Contents

1. Time Line

- Pre-1980's: The Execution Era
- 1980-2000: Porter's Forces
- 2000-2018: Information Rules
  Therefore data is oil (Data Moats)

2. The New Moats

- This bring us to Jerry Chen!
- VC at Greylock Partners (very well respected firm)
- Article is written at the tail end of the "data is the new oil" period
- He begins by stating that to understand profitability you have to understand "moats" – what is this?
- **Moat** turns out to be a fancy way of describing a long-term profitable force! His examples:
  - IP / Trade Secrets
  - High-switching costs
  - Brand and customer loyalty (remember the story about "Nobody is fired for buying IBM?"
- What is his investment thesis?
  (1) The next generation of "defensible moats" is going to be in "systems of intelligence" which sit upon two other data layers: "systems of engagement" and "systems of records"
  (2) The next generation of "defensive moats" will be built along traditional company verticals (HCM, CRM, IT, ERP, Finance, etc.)
- Lets discuss each:
  (1) **Systems of Record:** Data is valuable, but only to the extent that it is useable and accessible. Every generation there is usually a dominant player in

each category. Chen would (probably) say that these companies are defensible moats, just not that strong a moat.

- CRM (Customer Relationship Management) (Salesforce)
- HRM (Human Resource Management Systems) (Workday, peoplesoft, zenefits, trinet)
- ERP (Enterprise Resource Planning) (SAP, Oracle) – generalized logistic software

(2) **Systems of Engagement:** How do users interact with the systems of record?
- Slack, WeChat, etc.
- Each successive generation builds on the previous.
- For these system, the argument is that "holding" data is different than "exposing" data to the user.

(3) **Systems of Intelligence:** Leveraging the underlying systems of engagement and intelligence, this generation of companies will mix and match data in order to provide value.
- This moat is defensible because there is a virtuous cycle of model performance.
- Consider the case of modeling customer retention. If there is a company that does it "best" you will use them, which means they get better data which means that they will be better and thus other companies are more likely to use their product.

- In essence, Jerry is arguing that data isn't the moat – but the algorithm is the moat!
- We will call this period the "algorithmic moat period" (2018-2020) and it was *short*.
- Following Jerry's argument, where would you want to invest? What companies are going to be profitable?
  - Those companies that build upon these systems of engagement and record that can directly provide business value by combining multiple data sources to build models which a single other source isn't able to do on their own.
  - You want to invest in companies that (a) provide obvious business value, (b) can use multiple data sources and (c) have an algorithm / model which benefits the more data you get.

The Empty Promise of Data Moats

Andreessen-Horowitz.

- Current "Best regarded" venture capital firm

- Marc Andreessen: founder of Moasic (Netscape/Firefox) which was sold to AOL for 4.3 Billion. He then founded Loudcloud/Opsware (with Horowitz) and sold that to HP for 1.6B in cash
- Ben Horowitz: Engineer at SGI (NVIDIA of the 80's and 90's) and then Product Manager at Netscape

  Wrote a book called "The Hard thing about Hard things" – considered a class in the startup world.
- Invested in a ton of famous companies (airbnb, Twitter, Facebook, Zynga, Substack, Lyft, Roblox, Figma, etc.)
- "a16z"

The Empty Promise of Data Moats.

- From 5/2019 – right around when "algorithmic moats" were killed as a means of creating sustainable profit.
- Article hits quite a bit on the "lack of value in data" as well as the ability of an "algorithm" to be a data moat.
- They define a "network data effect" as the defensible moat that Jerry Chen spoke of. You have a model that gets better with additional data and therefore you can keep ahead of your competitors on the basis of its performance.
- They call the previous argument around "algorithmic moats" a "flywheel" effect.
- However, this article says that what they are seeing is that data is more of a "liability" and isn't going to create a moat:
  - Spending resources
  - Marginal information collection costs skyrocket after a while
  - The value of marginal information also goes down
  - Algorithims don't perform linearly better with more data

    They talk about a support chatbot – once you get the easy support questions working, everything else is one-off and it isn't an 80/20 rule, but much much worse.

Table 1. Breakdown of question distribution for chatbot

Pct	Question	Value
10%	How do I Update my Account	PROFITABLE
3%	Other Question #1	BREAK EVEN
1%	Other Question #2	NOT PROFITABLE
.1%	Other Question #3	NOT PROFITABLE
.01%	Other Question #4	NOT PROFITABLE
.001%	Other Question #5	NOT PROFITABLE

  - Data becomes stale with time – and as companies see value in that data just keeping up becomes more and more costly.
- This doesn't mean that algorithmic moats don't exist – just that they aren't as common as expected.
- The final piece of advice that they state "Greater long-term defensibility is more likely to come from *packaging differentiated technology*"

  Advice which could have come directly from information rules!

The New Business of AI.

- From 2/2020 – about 9 months later.
- Still from a16z.
- Their argument: Many people believe that AI based businesses will behave like traditional software business
- What is traditional software? *Information Good*
  - High up-front cost, zero cost of reproduction
  - Price at a customer's willingness to pay
  - Differentiation and Price Discrimination are incredibly important and powerful
- However, they have found AI Businesses do NOT look like this. Why? Mainly Marginal Costs are too high, meaning that gross margins are much lower than expected
  (1) Cloud infrastructure costs ($$$) 25% of revenue!
    - *Training* can cost thousands to millions of dollars. It's NOT a one-time cost. As more data comes in you constantly need to retrain and verify. "Data Drift"
    - *Inference Cost* Even once you get a model trained, the evaluation costs can be incredibly high. As an example, when I was at Sega we were looking at a few different ways to run our nightly LTV calculations. One of them was a bit more accurate but we would have had to use a much more costly set of servers to do it – so we didn't.
    - *Storage Costs* Models which are trained on rich media (images and video) tend to use incredible amounts of storage.
    - *Other cloud operations* Egress/Ingress, model verification, running APIs and expensive servers to use RAM to lower latency all add up rather quickly.

  (2) Human Intervention (Consulting and In-the-Loop) mean that that reproduction isn't free (10-15% of revenue)
    - Thorny Edge Case Problems
      * Part of the "Consulting Problem"

        * Constantly having to adapt your model and techniques to different systems, data generating processes, etc.
        * Customer-by-customer you need to make costly adaptations.
        * Anyone will put anything into a model and you have to be constantly checking and on the look out for new ways that the system will break.
    - Human in the loop models ($$$)
        * Many times the model will require a human in the loop to verify, etc.

- In the end this article is saying that AI companies are looking more like "Service" companies than "Tech" companies.
    - For VC's "service" is a four letter word!
    - VCs do not want to invest in consulting firms, they want to invest in technology firms.
    - Why? For the reasons above – there isn't the built-in scalability that you would have with a "technology" company.
    - Consulting firms do not behave by as Information goods and thus they aren't as profitable
    - This means that the exit that they will get is significantly lower than what they would otherwise.
- There are two conclusions in this article: one aspirational and one operational.
    - Operational #1: Lower *custom* costs (decrease model complexity, decrease problem domain) to appear more like a technology company. Doing this will increase your margins
    - Operational #2: Price in variable costs (raise prices). Consulting is *not an informational good* so price not on willingness to pay, but on cost to provide the service.
    - Operational #3: Embrace Services, but treat them as an Extra and charge as such!

3. Asserts.ai – Strategy Proposition

- Use data in order to understand infrastructure
- Leverages other systems and applies "magic" on top of it.
- Specifically prometheus, k8s api and other common services
- How does this fit into the above framework?
- Do we think that this is a data moat or not?

4. Ambient.ai – Strategy Proposition

- Use AI in order to make your security more efficient.
- They use existing cameras, not requiring you to make any additional hardware purchases.
- How does this fit into the above framework?
- Do we think that there is a data moat or not?