the sales team as quickly as possible. CrowdFlower built a solution based on its proprietary business data workflow, which leverages a pool of millions of human contributors to research companies on demand.

The Solution

Autodesk sends data from a variety of sources to its sales team - purchased lists, existing customers, and inbound interest. Autodesk's Enterprise Data Management team ensures a consistent set of data to allow the sales team to find and pursue its best sales leads.

Any customer records missing key data are automatically sent to CrowdFlower via a direct API connection. The crowd first finds the official URL of the target company, and uses that (as well as additional data sources) to identify the company industry, size, and any parent companies. If necessary, additional information like company address, phone number, and description could be added, or company executive information like team names, titles and bios.

Answers from multiple contributors are cross-checked, integrated into a single answer, and sent back to Autodesk. In October 2012 alone, CrowdFlower processed 453,403 records for Autodesk.

Unlike Autodesk's relationship with other legacy data providers, Autodesk is able to purchase and retain data from CrowdFlower rather than licensing it. That means Autodesk doesn't need to re-up every year - they own their own data.

The Results

CrowdFlower improved the percentage of customer records with complete information from 70% to 85%.

"CrowdFlower was part of the overall solution that got us from directional to transactable - you can't transact on 70%." This translates into more sales leads with complete data that their sales team is incentivized to pursue.

According to Booher, CrowdFlower's advantage comes down to data completeness and accuracy. "The question is, are you getting the yield and accuracy you need from current data providers? Similar to the challenges we faced, my peers at other companies report yields of 65-70% from other data providers. The question is why? The answer is applying innovative methods towards existing problems, and that's where CrowdFlower has an advantage."

"So that's when we said: 70% is not good enough. How do we get higher?"

About CrowdFlower

CrowdFlower combines human intelligence with the scalability and efficiency of computer algorithms to offer quality-ensured processing of business information. CrowdFlower's platform provides solutions to a wide variety of data needs for enterprises such as product catalog enhancement, content generation, image moderation, and business listing enrichment. With 5 million Crowd Contributors completing millions of judgments each month for over 500 customers, CrowdFlower is the leader in enterprise crowdsourcing. The company has successfully worked with well-respected enterprises including Apple, AT&T, Autodesk, eBay, Ford, LinkedIn, Microsoft, Sears, Toshiba and Twitter.

For more information, visit www.crowdflower.com.



CrowdFlower, Inc. • 2111 Mission Street, Suite 302, San Francisco, CA 94110 • (415) 471-1920

Copyright © 2013 CrowdFlower. All rights reserved. CrowdFlower is a registered trademark in the U.S.A. and certain other countries. All other trademarks or registered trademarks, product names and company names or logos cited are the property of their respective owners.