

Customized Data Entry Solution Supporting the Building of Personalized Seeds for a Biotech Company



Building an Accurate Genetic Database To Generate Diversity

An innovative biotechnology company was looking to streamline data extraction from 1,965 multi-page agricultural PDFs to build a genetic database of plant inherent potential. Using this genetic knowledge, the company would develop “personalized seeds” by matching it with a farm’s unique growing conditions to achieve optimized field performance.

Parsing Genealogy Data from Narrative Form

The main challenge that this company faced was due to the sheer volume of data written in narrative form. In total, 13 different fields needed to be extracted from the documents’ various formats. While most of the information was easily identified, the location of genealogy data varied between PDFs. The company emphasized that high accuracy of data along with a quick cycle time was crucial in order to successfully develop different seed variants.

Due to the age and format of the documents, ARDEM recognized the use of automation would increase the risk of poor data quality. Putting together a highly skilled team versed in both analytical and technical skills would provide the best solution. Portions of the required data could be found in specific fields, however, some of the haploidization details required biological understanding to interpret and parse the correct information.

Key Benefits

- Extensive Quality Assurance
- Skilled team versed in analytical and technical skills
- Quick Cycle Time

What are the Challenges of Data Entry?

- ⇒ Good enough isn't enough when it comes to data entry which is why maintaining accuracy and consistency is the key to success
- ⇒ Before data can be utilized effectively it needs to be cleaned and normalized.
- ⇒ Data becomes less valuable over time which is why slow cycle times hurt businesses that have strict deadlines. Often, the people tasked with the data entry job are not specifically trained for efficiency.

The biotechnology company found that being able to determine self-pollinated plants would not need two different parents in order to keep certain desirable traits was important in developing the genetic database.

To maximize accuracy, a combination of effort-based extraction as well as programmed validations were used as a check and balance system for increased accuracy. Dual-key entry along with multiple verification steps are built into the workflow and a tight data security policy ensured all information processed was kept confidential.

ARDEM's Complete Data Entry Solution

By leveraging skilled people along with programmatic technology, ARDEM was able to create a process that is more effective, faster, and precise at a lower cost. An open line of communication allowed for any questions to be cleared up quickly allowing for a smooth completion of the final data set.

The Head of Digital Technologies remarked that no errors were detected in the final dataset and it was easily uploaded into their current database. The plant breeding technology company was extremely satisfied with the result and was eager to call on ARDEM for any similar projects in the future!



Similar to this Biotech Company's values - ARDEM defines the most efficient process and executes a solution to optimize both time and cost.

About ARDEM Incorporated

ARDEM Incorporated delivers success companies looking to efficiently streamline their processes and improve ROI. We implement the world's most sophisticated Technology platform combining human interaction with advance technologies to automate business processes and lower operational costs.

ARDEM Incorporated  <https://www.ardem.net>

888-359-2679 | 908-359-2600

 @ARDEMInc  <https://www.linkedin.com/company/ardem-incorporated>  <https://www.facebook.com/ARDEMIncorporated/>  pricing@ardem.net

Copyright © 2018 ARDEM Incorporated All rights reserved. ARDEM Incorporated, the ARDEM Incorporated logo, and among others are either registered trademarks or trademarks ARDEM Incorporated. in the United States and/or other countries. Other product names used in this publication are for identification purposes only and may be trademarks of their respective owners.

