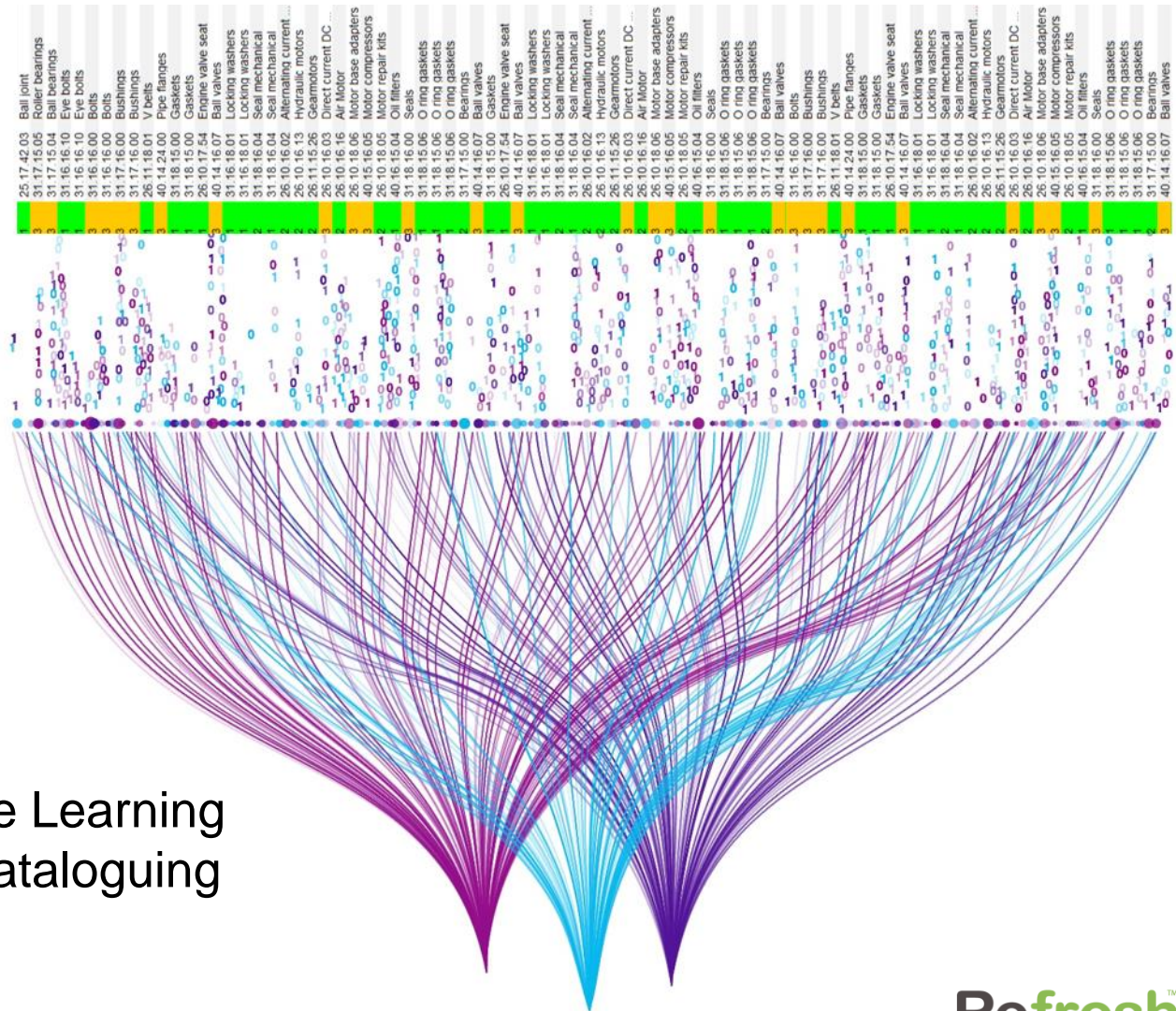


Refresh™

Machine Learning
Big Data Standardization
Enterprise Software



10 years of AI & Machine Learning
For Spend Analysis & Cataloguing

Copyright 2011-2022, Fresh Software AG, all rights reserved

Refresh™
YOUR DATA MASTERED

This presentation is the confidential property of Fresh Software A.G. - www.freshinternational.com

A.I. WARS

Last year, one of the worlds largest companies went out to tender for specialist software to help them classify millions of products...

This wouldn't be an interesting story except that more than 30 of the worlds best A.I. companies were invited. Some had expected Watson to do it... but a small company from Switzerland won.

The goal seemed simple enough – take the millions of line items that this company buys each year, and classify each one to the correct spend class. While you're at it, mine out key attributes and re-describe the items correctly where the descriptions are poor.

Live demos on real data by WebEx meant that after a few months only 5 software companies were left shortlisted.

The final showdown was a live demo week where each finalist was given a blind sample of thousands of items to classify in real-time at the customer HQ and recorded in detail for clarity and fairness.

**We emerged the winner.
This paper aims to explain why.**

1. You need lots of CLEAN training data to train an A.I. model



Machine Learning algorithms, and especially Deep Learning ones, require lots of training data to accurately predict the next item's correct class - ideally tens of millions of lines of clean data from lots of very large companies. But to get this data takes many years and many successful cleansing projects. It should also cover all technical classes or 50 000+ diverse UNSPSC's, with each class equally represented by about 500 good examples. This means you need tens of millions of clean, relevant and up to date lines – hard to get.

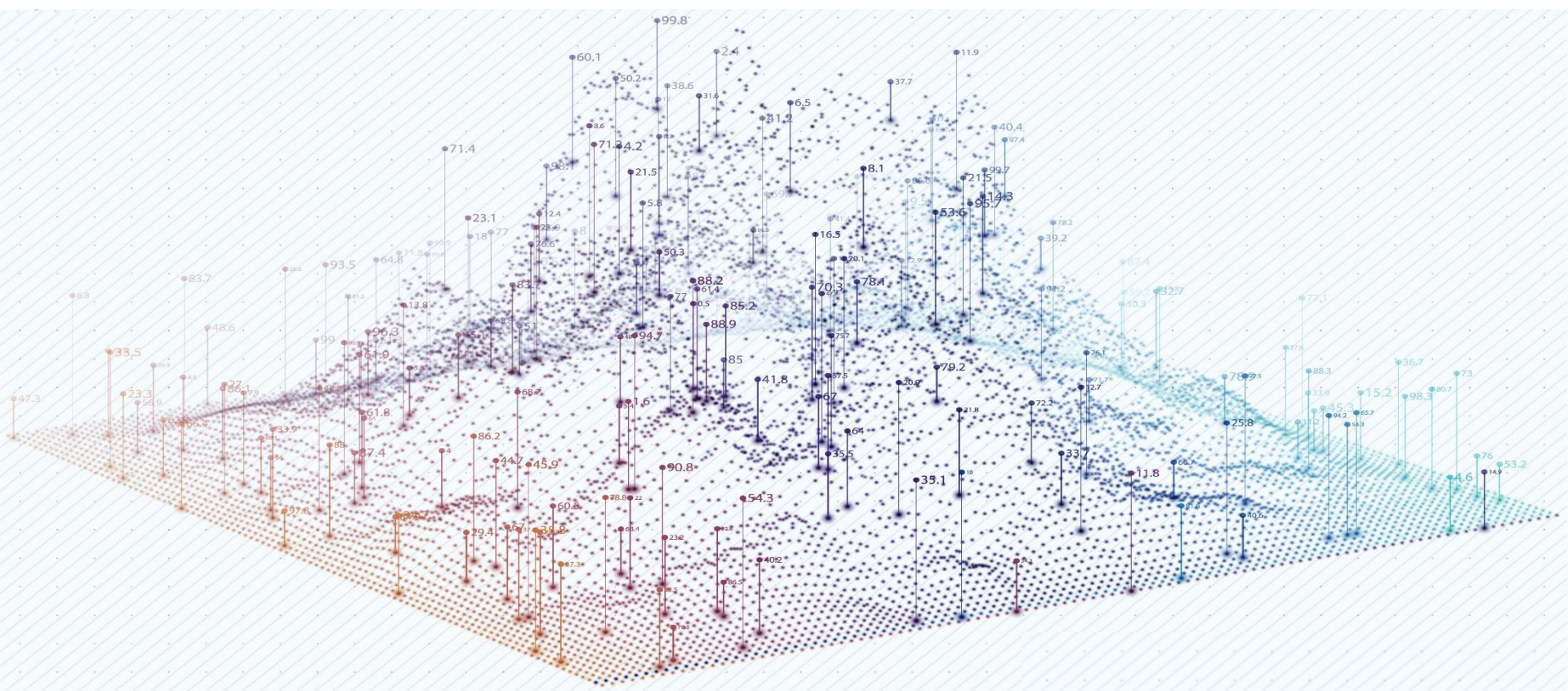
Refresh™ is the world's leading product auto-classifier and we've collected only CLEAN data for 10 years.

2. It is NOT good enough to use ‘the latest’ M.L. algorithm



Our chief programmer here at Fresh likes to remind us that while he was doing his post-grad in computer science in California, Neural Networks and Expert Systems were all the rage. That was 20 years ago. They are back in fashion again today... and still not very good at spend classification. This is because they were never written for *few texts and many classes*. Neural Nets are better for complex inputs such as voice, images or facial recognition, not so good for a purchase order text of 40 characters. Similarly, Word2Vec was written at Google for massive texts and few classes, but we have exactly the opposite situation here. Most “state of the art” A.I.s were just not made for product text. You need a bespoke one built from scratch.

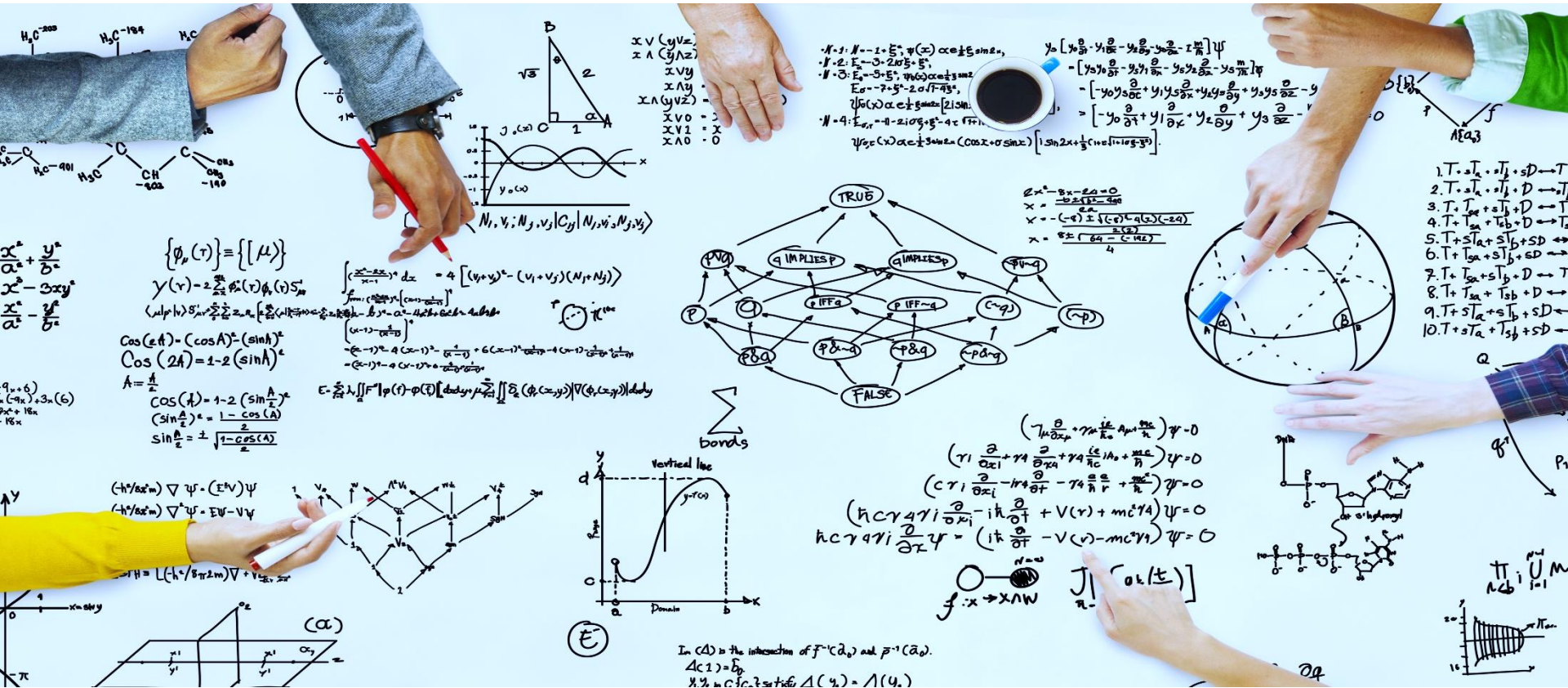
3. Building a bespoke machine learning algorithm takes time



The success of any machine learning system is actually heavily skewed towards the ingredients and not the methods of combining them in a recipe: knowing the problem, asking the right questions, thoroughly understanding the data, and combining results from your test algorithms with expert knowledge are all vital. Finding better features in preprocessing also brings massive gains, as does improving the training set. All of these things take time, trial and error feedback, updates, deep testing, and knowledge transfer.

We've been building our algorithm for a decade and every day we still see gains from years of fine-tuning.

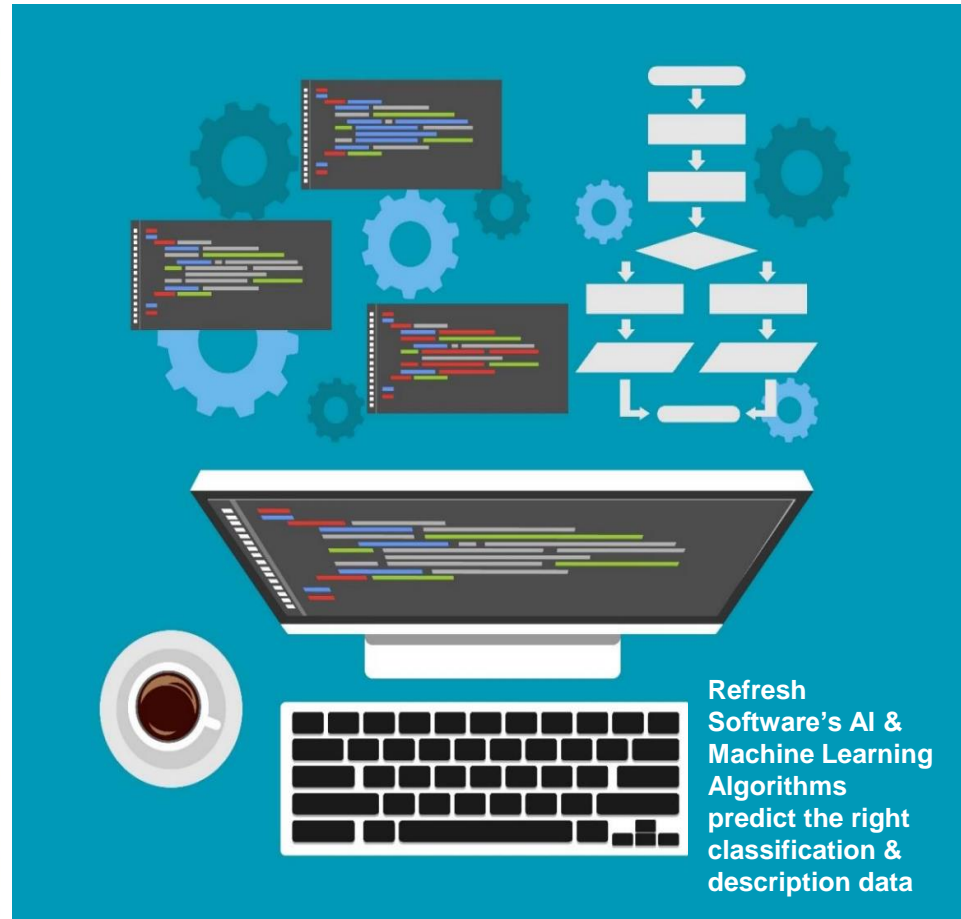
Building vs. Buying your A.I. System



There are two massive challenges companies face in DIY Machine Learning: data and talent. To build good A.I., companies need lots of their own relevant and clean training data. But maybe you haven't had a data culture for very long, or you've grown by acquisition, or you just don't have your data layer in shape yet. Even if you do have talented data scientists to build a bespoke A.I. model for you, they will still need clean training data. So using a tool like our Refresh Desktop classifier would effectively get you there faster.

This is usually a quick-win opportunity, so why not give us a call? We can help you get results from day1.

Just like Refresh™ helps these customers today:



Our A.I. algorithms, built over a decade, have learned from classifying hundreds of millions of real customer items



This automation gives you real-world savings from day 1

Available today. Installed in an hour. 1st Results in a week.



[@materialcleanse](https://twitter.com/materialcleanse)
www.freshinternational.com/contact
accounts<at>freshinternational.com

Contact us for a free live demo.