

Frictionless Data

<http://frictionlessdata.io/>

[2m video](#) - [6m video](#)

The Problem



We want “insight”

Whether it’s into
how to stop global warming
or ...
the quickest way to get to work

We want to use data to do that



How much CO₂ is emitted, what renewables could we use – and how much do they cost?

BUT THERE'S A PROBLEM

To understand the problem I
want you to think about
💡insight like a 🍰 cake

🍰 we combine together ingredients to make
delicious cakes

💡we combine together data to make delicious
insights

		
Raw ingredients	Corn grown in a field. Hens laying eggs etc	Statistics agency does surveys, geographers make maps etc
Logistics Getting those items from source to you	Loading them on a truck, put them on a ship, take them to a shop where you can buy them and take to your kitchen!	Move data from system to system across the internet, put them in big databases, then you download them and put them in your own spreadsheet, database or program
Preparation & Baking	Mix all those ingredients together and put them in the oven	Analyze that data, combining different datasets together, and finally do some processing to render out a beautiful infographic map!
Consumption	Ummm, i love chocolate!	Look at that infographic go!

The Cake!

Imagine that to make a cake you'd first have to collect the ingredients – go to the farm and collect the corn, mill it, cart back to your house **AND only then** could you start baking!

Rather than a quick trip to the store and a couple of hours in the kitchen you would spend weeks!

This is where we are with data today!

With data

We spend most of our time
collecting and preparing the
data AND are left with only a
small fraction to work on
actually turning it into
delicious insight!



SUMMARY

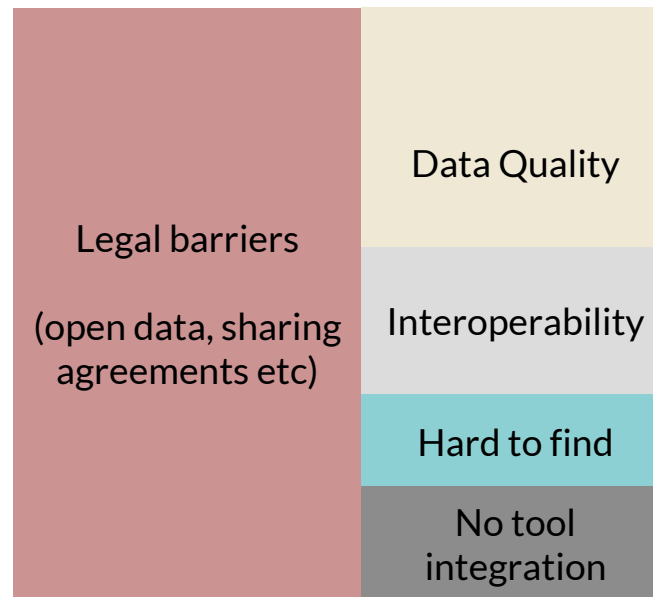
There's huge amount of friction in getting and using data ...

Which stops us getting insight to solve important problems...

SO ...

**What can we do about
it?**

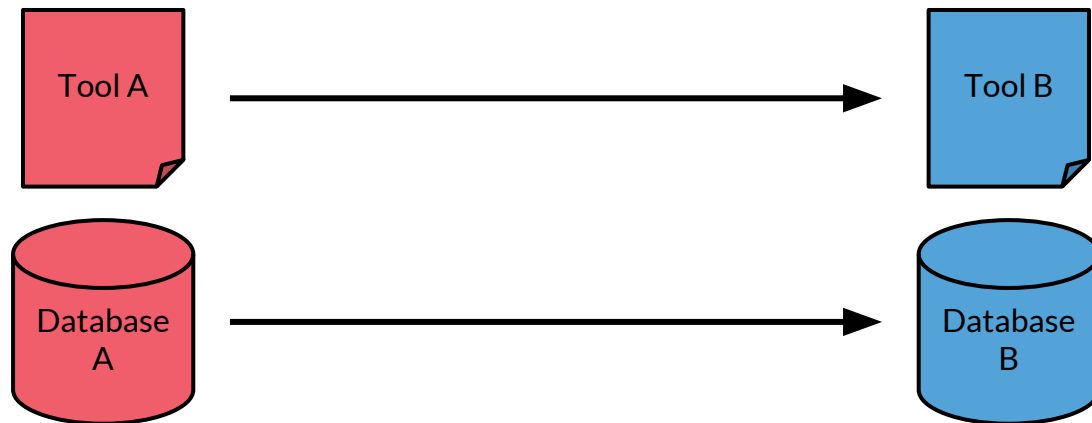
There are lots of sources of friction – and lots of things we could improve ...



Focus on just one

Logistics

Eliminate friction getting data A \rightarrow B

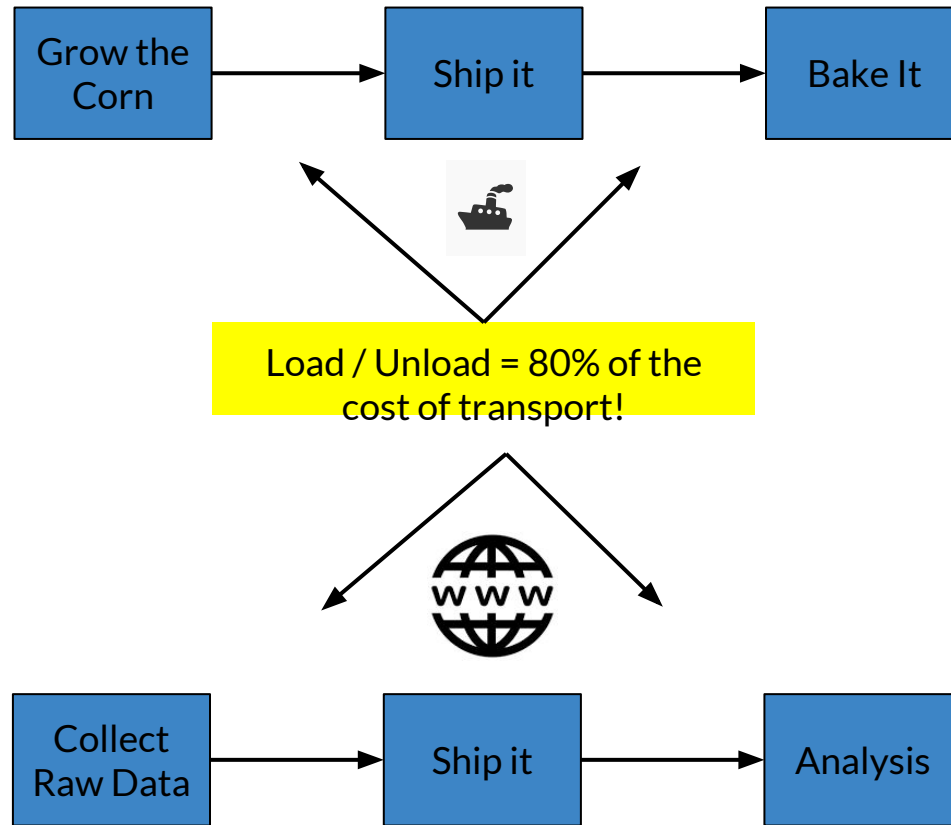


Goal

Ship data from tool A -> tool B
With zero friction



Cakes and Insight - Revisited!



A black and white photograph showing a group of men, likely banana workers, walking in a line. They are carrying large, heavy bunches of bananas balanced on their heads. They are dressed in simple work clothes like shirts and trousers. In the background, the wooden structure of a ship's deck is visible, with a railing and a person standing on an upper level. The scene is outdoors, possibly at a dock or a large warehouse. The text "Shipping 1955" is overlaid in a large, white, sans-serif font, with "Manual, Slow, Costly" and "(and Dangerous)" in a smaller font below it.

Shipping 1955

Manual, Slow, Costly

(and Dangerous)



OPEN KNOWLEDGE



In just twenty years from 1956 to 1976 containerization increased productivity by 7000x (Tons Per Man Hour went from 0.6 to 4200 and hours in port from 504 to 18 – weeks to less than a day!)

Shipping Today

Automated, Fast, Cheap

(and Safe)



OPEN KNOWLEDGE

Can we do containerization for data



Yes we can!

Data Packages
=
Containers for Data

Bananas



Steel
Container

Spreadsheet



Virtual
Container



It's about the tools!



OPEN KNOWLEDGE

Validate >
Store and search >
Import to your tool >
etc>

Labels on the outside of the
box

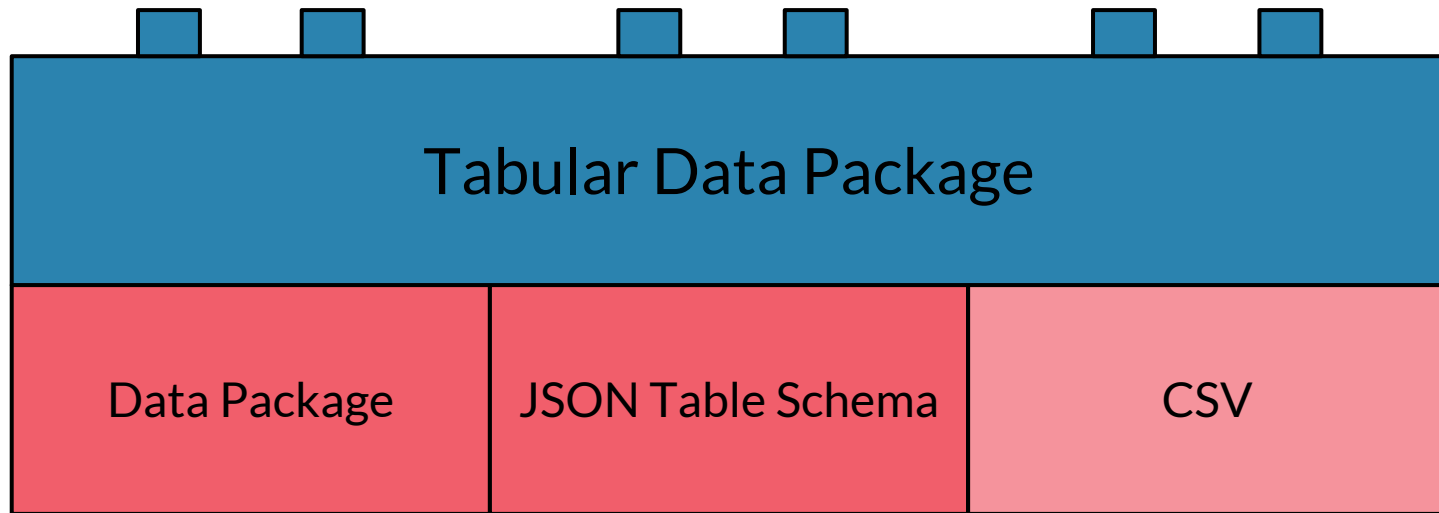
=>

Allow tools to automatically
process the contents!

Data Packages

Extendable and customizable
to suit your data

Tabular Data Package



<http://frictionlessdata.io/guides/tabular-data-package/>

<http://frictionlessdata.io/guides/data-package/>

<http://frictionlessdata.io/guides/json-table-schema/>

Summary

1. We want to turn data into insight – quickly, easily, reliably
2. Today, there's huge amounts of friction – you can spend weeks preparing data to be able to do a few hours of analysis
3. By containerizing data – putting it in data packages – we can dramatically cut the costs of acquiring and integrating data creating a world of frictionless data
4. => More insight, more efficiently
5. Together with partners around the world we're implementing the frictionless data future today

To find out more and start
using data packages today -
visit:

<http://frictionlessdata.io/>