The Product/Data Fit Strategy

Column	@August 11, 2021 12:50 PM
Due Data	
Readings	
Status	
Topic	

the transformation of an idea into a product based on data must work the tensions between three dimensions: profitable, possible and desirable.

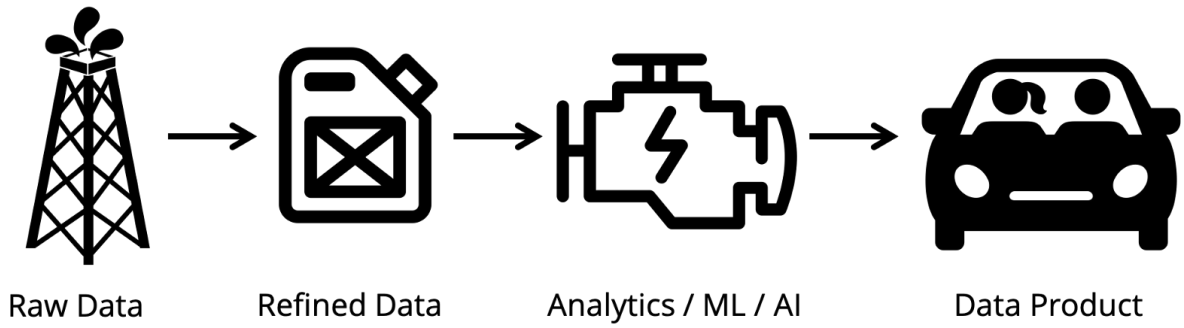
The conceptualization phase requires that visions live not just as flat perfect things for board room PowerPoint, but as tangible things co-existing with all of the dynamic tensions and forces in the world.

with products powered by machine learning and artificial intelligence (i.e. data products), the experience is no more linear or based on static business/design rules. The experience evolves according to human behaviors with constantly updating models fed by streams of data

Data is the new oil ... however most people don't go around shopping for crude oil

> Data is just like crude. It’s valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity. — Clive Humby

ML and AI are the new combustion engine, and data products are the new cars



Sometimes, parts of this chain can be outsourced.

For example, many companies successfully sell ‘analytics’ or ‘insights’. These are essentially data refineries: their product is refined data or sometimes even the engine

Then, other products use them to generate value in the market.

ROI on machine learning models:

- The investment in the model includes the cost, in time and dollars, of acquiring and storing the data. It also includes the time and cost of refining the data and training the model.
- The return on the model depends on two components:
  - The accuracy of the model
  - The business value generated from a correct prediction (in dollars, clicks or another quantifiable metric), and the business cost of a wrong or inaccurate prediction.

The key of data strategy is: focus on increasing the return, not increasing the investment. This sounds obvious, but is often lost in the hype around data and AI

Some people focus exclusively on the amount of data. They are the ones who always complain that "We need more data!" or brag about how "We are generating so much data."

But these phrases are often markers of a poor data strategy. They emphasize the investment instead of the return.

The real goal is to build better cars with more efficient engines, not to accumulate more crude oil.

Bottom line:

- Data strategy is about building better products, not accumulating more data or using more sophisticated methods
- To do that, you must understand how your product uses data to generate value, and focus on increasing the ROI on your models
- One way to do that is by improving the accuracy of the models, but you will quickly run into diminishing returns.
- Another way is by finding more valuable questions that your data can answer, which can result in better ROI even with less accurate models.

- This is how you find product/data fit: iterate to simultaneously increase the value of your data, your models and the questions they are tackling