Title Slide:

Slide:

Out in California, where it's warm, some researchers at USC recently decided to calculate humanity's total storage capacity—how information we could store on every piece of paper and every hard drive in the world. They came up with the number 295. 295 exabytes. One exabyte is equal to a million terabytes. Still, reading that number doesn't really help us understand the scale of this information: our brains are not wired to handle data on that scale. Our ancestors on the savannah didn't have to comprehend billions of things at once, but nowadays we do.

Slide:

Though our brains can only handle about 30 pieces of information at once, our own genomes have 3 billion base pairs. It's been a decade since the full sequence of the human genome was published, but still no one has figured out how to make my hair blond. People hoped that once we had sequenced the human genome, we could use that knowledge to create gene therapies to cure diseases. But it hasn't, and a big reason for that is the sheer amount of data: 3 billion base pairs.

Slide:

The sequence of the human genome alone is an overwhelming amount of data: it would take 9.5 years to read the entire sequence of one person's genome aloud without stopping. Still, it has a huge amount of potential to improve people's lives.

Slide: [image of genetic code browser]

So the question becomes, how can we make 3 billion data points more manageable? How can we present this information in a useful and accessible way? One way is with a picture: this is a browser for the human genome. It displays genetic data at three different levels. It displays the big picture and the fine details at the same time: a researcher can look at a single gene closely and see where it falls in the entire chromosome simultaneously.

Slide:

That browser was created by Ben Fry, a CMU alumnus. He was a design major with a computer science minor when he was at CMU. He got a PHD from MIT's Media Lab: his dissertation proposed that data and design should be brought together in a process called computational Information Design. Fry owns his own design firm and has created useful and beautiful visualizations of complex data. Besides his visualizations for genetics, he has been involved with visualizations of geography. For instance, he made of all the roads in the continental United States. He's also created illustrations for SEED magazine and had his work on display at the Museum of Modern Art.

Slide:

Data can be overwhelming and lead to bad decision making. A failure to understand economic data contributed to the recent global recession. But data is ubiquitous now, and if we can use it, it can help us make better decisions and teach us about ourselves.

Slide:

Nowadays much of the work people do looks something like this: people have to make leaps from data to conclusions and then somehow communicate those conclusions in an understandable way. Ben Fry's work draws the arrows in the diagram: to make the leap from data to conclusions, he creates visualizations of data that help people find useful patterns. To make the lead from conclusions to communication, Fry incorporates the aesthetics of design.

Slide:

Data can be ugly, overwhelming, confusing. Who wants to spend their time staring at a spreadsheet? But data has patterns. And people enjoy finding patterns—it's how we understand things, putting together pieces of information to create a larger meaning. Humans are naturally pattern-seekers, and patterns are beautiful to us. Ben Fry's work as a designer is to uncover this beauty. He finds beautiful and useful patterns in data and shares them with the world.