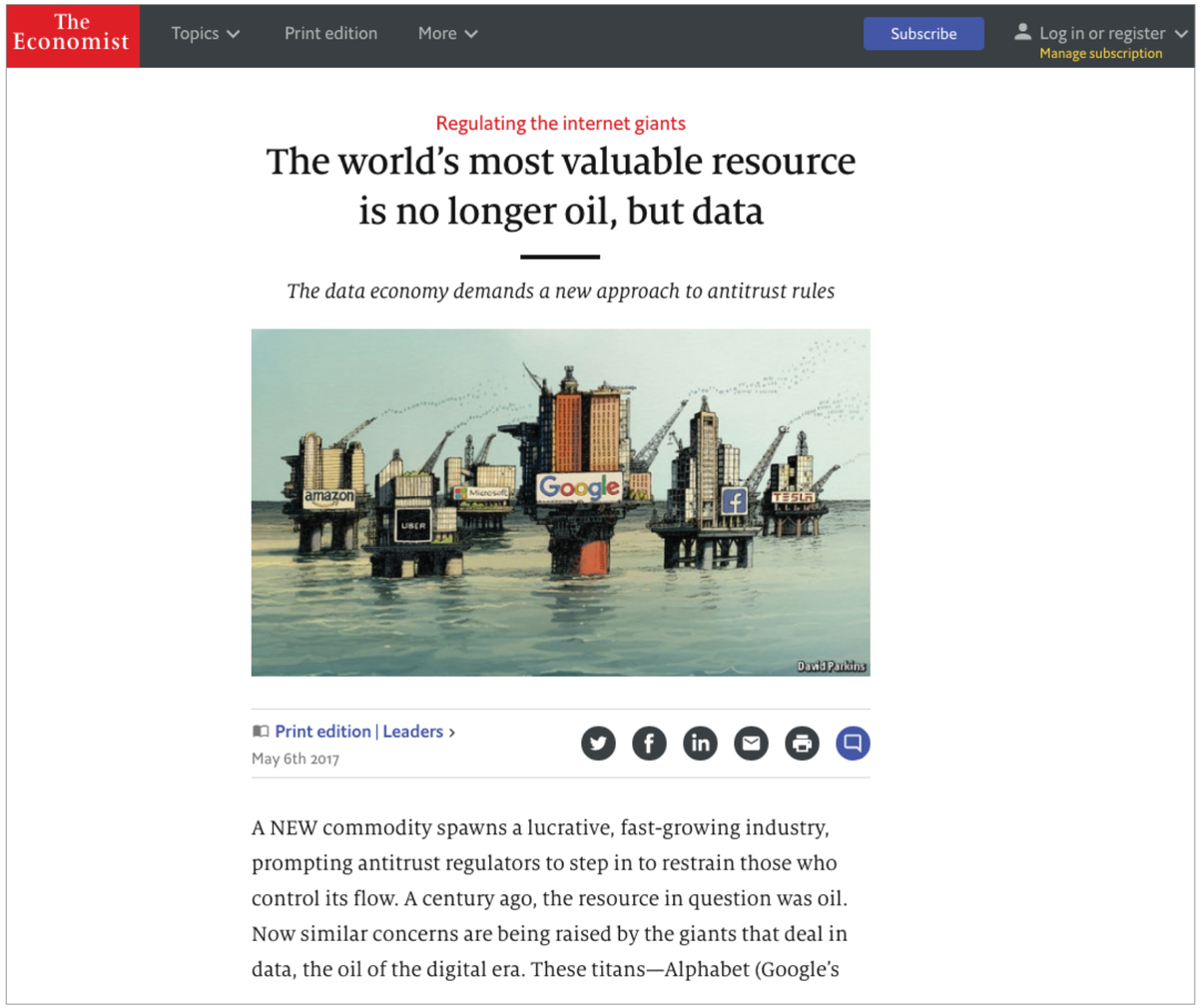
Daikin Houston November 2, 2017

“Datafication” — the rise of big data and the application of AI to everything

Hugh Dubberly
Dubberly Design Office

presentations.dubberly.com/Daikin_IoT.pdf

For the 21st century, data will be what oil was for the 20th century.



The Economist | Topics | Print edition | More | Subscribe | Log in or register | Manage subscription

Regulating the internet giants

The world's most valuable resource is no longer oil, but data

The data economy demands a new approach to antitrust rules

Print edition | Leaders ›
May 6th 2017

A NEW commodity spawns a lucrative, fast-growing industry, prompting antitrust regulators to step in to restrain those who control its flow. A century ago, the resource in question was oil. Now similar concerns are being raised by the giants that deal in data, the oil of the digital era. These titans—Alphabet (Google's

— The world's most valuable resource is no longer oil, but data
The Economist, May 6, 2017
https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource

"If you went to bed last night thinking you're an industrial company, you're going to wake up this morning as a software and analytics company."

— Jeff Immelt, former Chairman and CEO, General Electric

Minds + Machines 2014
www.ge.com/stories/industrial-internet

GE sold its appliance group for $3.3 billion.
Google bought Nest for $3.2 billion.

GE APPLIANCES

- 12,000 employees
- $5 billion in revenue
- 7 to 10 million appliances per year
- Appliance Park factory complex

nest

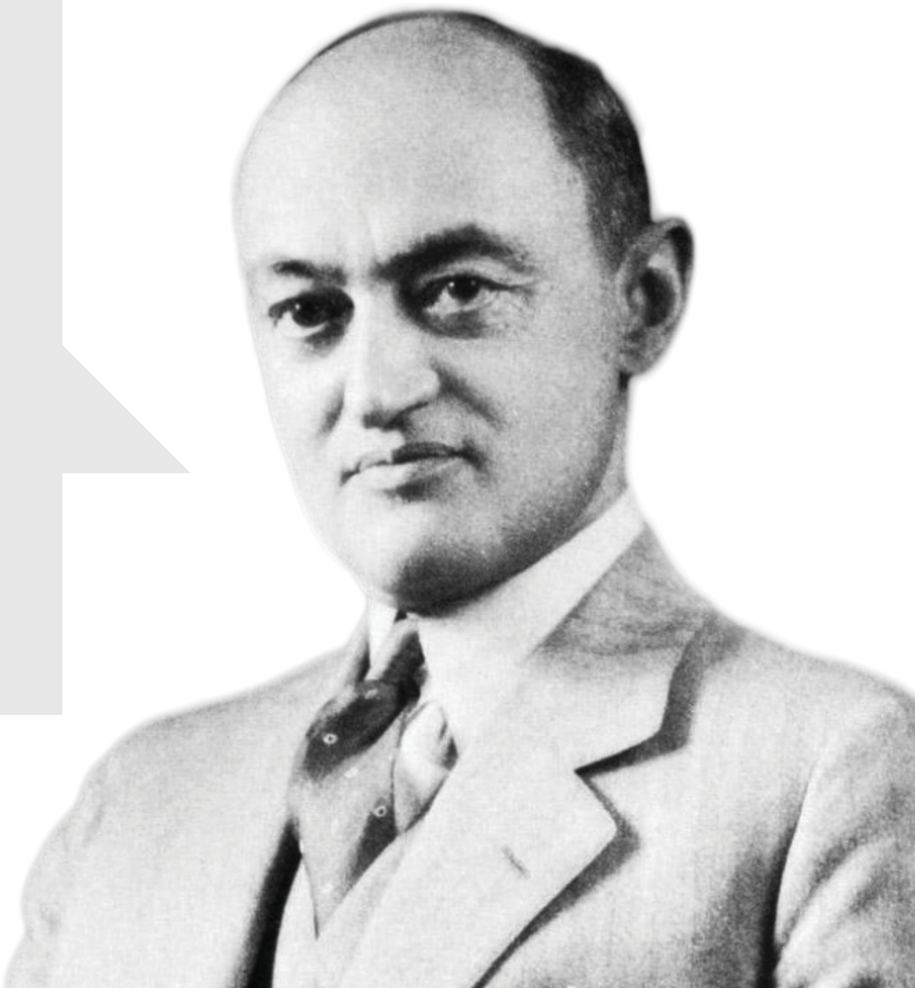
- 250 employees
- $250 million in revenue
- 480,000 thermostats per year
- Manufacturing outsourced to China

“Creative Destruction is the essential fact about capitalism.”

“Capitalism, then, is by nature a form or method of economic change and not only never is but never can be stationary....

The fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers’ goods, the new methods of production or transportation, the new markets, the new forms of industrial organization...

that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one.



—Joseph A. Schumpeter, (1942) *Capitalsim, Socialism and Democracy*, pages 82-83.

We have seen five industrial revolutions; what will be the sixth?

1
1780 - 1840
First industrial revolution
- iron making
- cotton spinning
- water power
- steam power

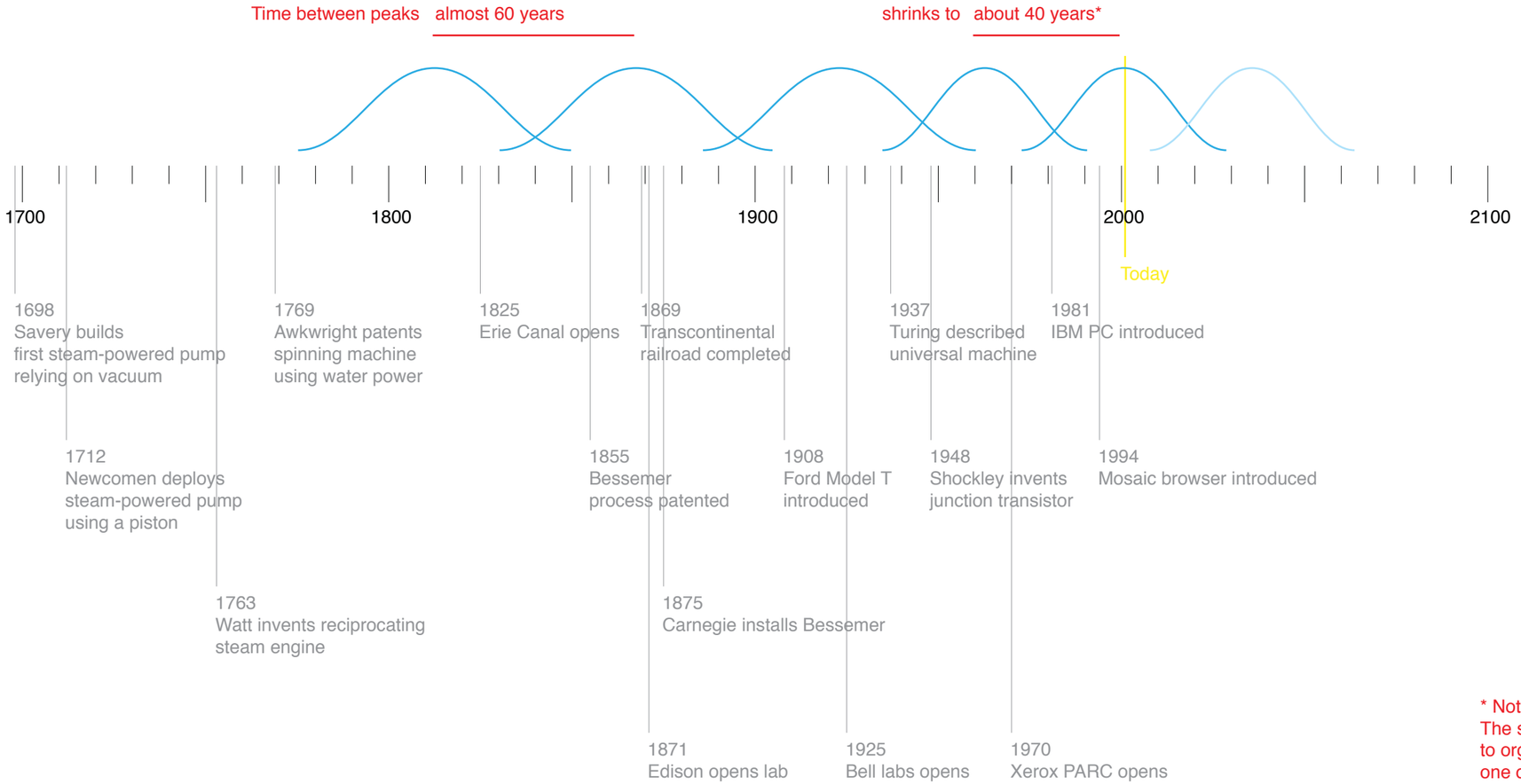
2
1840 - 1900
Second
- steel making
- railroads

3
1900 - 1950
Third
- electrification
- internal combustion

4
1940 - 1980
Fourth
- petrochemicals
- plastics
- electronics
- computing

5
1980 -2020
Fifth
- personal computers
- internet

6
What comes next?



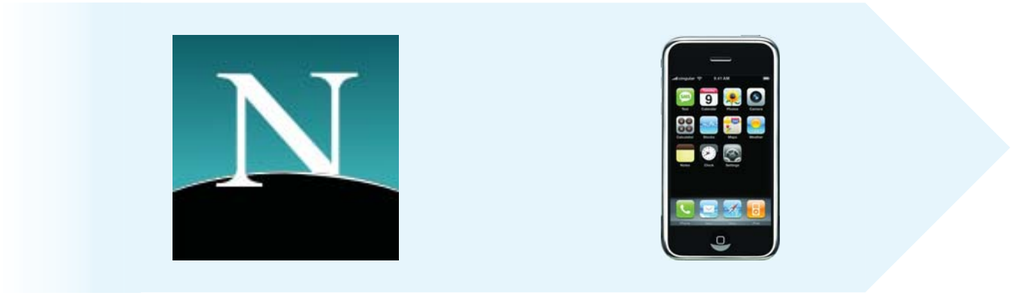
* Note:
The shrinking cycle time may be due to organized research and possibly to one or more “network effects.”

Google CEO Sundar Pichai has predicted “AI First”.

PC Web Mobile “AI First”

Google CEO, Sundar Pichai spoke at the #MadeByGoogle event on October 4, 2016

Siri co-founder Dag Kittlaus is focused on “assistants”.



Web Mobile Assistants

Siri co-founder Dag Kittlaus unveiled Viv at TechCrunch Distrupt NY 2016

In each era, the dominant technology is a “platform”— a system on which others can build.

Productivity Applications	Web-based Services	Mobile Apps	Monitoring + Prediction Services
			
PC	Web	Mobile	AI + Data + IoT
~ 1981	~ 1995	~ 2007	Today

In the early 1980s,
personal computers changed the way business is done.
Think of this as *going digital*; everything is becoming a computer.

> *"...software is eating the world.*
> *...we are in the middle of a dramatic and broad technological and economic shift in which software companies are poised to take over large swathes of the economy...*

— Marc Andreessen, founder, Netscape and Andreessen-Horowitz

In the mid-1990s,
the internet changed the way consumers + business communicate.
Think of this as *getting connected*; everything becomes a web service.

"I envision a 21st century form of business where the everyday consumer is helping shape the social contract …

It's a business world that is moving from value-based transactions to values-based partnerships."

— Paul Polman, CEO, Unilver

In 2007,
smartphones made computing ubiquitous—and turned it into communicating.
Think of this as *always connected*; anywhere, anytime.

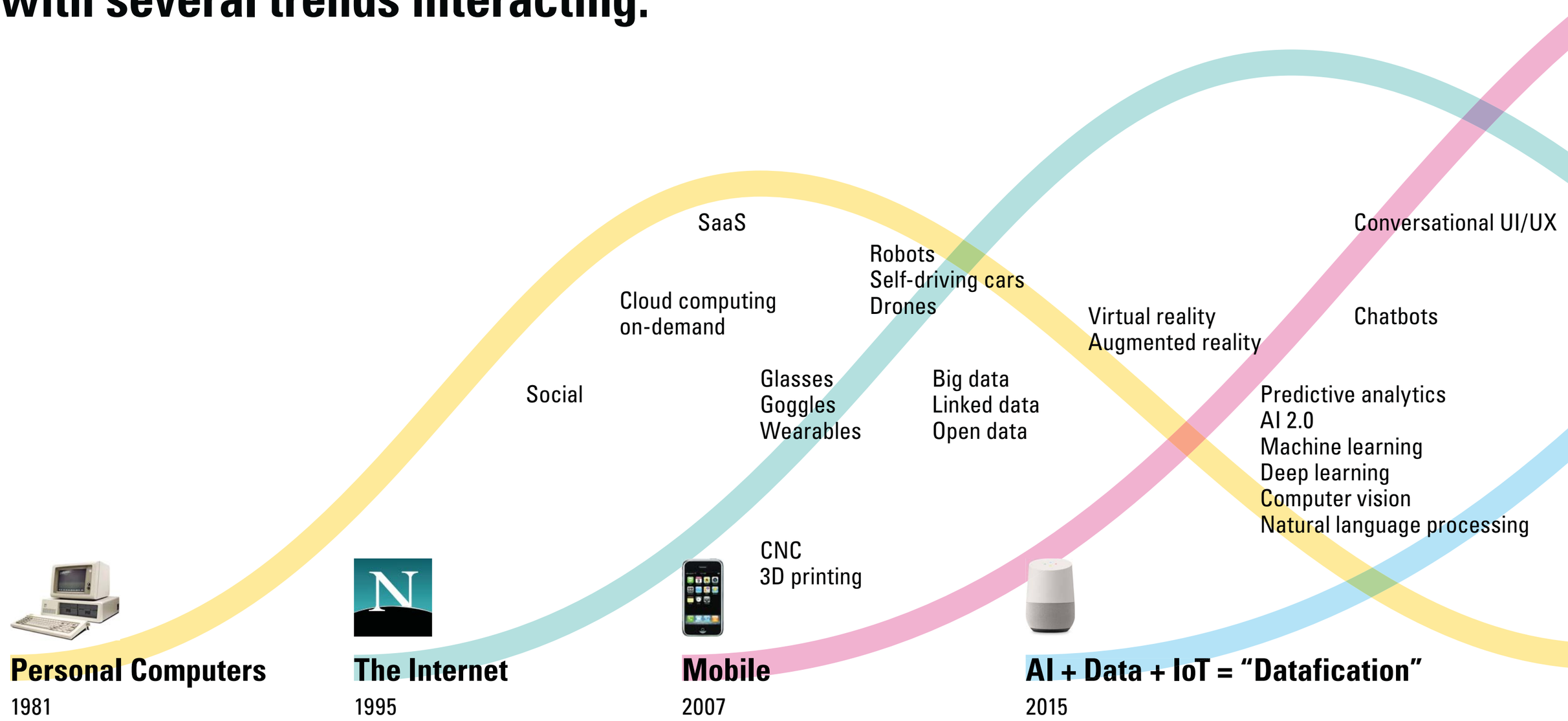
"Design has also evolved from the design of objects both physical and immaterial, to the design of systems, to the design of complex adaptive systems.

— Joi Ito, Director, MIT Media Lab

Today is like 1981, 1995 and 2007 all over again.
You can see the next wave coming. It goes by many names:

- Internet of Things (IoT)
- Internet of Everything, Cisco
- Industrial Internet, GE
- Smarter Planet, IBM
- Living Services, Accenture
- Platform World, Sapient.Publicis
- Social CRM or Social Business
- Digital Engagement
- Digital Transformation
- “Datafication”

The eras Pichai + Kittlaus describe can be seen as “waves”, with several trends interacting.



Combinatorial innovation explains how trends work together.

"We're in the middle of a period of... 'combinatorial innovation'...
In the 1800's, it was interchangeable parts.
In 1920, it was electronics. In the 1970s, it was integrated circuits.
Now what we see is a period where you have Internet components...
and capabilities to combine these components parts
in ways that create totally new innovations."

—Hal Varian, Google's Chief Economist and UC Berkeley Professor

“Datafication” is a series of trends; none capture the whole.

- **Sensor Revolution** — printing sensors on chips; installing measurement capability all around us.

- **Smart Things** — adding “intelligence” to everything, by building in microprocessors.

- **Internet of Things (IoT)** — connecting sensors and smart things to the cloud.

- **Big Data** — recording everything that happens in the physical world and online.

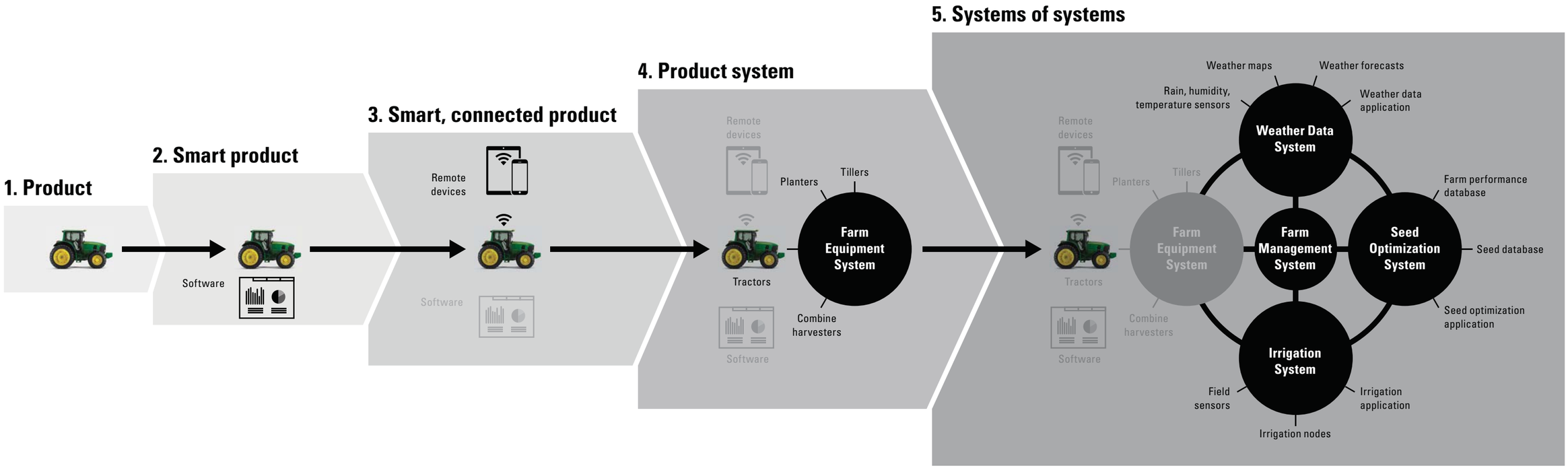
- **Cloud Computing** — putting massive resources online, so that the marginal cost of computation falls to zero.

- **AI, ML, DL, NLP, CV** — algorithms (often run in the cloud), making sense of the measurements we record.

“Datafication”

An example

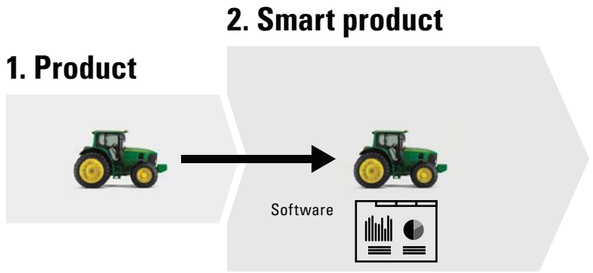
Harvard Business School professor Michael Porter writes about systems of systems.



—Michael Porter and James Heppelmann, How Smart, Connected Products Are Transforming Competition
Harvard Business Review, November 2014
https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition

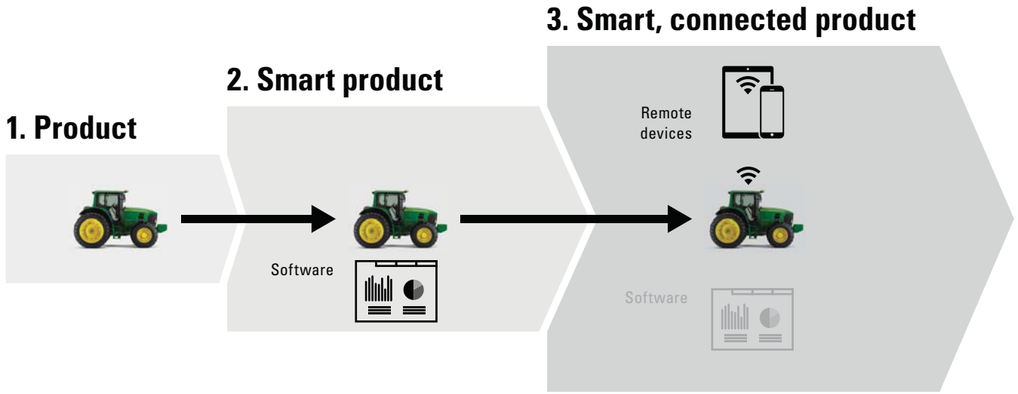
Products are becoming “smart.”

Product
+ Sensor
+ Computer
= Smart Product



Sets of smart products are connecting.

Smart Product
+ Network
+ Cloud Service
= Smart, Connected Product

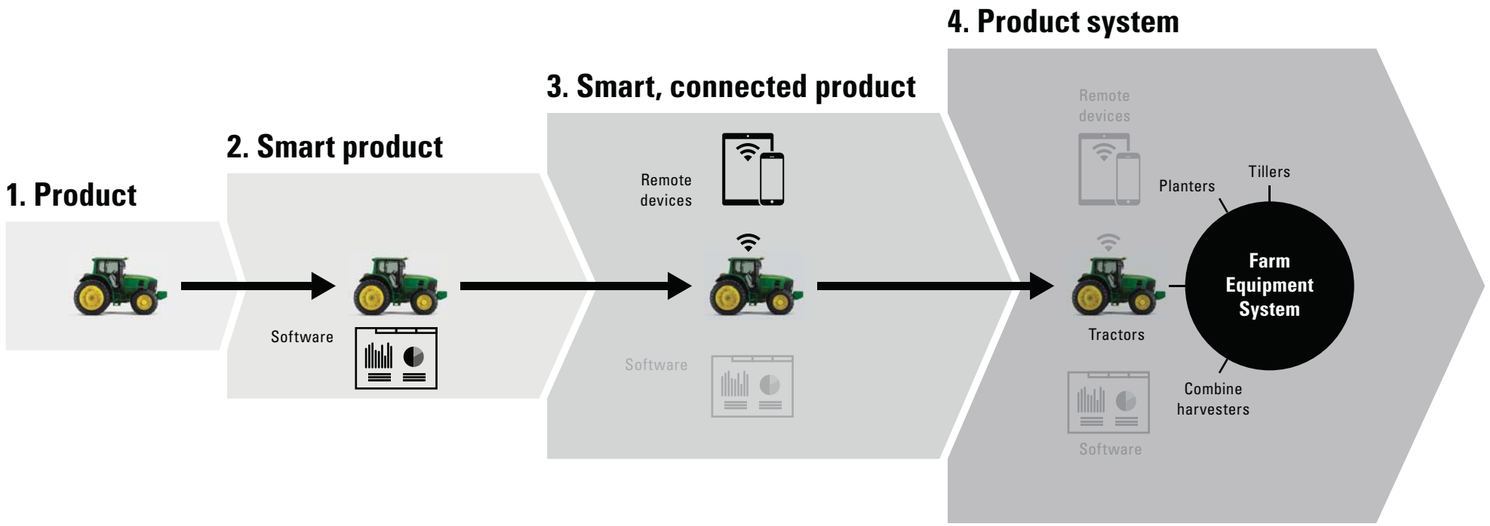


Sets of connected products form product systems.

Smart, Connected Product

+ other Smart, Connected Products

= Product System

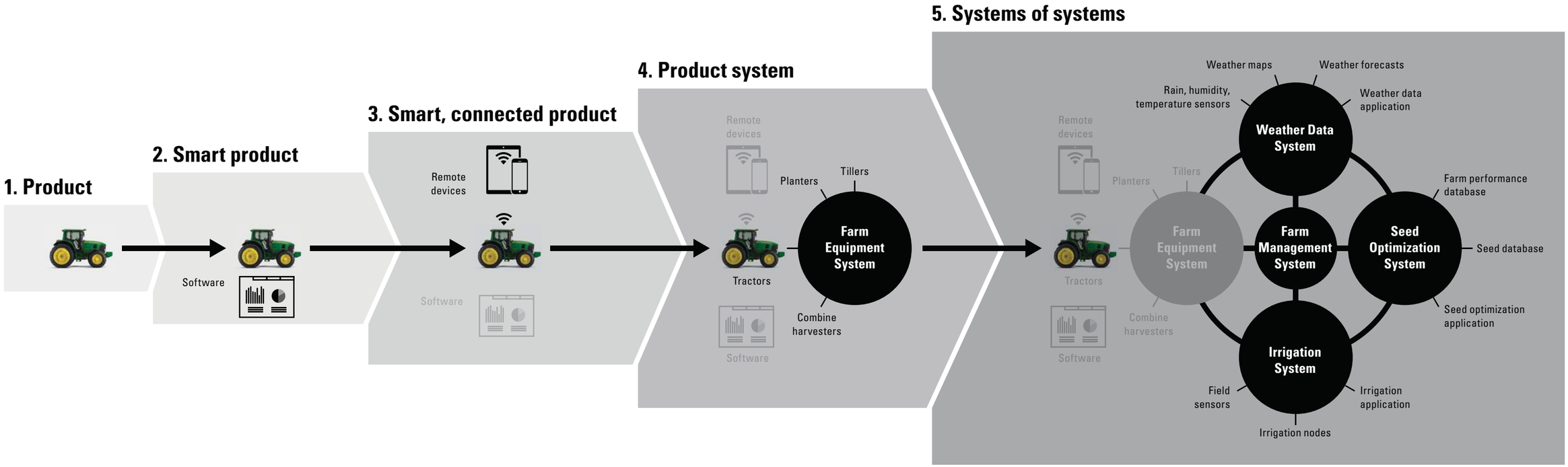


Systems connect to other systems, forming ecologies.

Product Systems

+ other Product Systems

= Product-Services Ecology

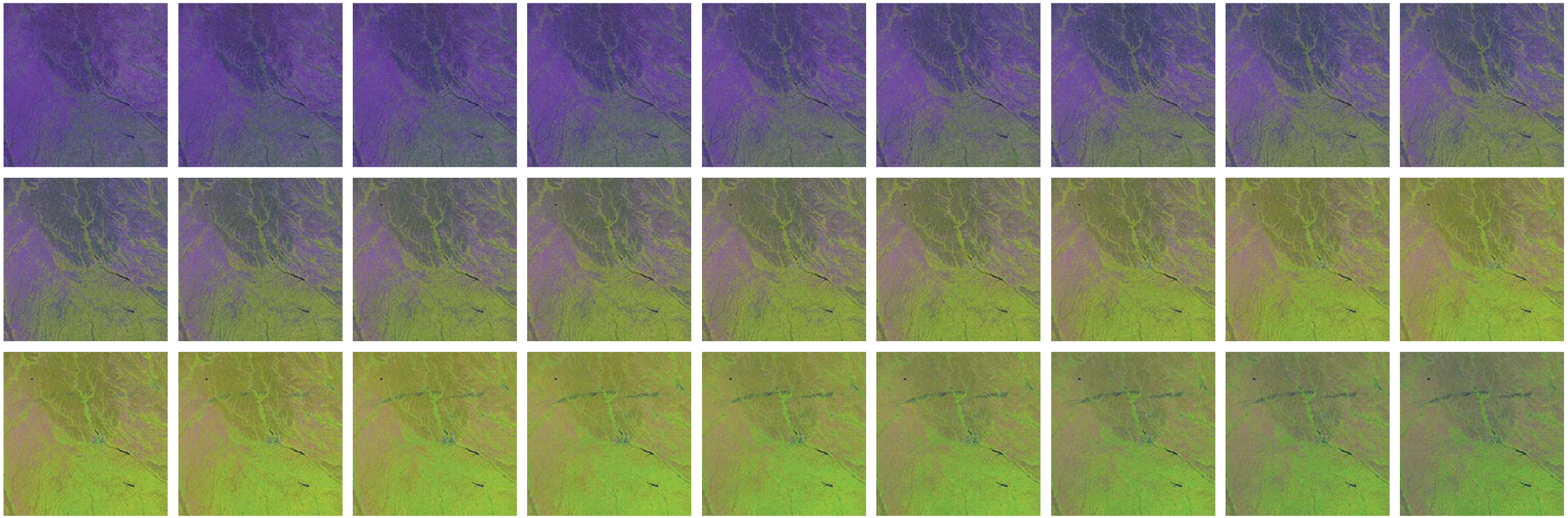


Farms are becoming automated factories.

Plants are attached to sensors, connected to networks, generating data.

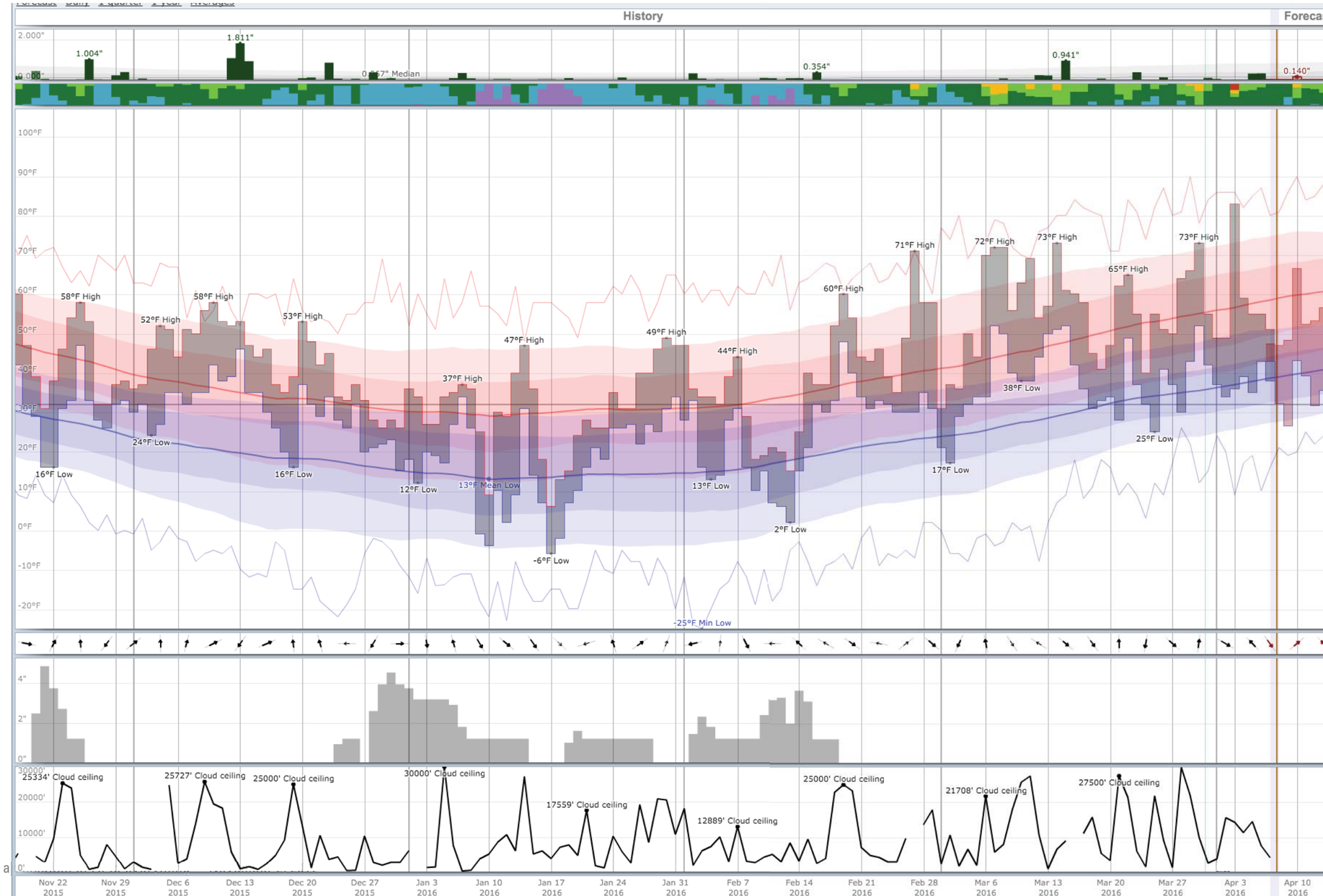
Macro view: processed satellite images of crop growth over time, e.g., central Iowa, March 29 to October 23, in 8 day increments.

Algorithms automatically align images,
remove clouds,
and detect vegetation.

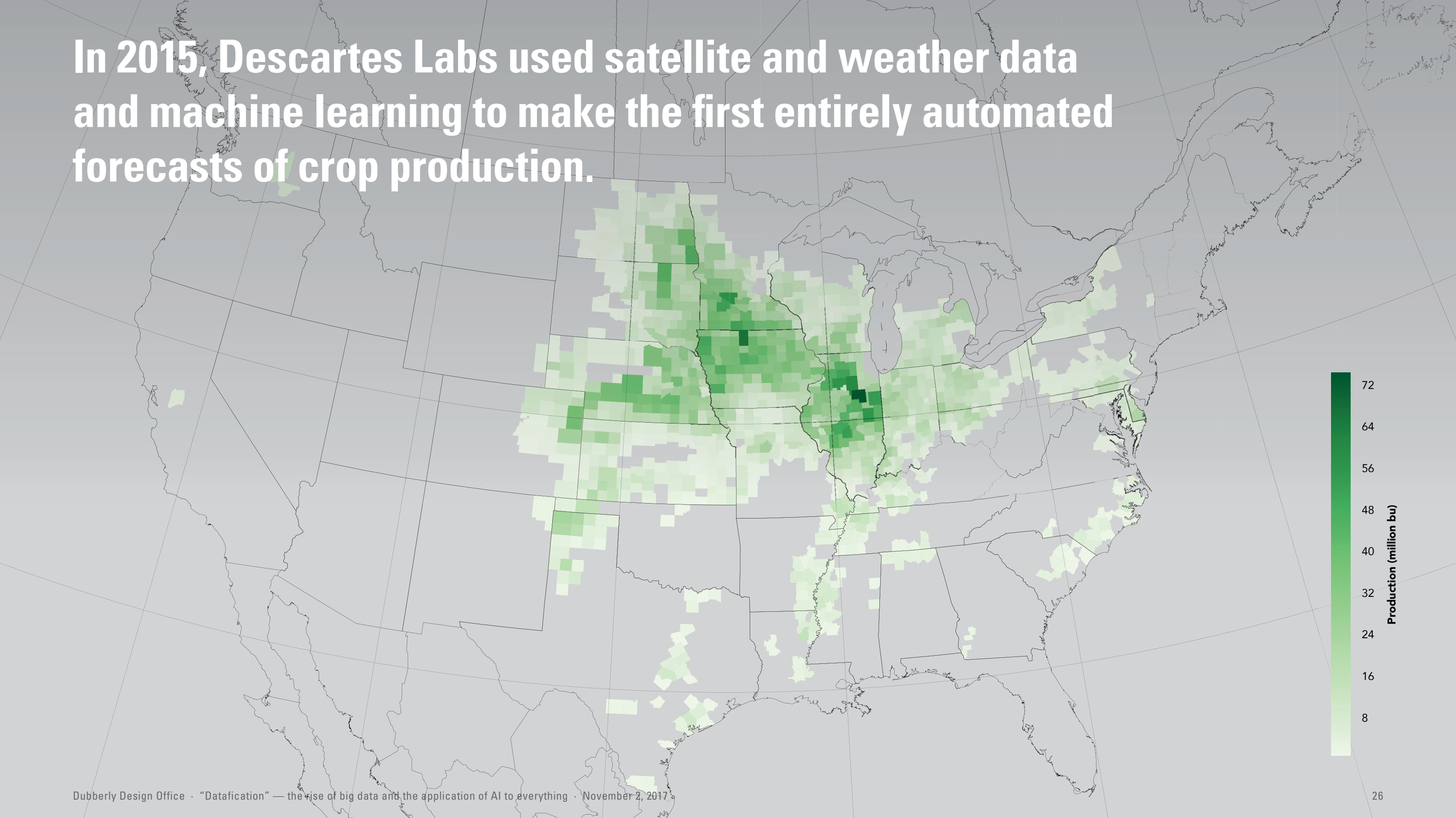


Daily weather data can augment machine learning.

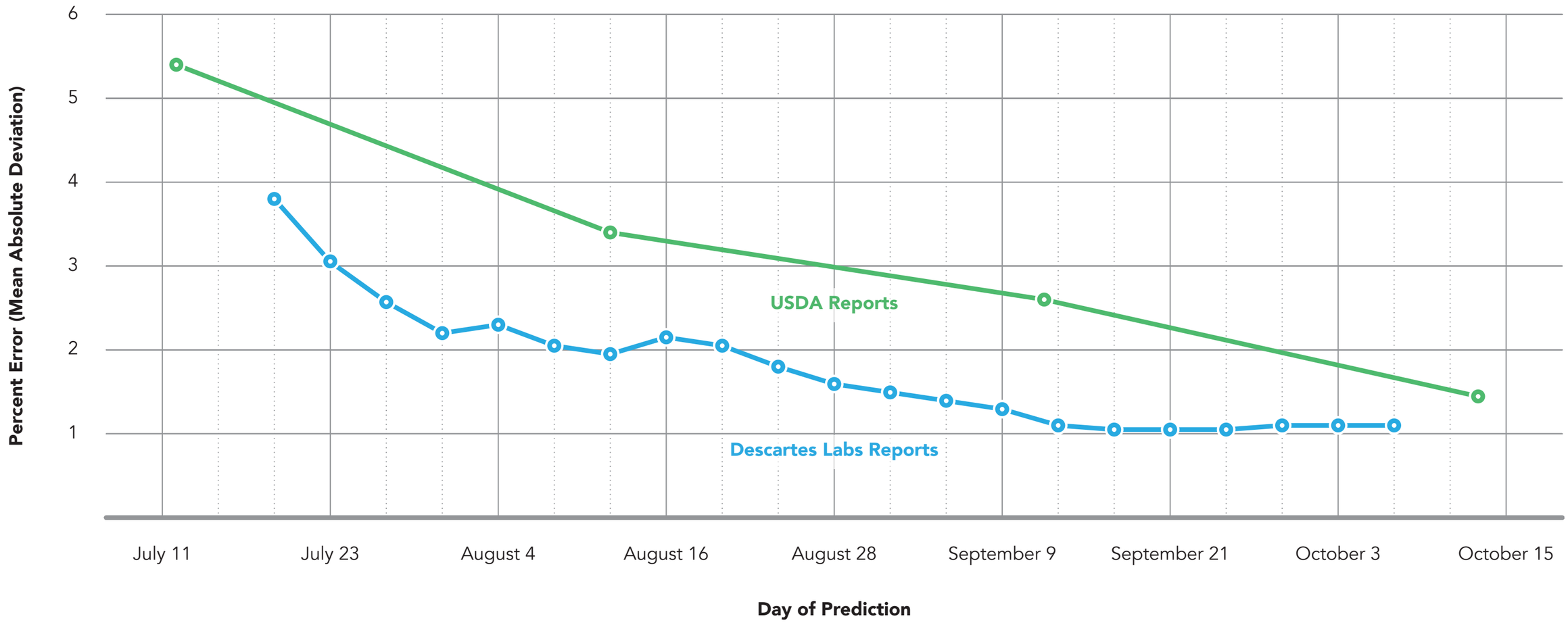
Precipitation,
temperature,
wind direction and speed,
snow cover,
and cloud cover
can aid forecasting.



In 2015, Descartes Labs used satellite and weather data and machine learning to make the first entirely automated forecasts of crop production.



Descartes Labs predicted US corn product — within about 1.9% of actual production later reported by USDA.



Descartes Labs is a signal of a massive change.

- Self-driving cars, trucks, and drones

- IBM Watson Health

- GE Predix and Siemens MindSphere

- Apple Siri, Viv (now Samsung), Amazon Alexa, Google Assistant, Facebook M, Microsoft Cortana

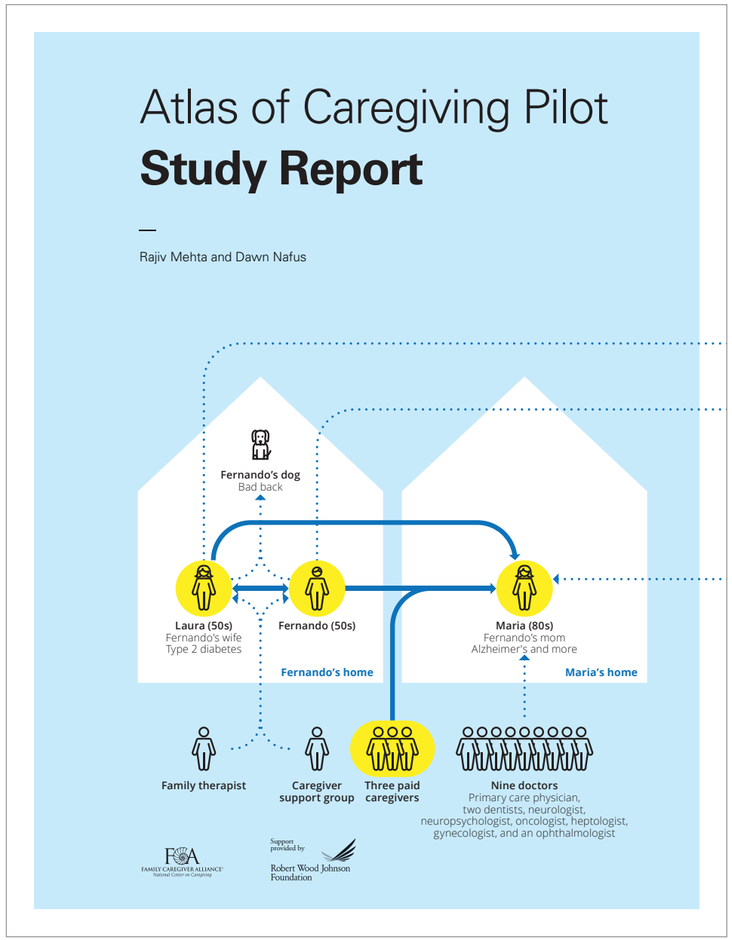
- FBI's Facial Analysis, Comparison, and Evaluation (FACE) Services has access to > 400 million photos.

Large, unique databases are inherently valuable.

data + algorithms = prediction

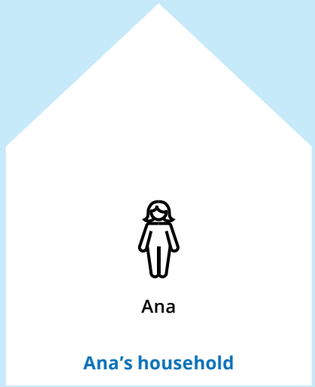
Another example

In 2015, Robert Wood Johnson Foundation funded a pilot study to look at new ways of measuring family caregiving.



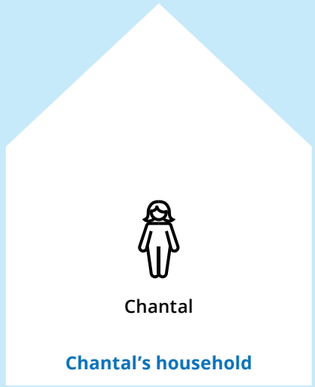
Robert Wood Johnson Foundation

We looked at 14 households, with 20 participants, with 21 chronic conditions.



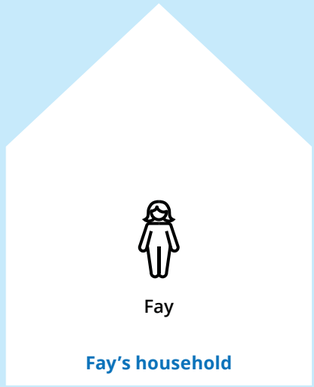
Ana's household

Ana (50s) has had **cystic fibrosis** since birth. She devotes several hours a day to care for her own condition. She also cares for her teenage son Albert, who has **depression**.



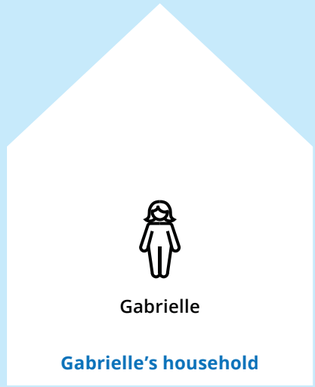
Chantal's household

Chantal (50s) has resigned work to care for her mother Debby (80s) who requires 24x7 care for **dementia**. Additional support comes from a paid home aide and other family members.



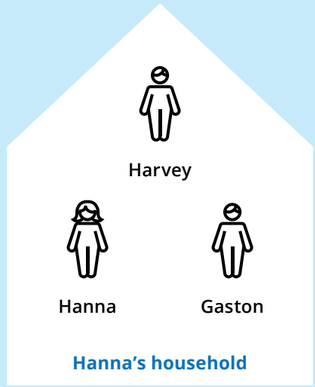
Fay's household

Only-child Fay (30s) cares for her mother Josephine (70s) who has **Alzheimer's**. With no one to help her, she has put PhD studies on hold to provide 24x7 care.



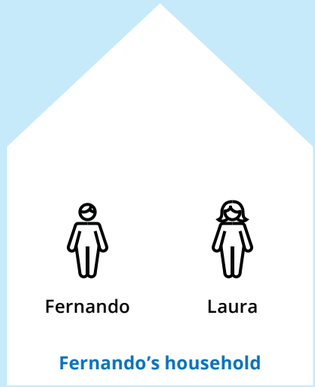
Gabrielle's household

Gabrielle (60s) is the primary caregiver of her mother Karen (101), who has **Alzheimer's**. Gabrielle also has health issues of her own and the sleepless nights and caregiving needs of her mother have taken a toll.



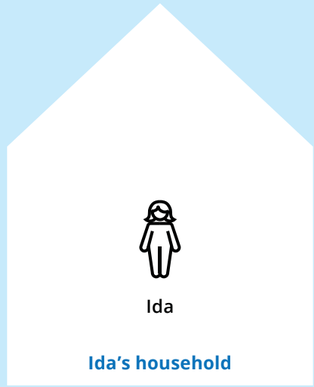
Hanna's household

Hanna and husband Gaston care for her brother Harvey, who has **epilepsy** and **pneumonia/sepsis**. Gaston also cares for his mother, while managing his own **chronic pain** and **edema**. Both Hanna and Gaston also work.



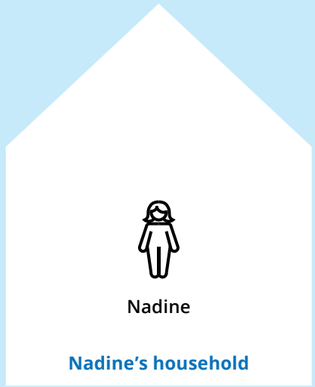
Fernando's household

Fernando and his wife Laura (50s) are the primary caregivers for Fernando's mother Maria (80s) who has **Alzheimer's** disease as well as other health conditions. Together, Fernando and Laura have built a care network to support Maria.



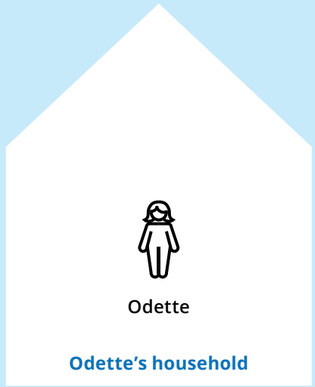
Ida's household

Ida (70s) cares for her husband Ian (70s) who has **Lewy Body Dementia** and **Dysautonomia**. They moved to San Francisco to be nearer to their children two years ago.



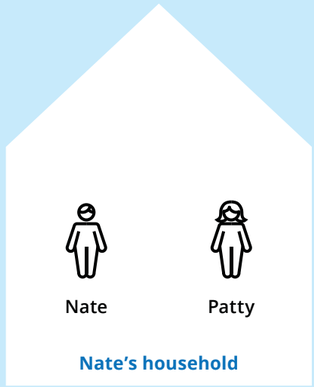
Nadine's household

Nadine (50s) lives with her husband Jerry and two teenage sons, Larry and Karl. Karl has **Type 1 Diabetes**. Nadine is his primary caregiver.



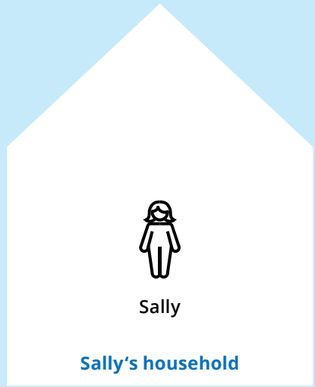
Odette's household

Odette (70s) and her husband Marco (70s) share their home with several other people: their son, son-in-law, and five tenants. Marco has **Parkinson's** disease. Odette is his primary caregiver, but several others are also involved.



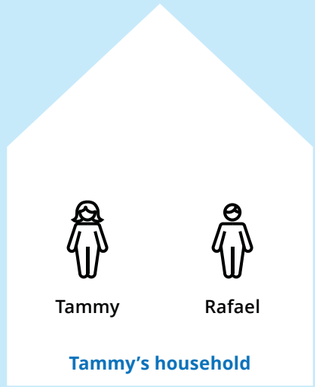
Nate's household

Nate and Patty, both in their 30s, care for each other. Patty has **multiple sclerosis** (MS) and Nate has **glioblastoma**, a terminal condition.



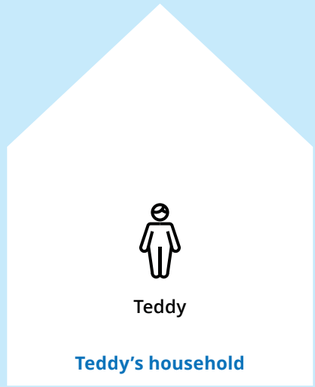
Sally's household

Sally (50s) cares for her son Pablo (20s), who has behavioral and emotional difficulties stemming from **XYY Chromosome Disorder**.



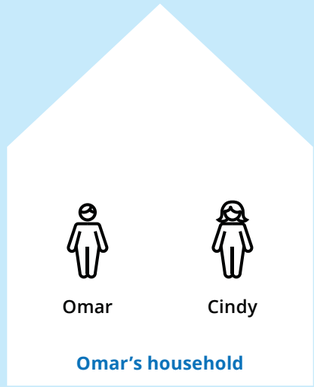
Tammy's household

Tammy (40s) and her husband Rafael (50s) care for their pre-teen children, Wanda and Sam. Wanda has severe **epilepsy** and **cerebral palsy**. She requires 24x7 care. Sam has severe **autism** and also requires a lot of care.



Teddy's household

Teddy (40s) and his wife are the primary caregivers for their two young sons, Van and Walter. Van has **Aspergers (ADHD type)** as well as **encopresis**, and Walter has **cyclical vomiting syndrome**.



Omar's household

Omar (40s) and his separated wife Cindy (40s) share a home with their young son Bob, who has **Aspergers**.

Using 12 sensors

Measuring 16 factors

Photographs
GPS and timestamp

Narrative Clip

Blood Volume Pulse
Calculated to derive heart rate

Acceleration X
Acceleration Y
Acceleration Z
Calculated to derive average motion

Electrodermal Activity
(EDA)

Skin Temperature

Empatica E4

Presence
At home or away

SmartSense Presence Sensor

Motion
In which room and how active

SmartSense Motion Sensors

Humidity
Temperature
Barometric pressure
C02
Noise
Indoor unit

Netatmo Indoor Weather Station

Humidity
Temperature
Outdoor unit

Netatmo Outdoor Weather Station

Over an average of 24 hours

Resulting in 5 GB of data—just from the watch.

The BVP sensor is running at 64 Hz.
That means it makes a reading every
1/64th of a second.
60 seconds comprise a minute;
60 minutes comprise an hour; and
36 hours is the maximum duration of
one of our study sessions.

In other words, one study session
comprises 2,160 minutes,
and just one of the sensors
is collecting 3,840 samples per minute.

That's 8,294,400 samples collected over
the course of one 36-hour session.

8,294,000	samples for BVP (at 64 Hz)
4,147,200	samples for X axis acceleration (at 32 Hz)
4,147,200	samples for Y axis acceleration (at 32 Hz)
4,147,200	samples for Z axis acceleration (at 32 Hz)
518,000	samples for EDA (at 4 Hz)
518,000	samples for skin temperature (at 4 Hz)
21,772,800	samples of raw data for one participant
×19	participants
413,683,200	or nearly half a billion data points

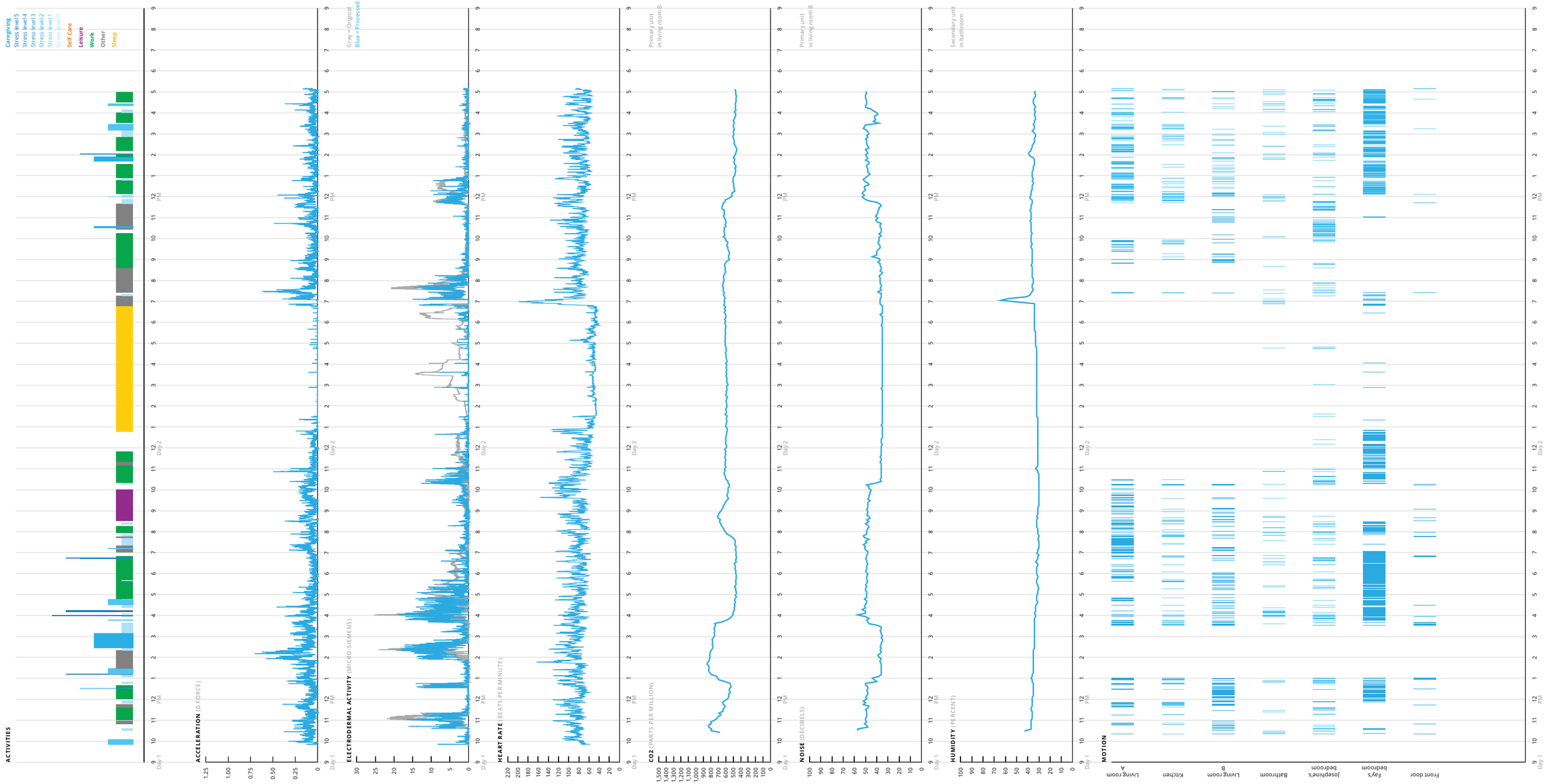
Photo log for Fay ×20 additional participants

Black squares replace recognizable faces to ensure privacy.

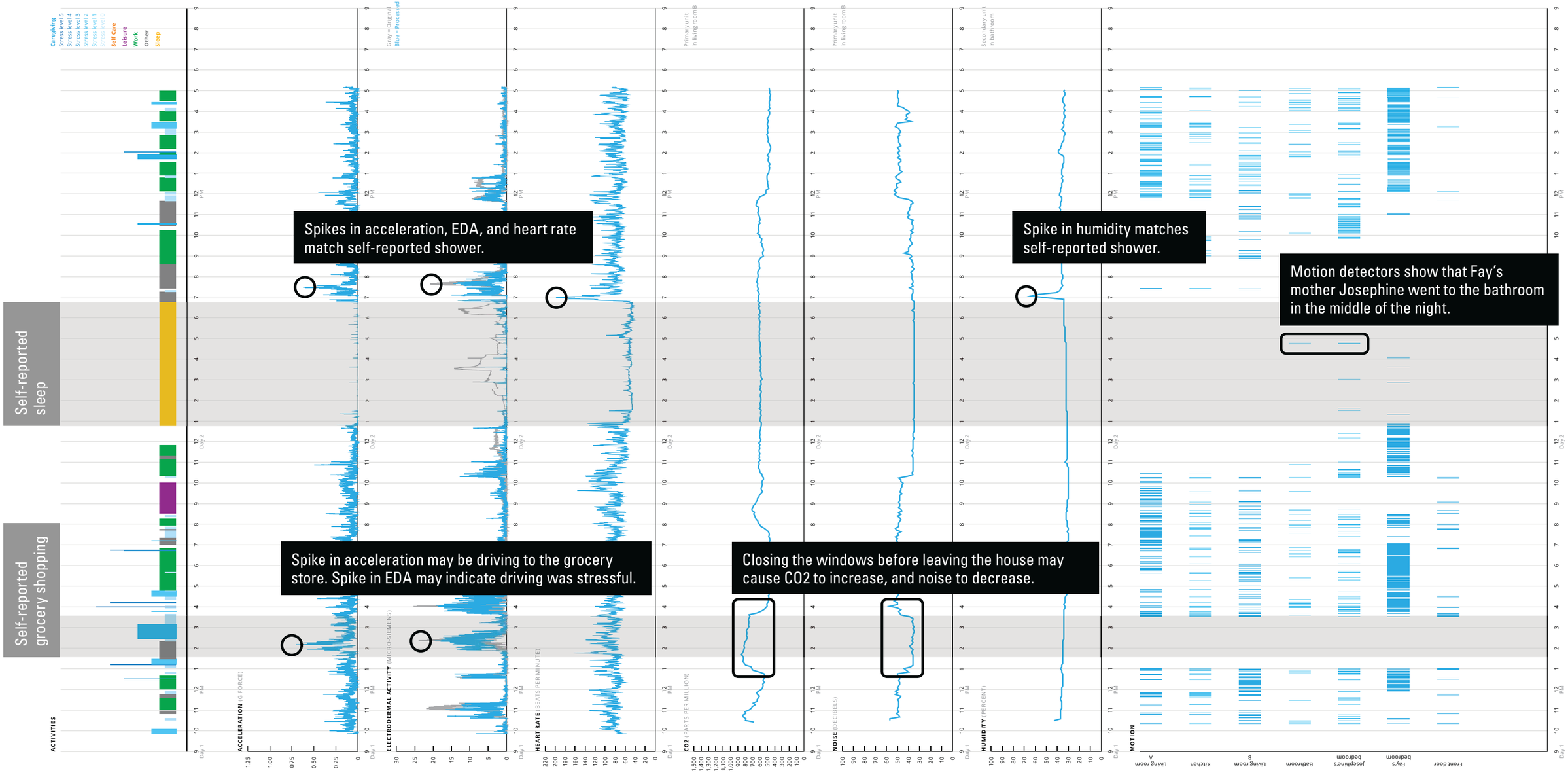
Gray squares indicate when participants turned the camera off.

White squares indicate the start and stop of the study.

Summary diagram for Fay ×18 additional participants

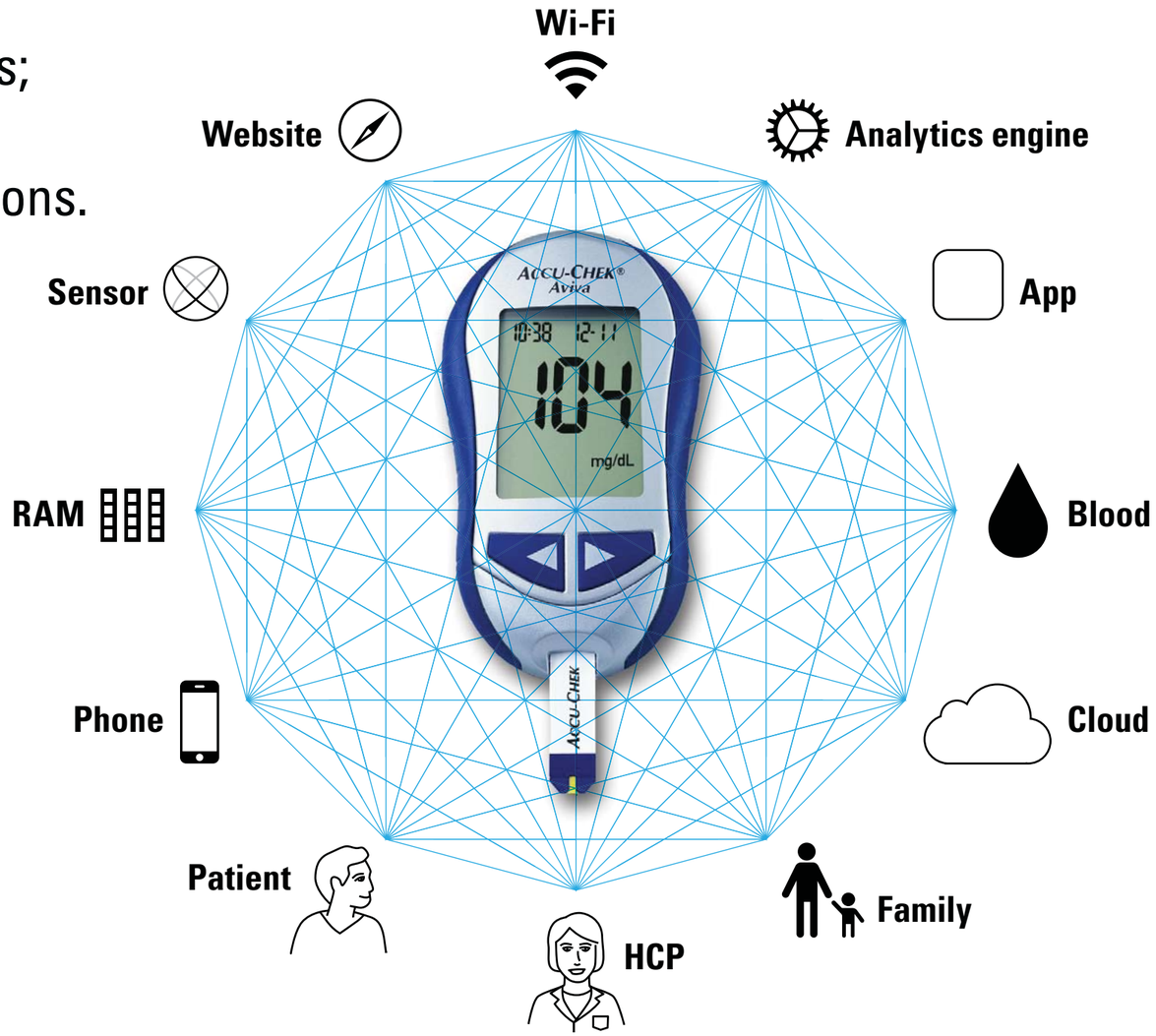


Analyzing Fay's summary diagram for insights—**morning**

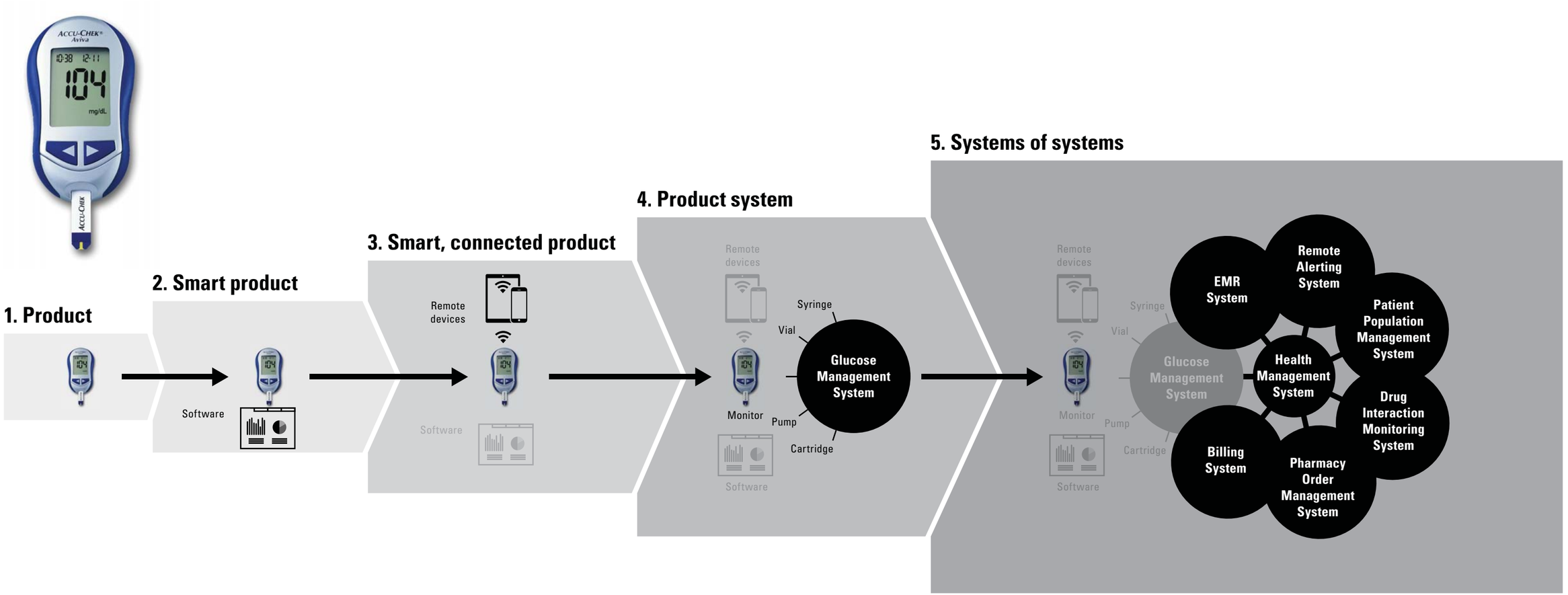


Medical devices are connecting to the cloud, too,
including pacemaker-defibrillators, auto-injectors, and glucometers.

Smart, connected glucometers
can alert family and HCPs about extreme lows;
they also collect data automatically
and help users see trends and make correlations.



In the future, medical products will no longer stand alone. Increasingly, they will exist in complex networked service webs.



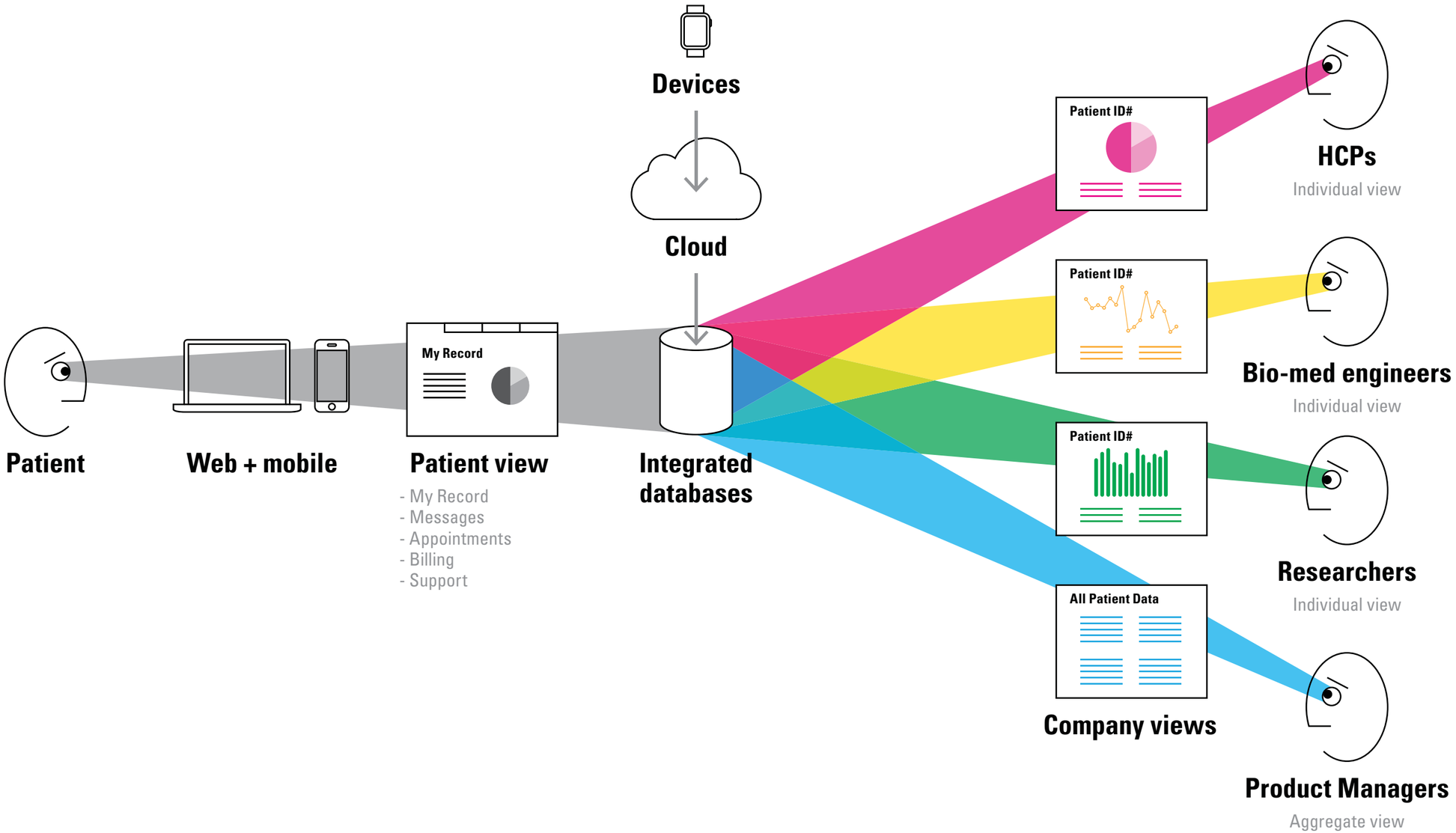
Glucometer + processor = computer that can run apps; e.g., bolus calculator, calorie estimator and tracker.

When a glucometer connects to a smart-phone, cost can come down, because the meter can build on the phone's processor and display. Plus data can be shared with family and HCPs.

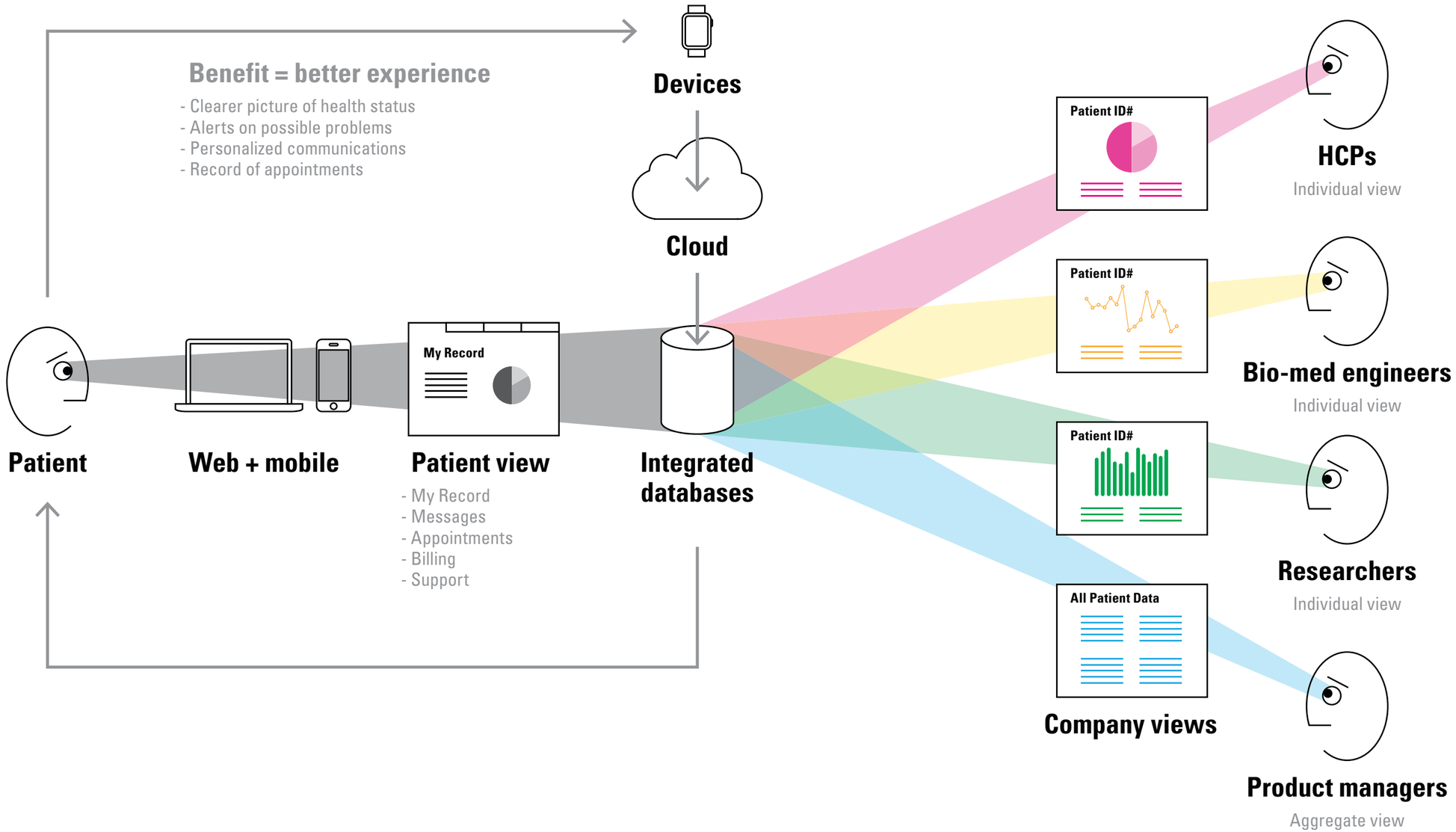
A near continuous glucose monitor can be coupled with an insulin pump, forming a glucose management system.

The glucose management system can connect with many other systems, such as EMRs, remote alerting, patient population management, drug interaction monitoring, pharmacy order management, and billing.

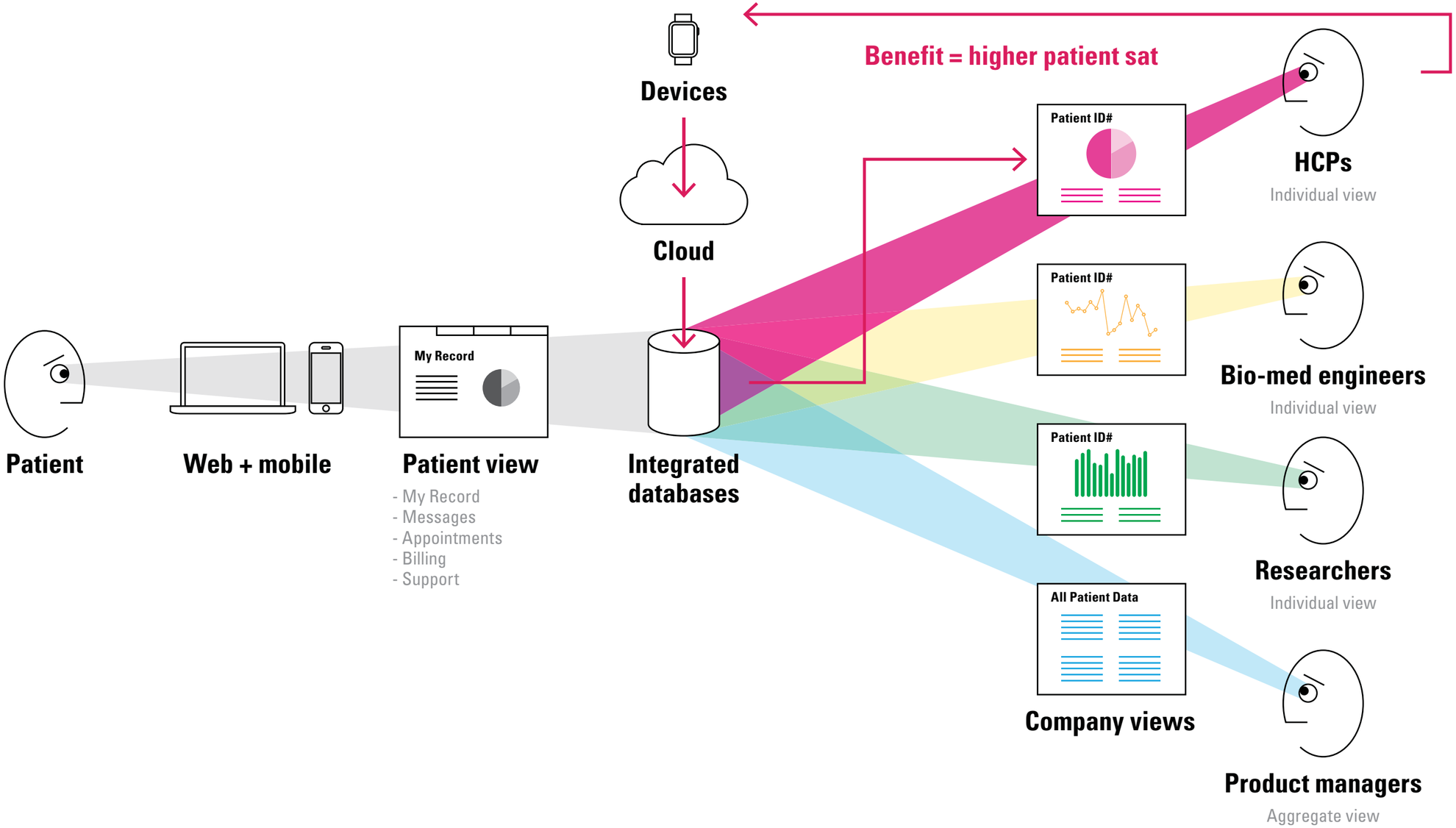
Unified patient and device data will afford useful views to many constituents.



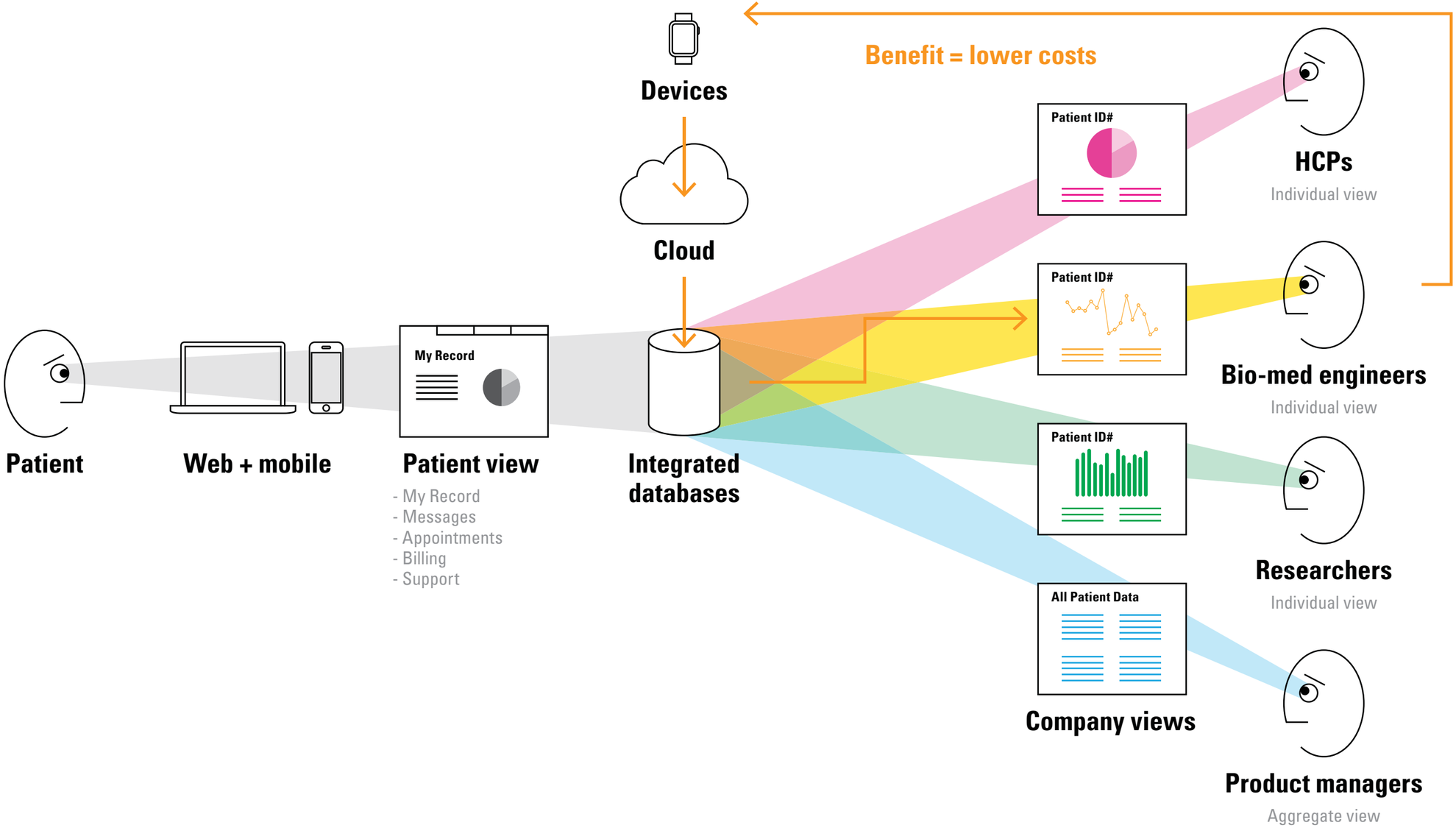
Patients can know much more about what's happening and can share information with family, friends, and HCPs.



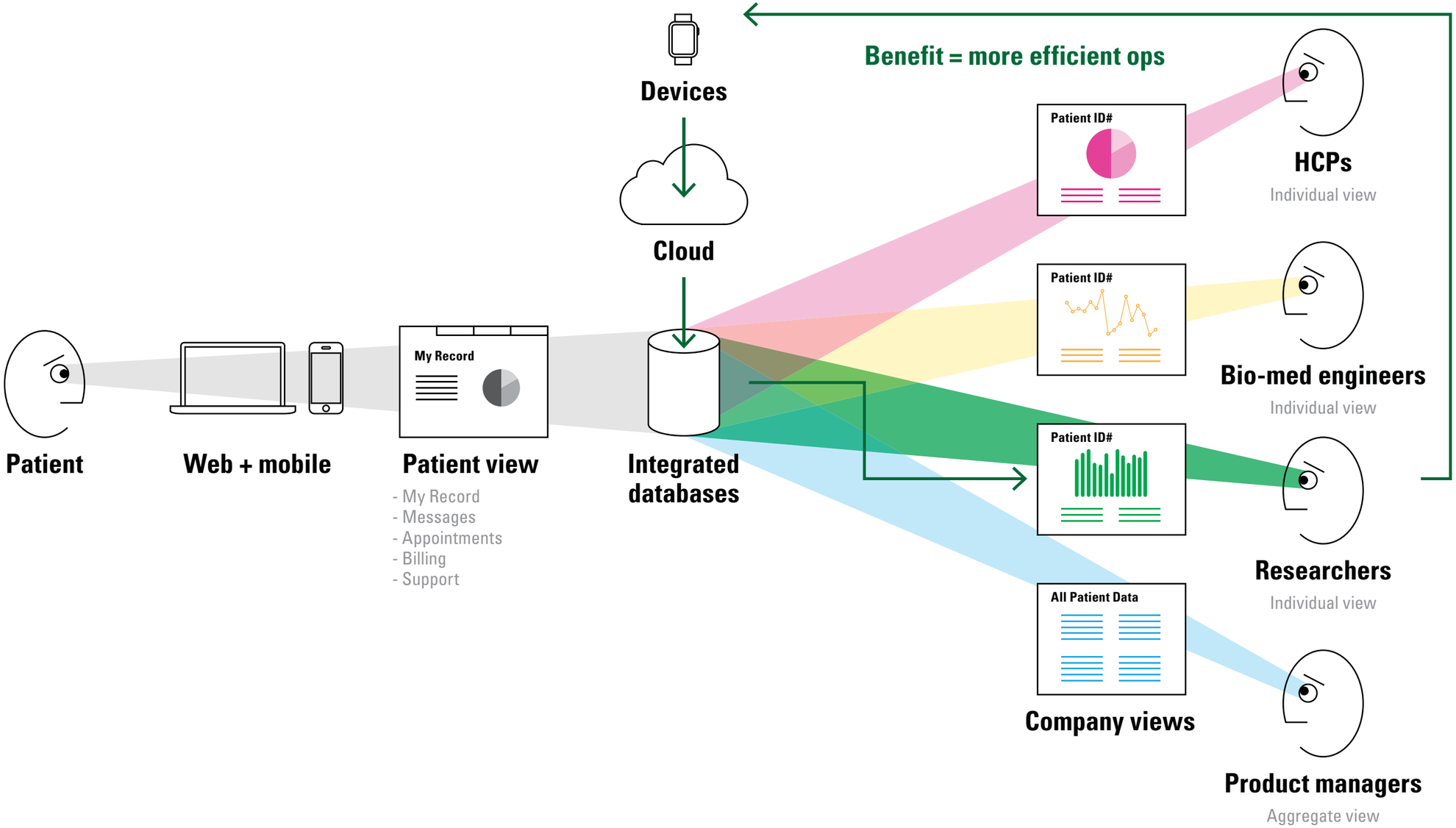
HCPs can receive a more holistic view of each patient and can manage groups of patients more efficiently.



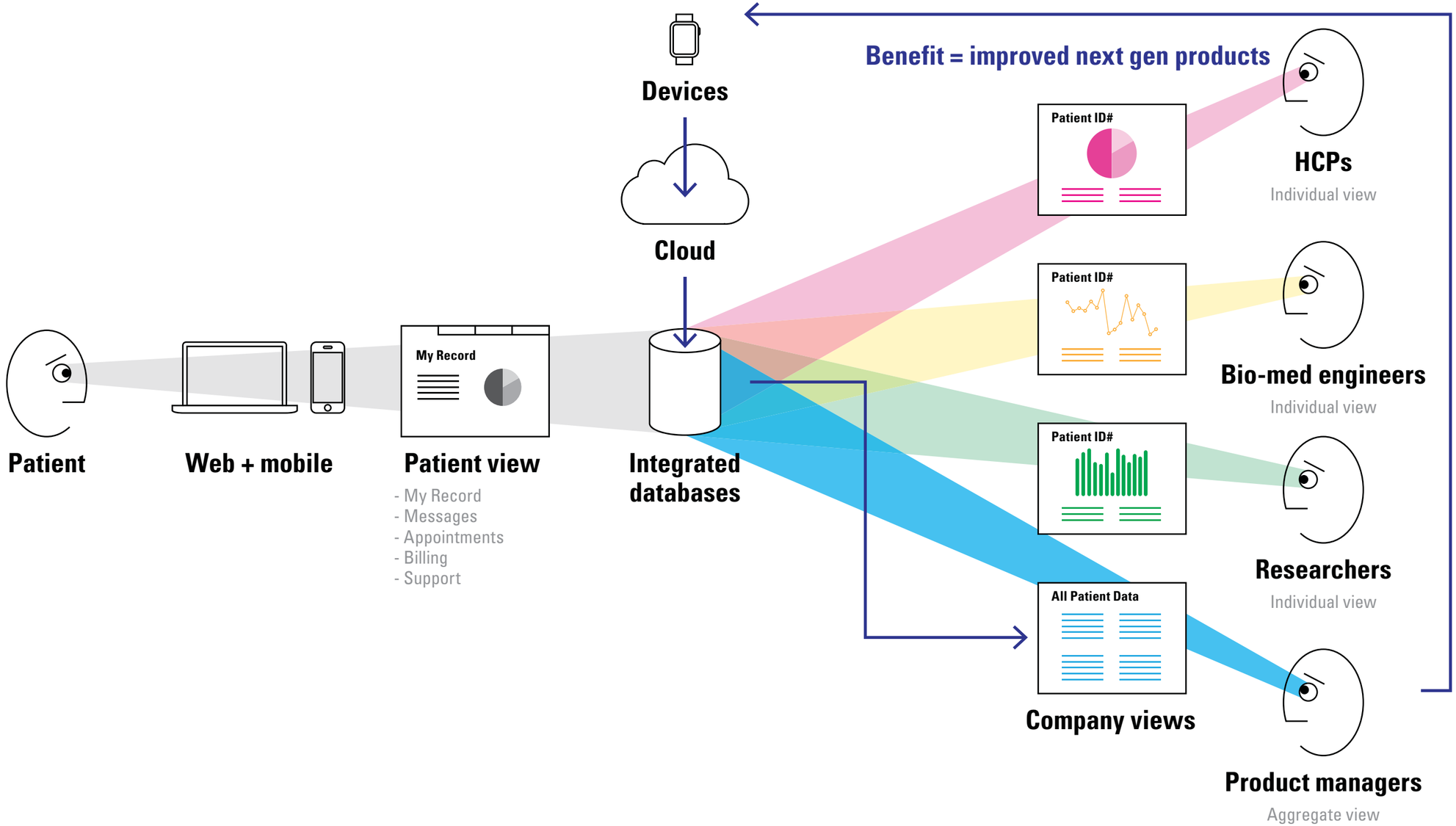
Bio-med engineers can better manage equipment, improve service, and reduce support costs.



Researchers can learn from aggregate data, to improve procedures and care-facility operations.



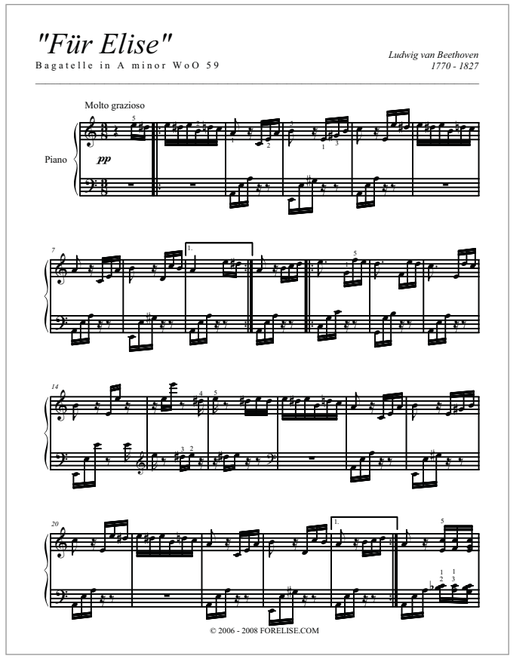
Product managers can get detailed usage data, to improve next generation products.

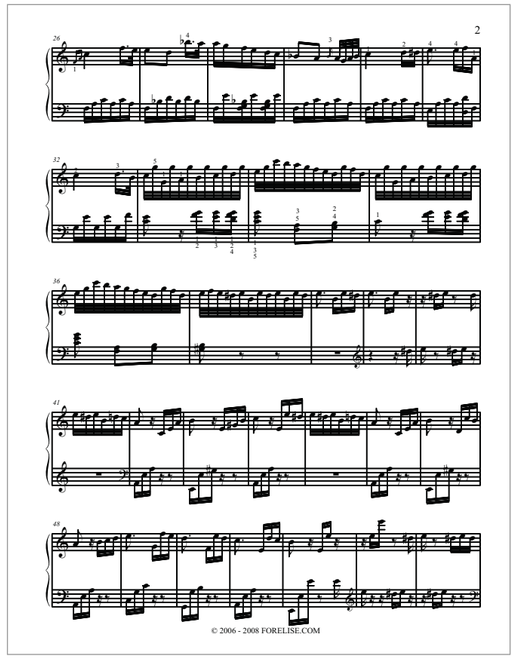


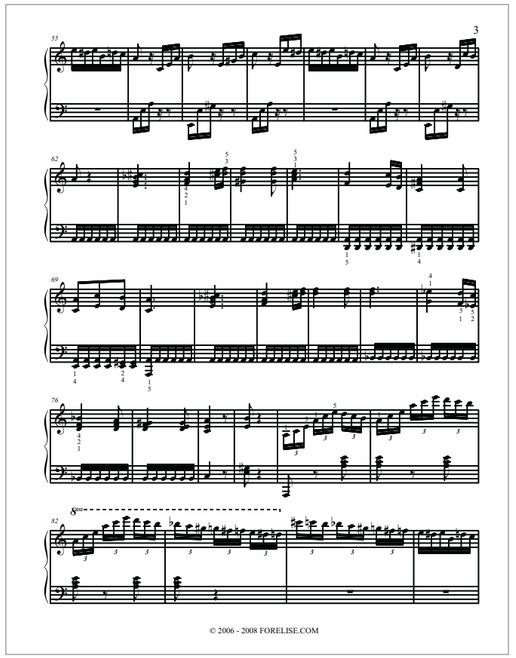
A final example

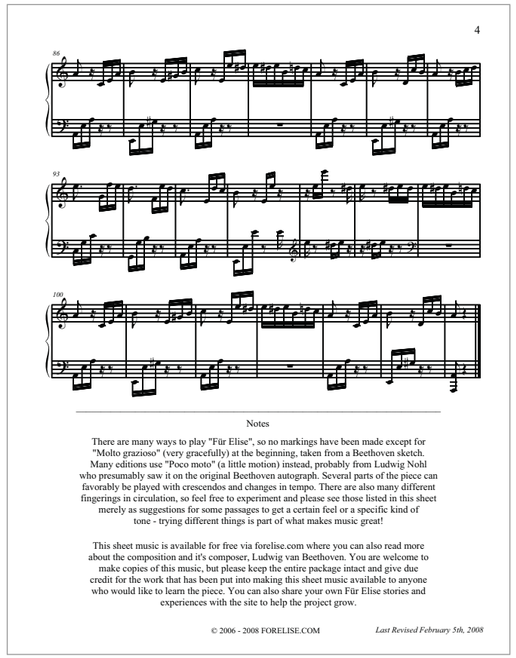
How do you measure “quality” in playing a piano?

An experiment was created to track the hand movement during multiple piano performances of Beethoven’s *Für Elise*.

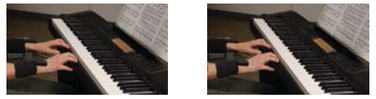
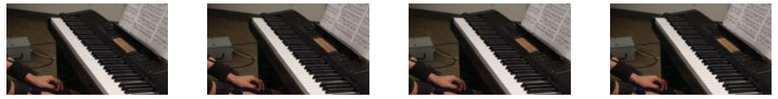
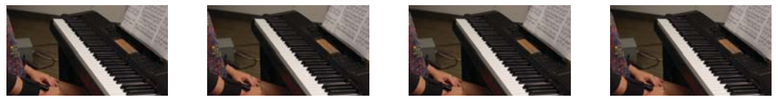
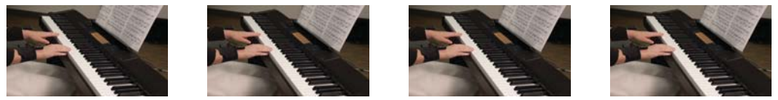






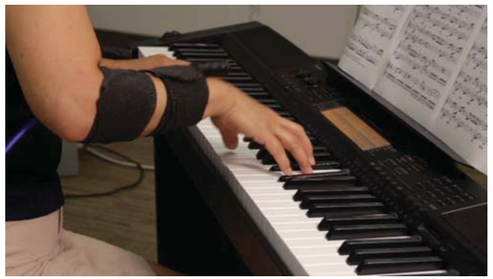


Twenty-five performances by seven performers were tracked and recorded.

- Grace 3 takes 
- Jamie 2 takes 
- Jiarong 4 takes 
- Katie 4 takes 
- Kelsie 4 takes 
- Sachiko 3 takes 
- ShanShan 5 takes 

Hypothesis: advanced players move their wrists to a greater degree.

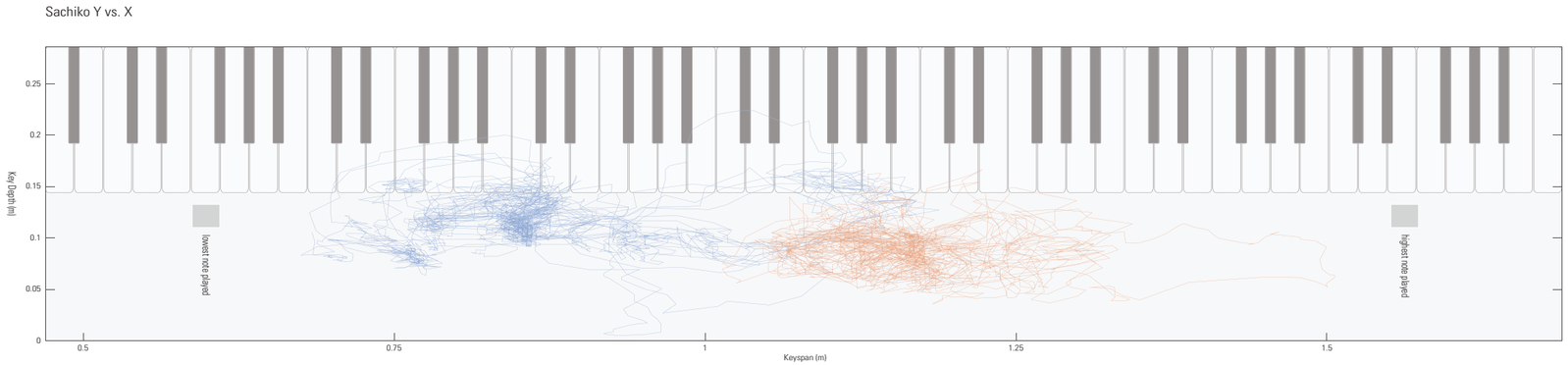
Listening and watching videos of sample performances by Jamie and Sachiko gave a clue.

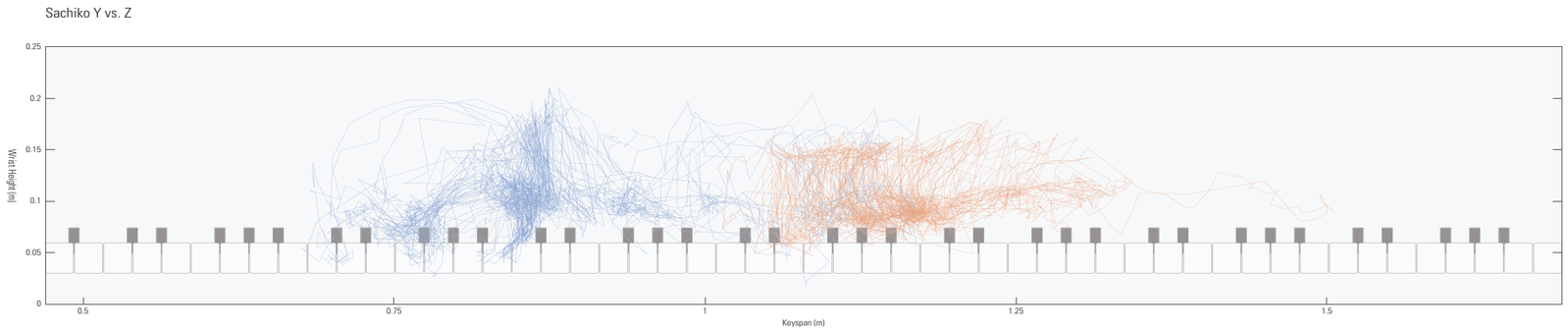


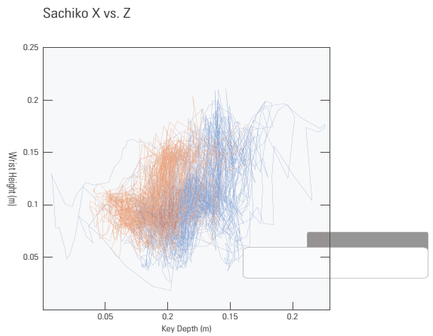
Sachiko’s performance was clearly better. Her wrists moved up and down as she played the piece, while Jamie’s wrists were relatively flat.

The orthographic projections are more revealing.

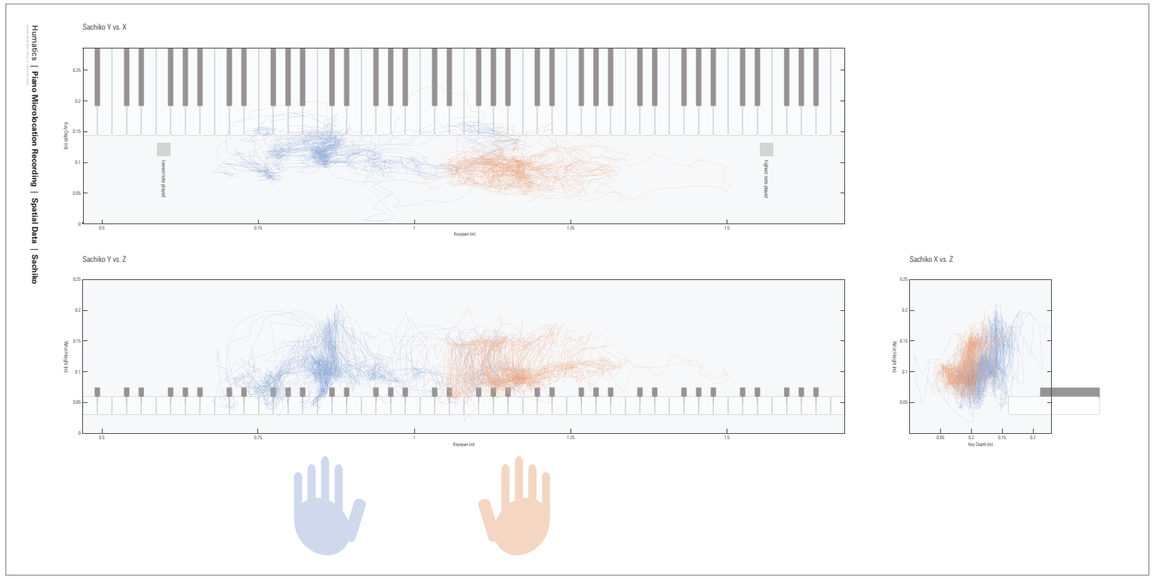
Humatics | Piano Microlocation Recording | Spatial Data | Sachiko



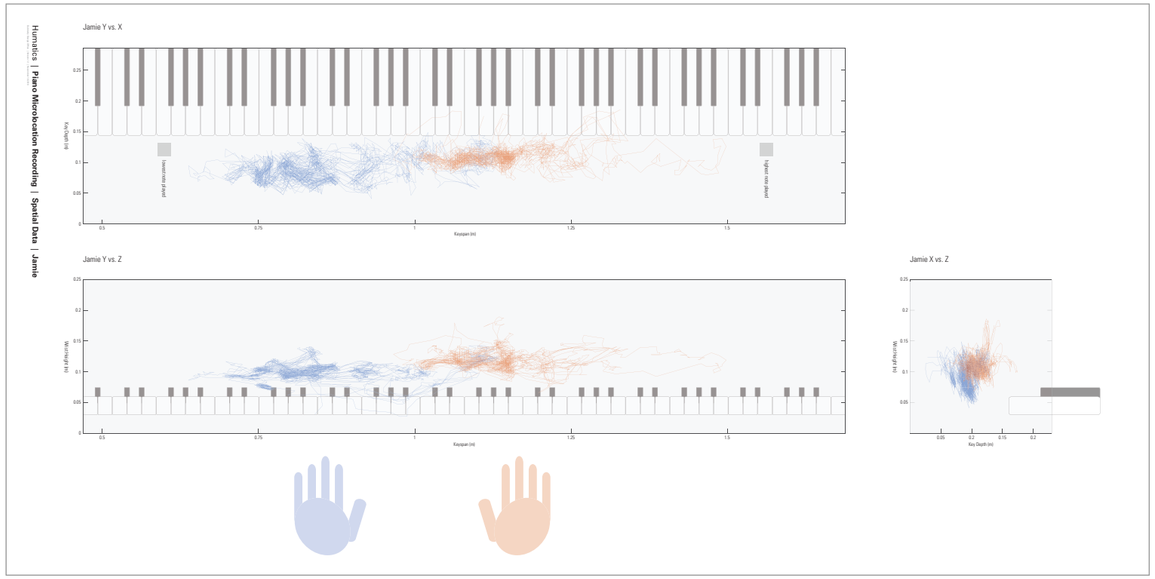




Comparisons show differences in movement in each dimension.

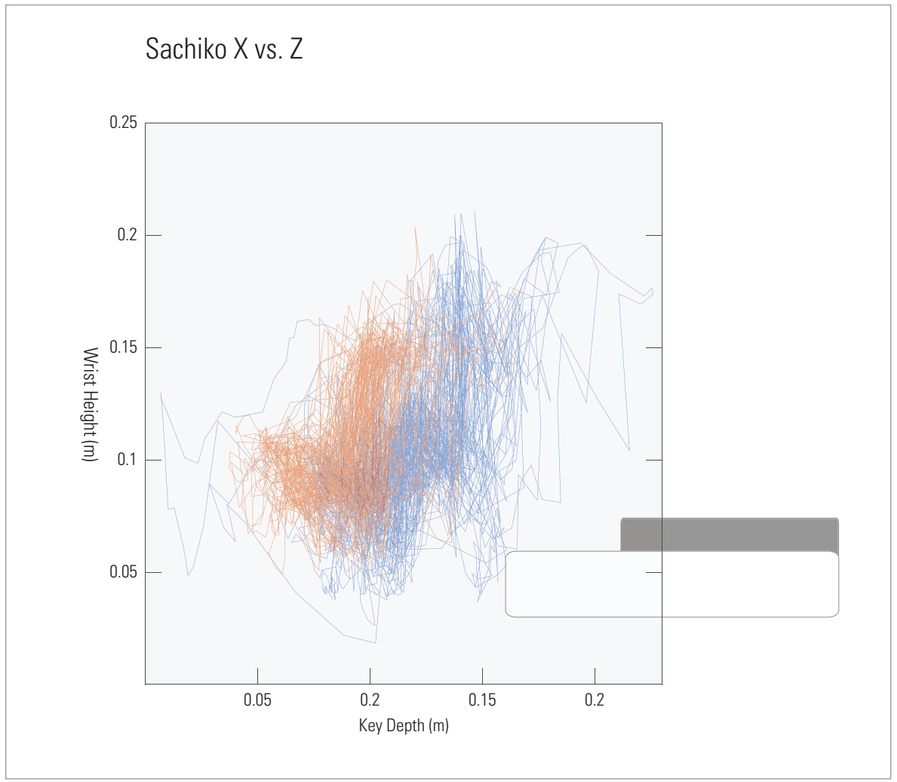


Sachiko

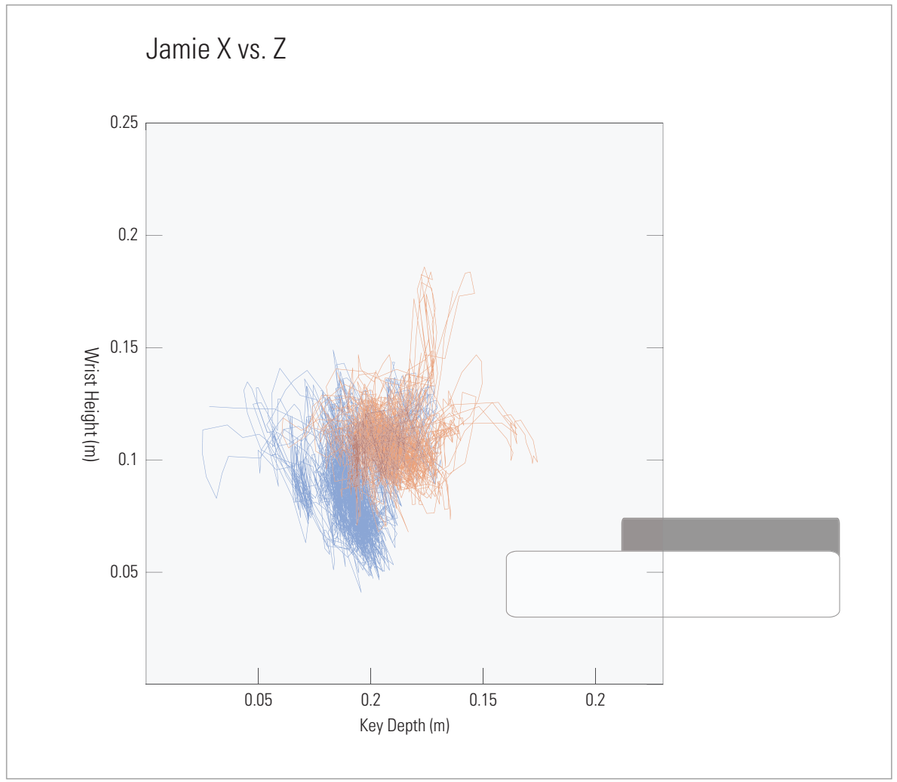


Jamie

The X vs Z plots are the most revealing.



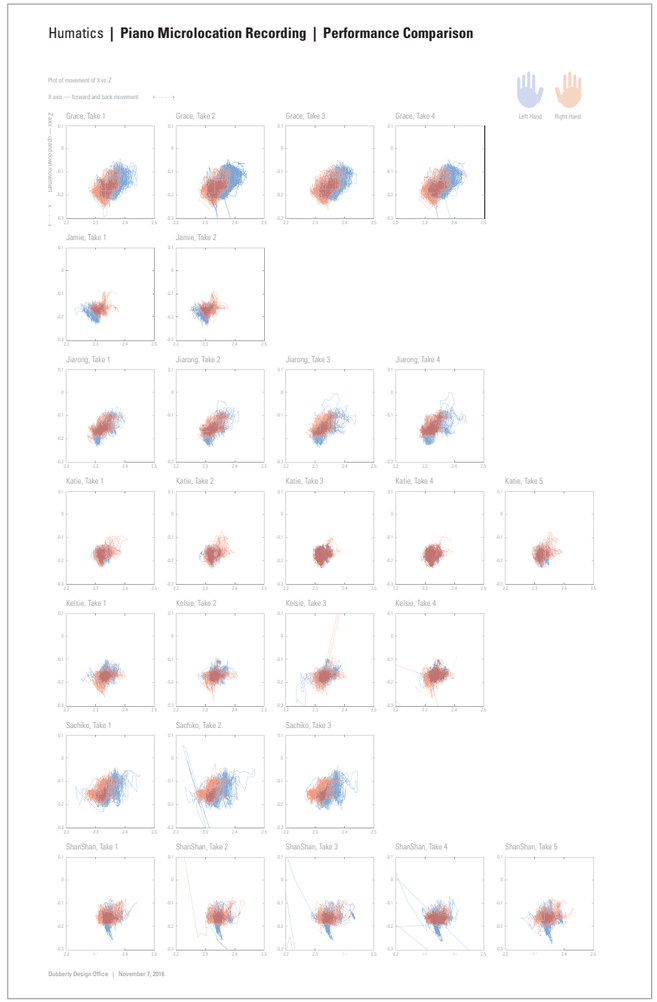
Sachiko



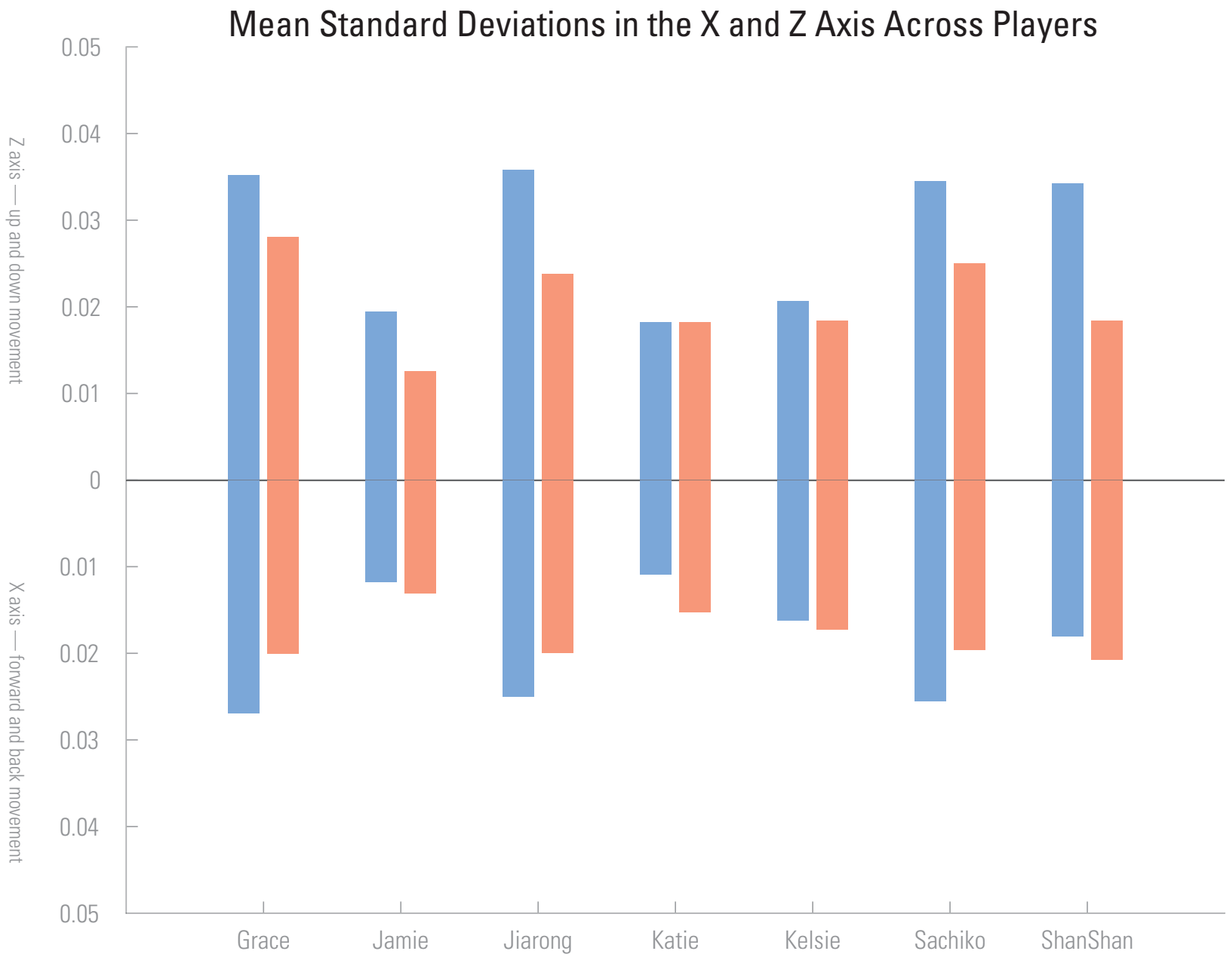
Jamie

X vs Z plots for all performances—
the differences are obvious.

- Grace
- Jamie
- Jiarong
- Katie
- Kelsie
- Sachiko
- ShanShan



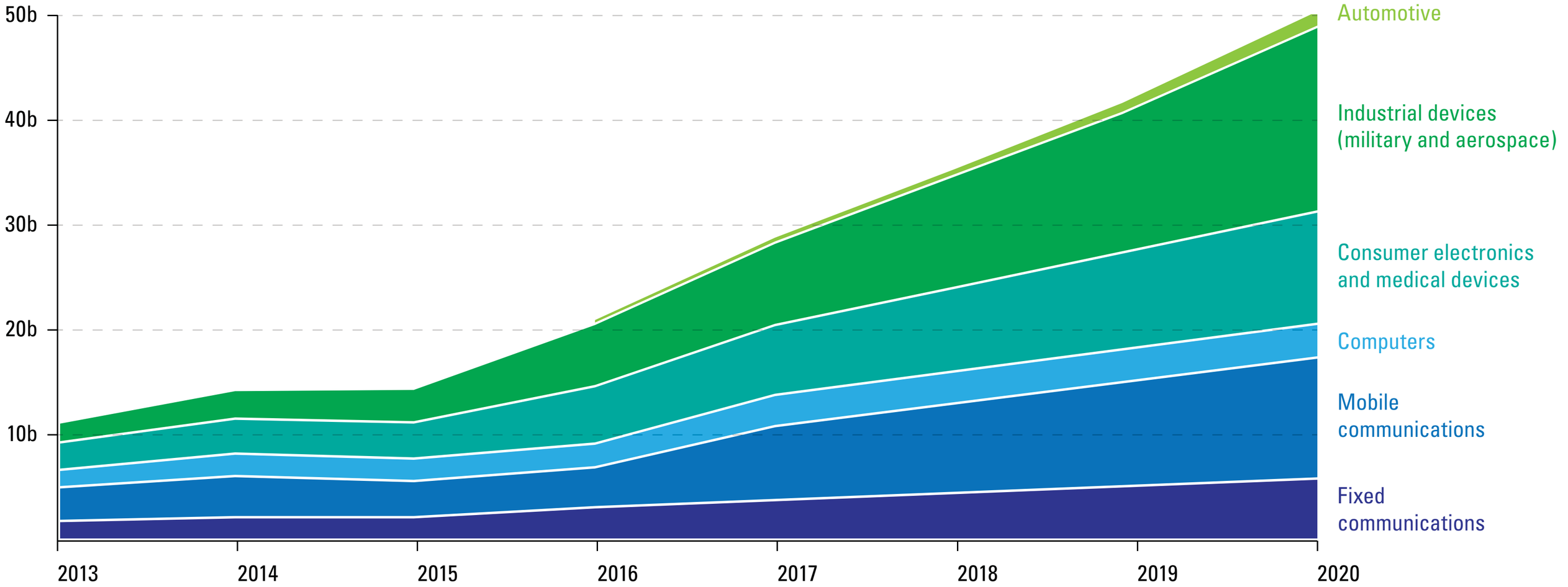
Calculating standard deviation shows a clear pattern.



The bar graphs show greater movement (particularly in the left hand) for Grace, Jiarong, Sachiko, and ShanShan—indicating they are advanced performers.

What does this mean?

By 2020, ~50 billion devices will be connected to the Internet; today, ~7 billion computers and tablets are connected.



Sources: The Economist and Cisco

Sensors will be everywhere — for example, Google + Levi's connected denim smart jacket

Jacquard Woven Gesture Sensor

Jacquard Tag

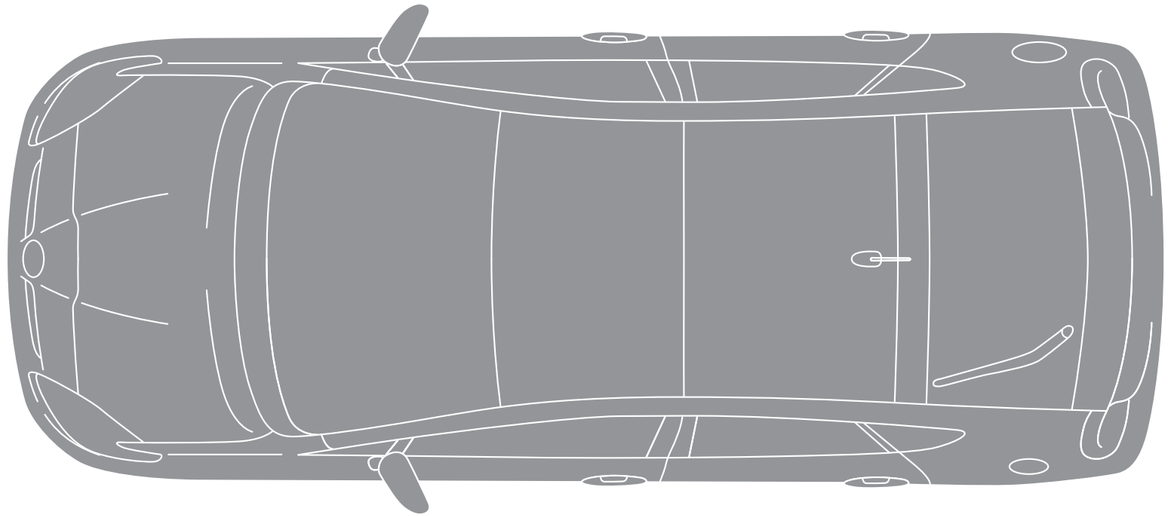
Jacquard App

Jacquard Services

Today's average car has:

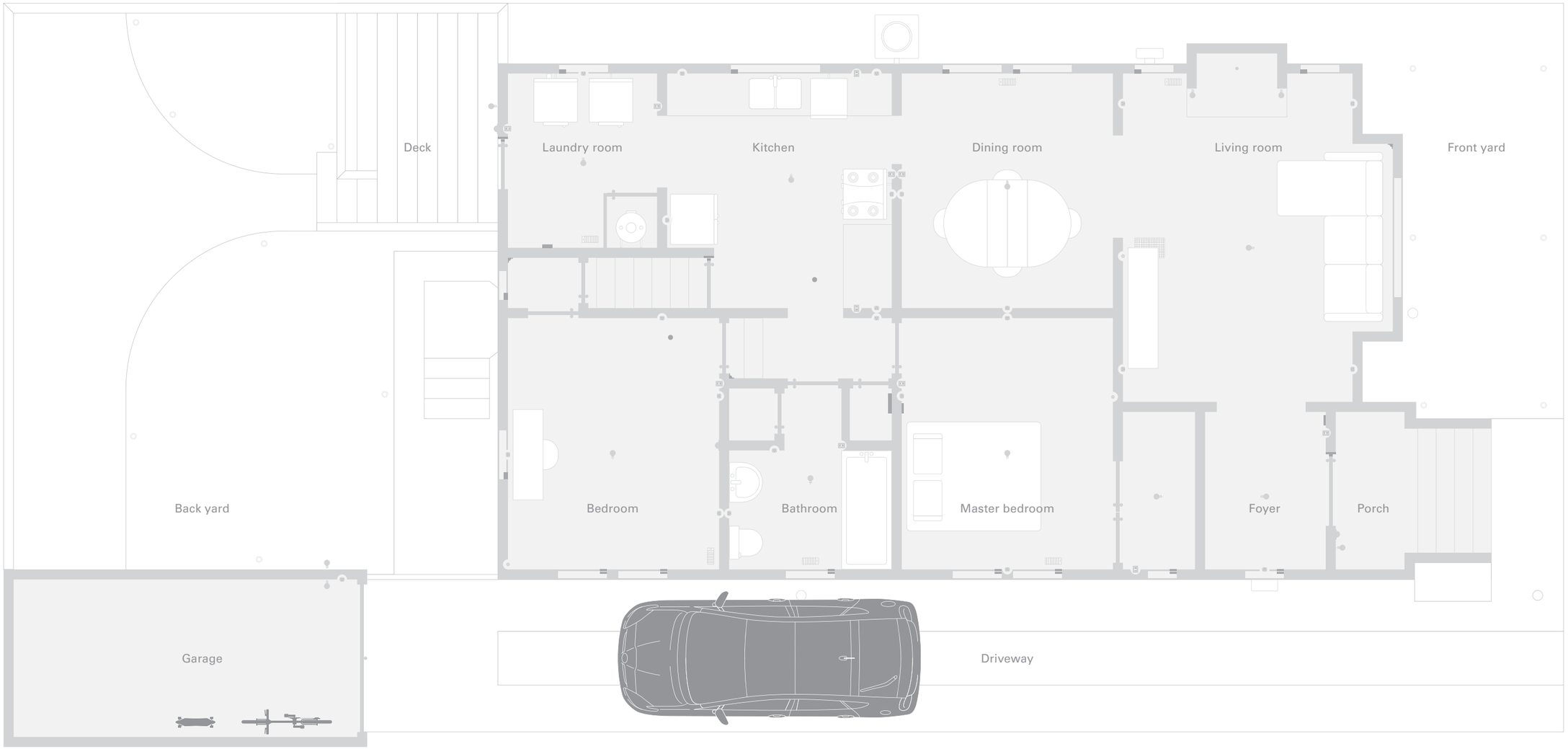
-1 engine
~7 small motors
(windows, wipers, fans)
~30 micro processors
(up to 100 for luxury cars) [1]
~60-100 sensors
(growing to 200 by 2020) [2]
~100 million lines of code
(up from 2 million lines in a generation) [3]

And it produces
"terabytes of data per car per day" [4]

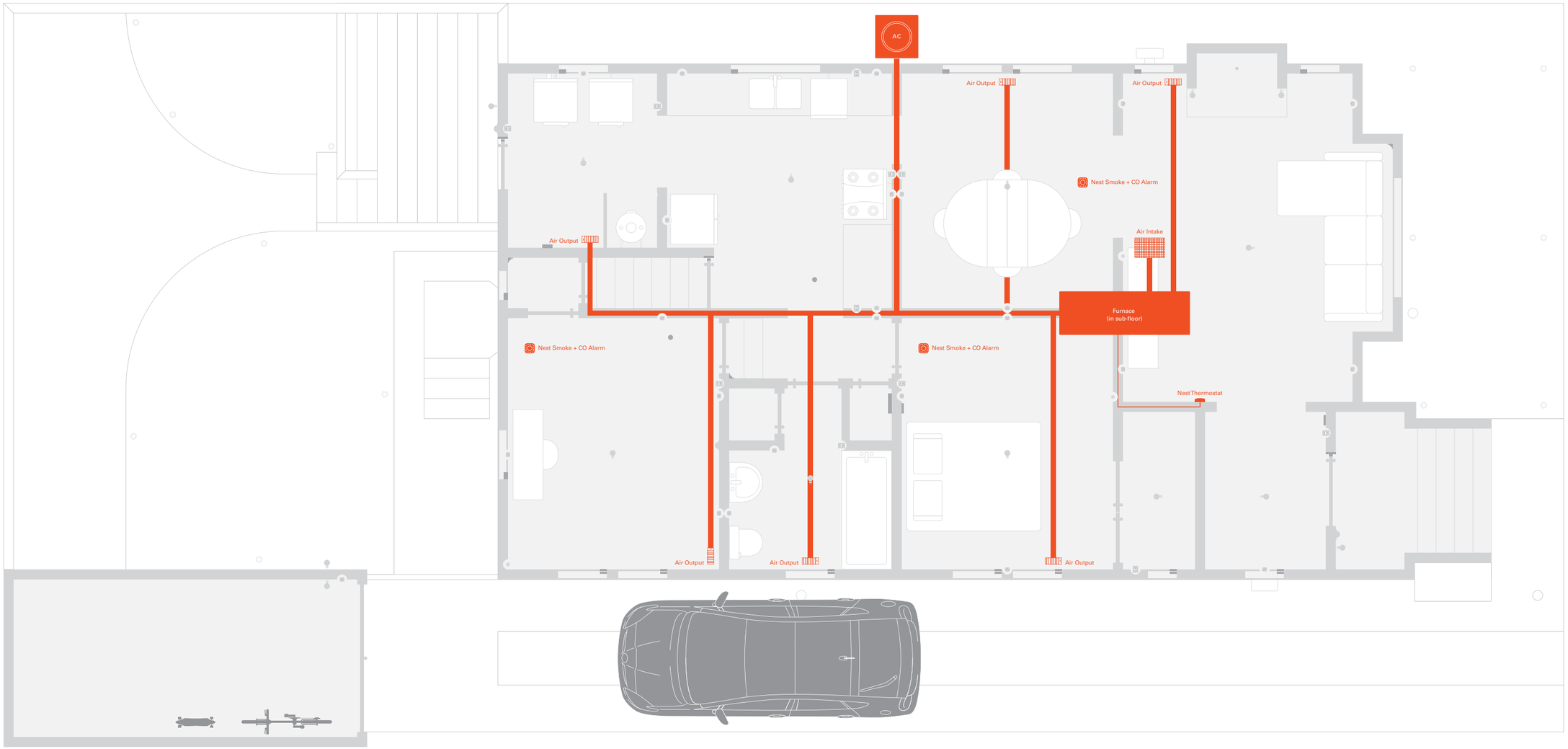


Sources:
[1] http://www.nytimes.com/2010/02/05/technology/05electronics.html
[2] http://www.automotivesensors2015.com/
[3] https://leithporsche.com/news/What+Makes+the+2017+Porsche+Panamera+Different3F+Computer+Code/7659/
[4] Parrish Hanna, Global Director of HMI at Ford (personal communications)

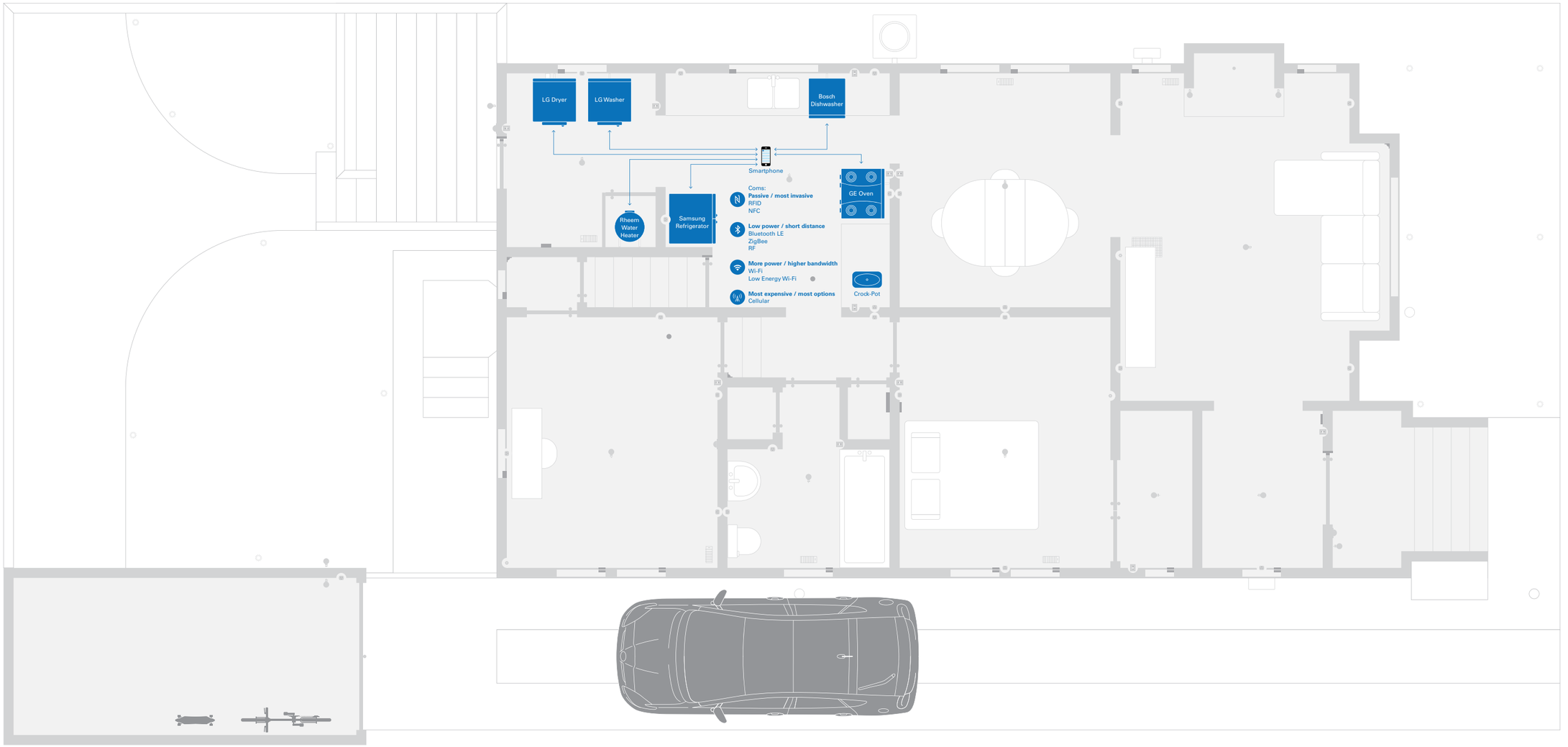
IoT devices in homes will produce and collect massive amounts of data.



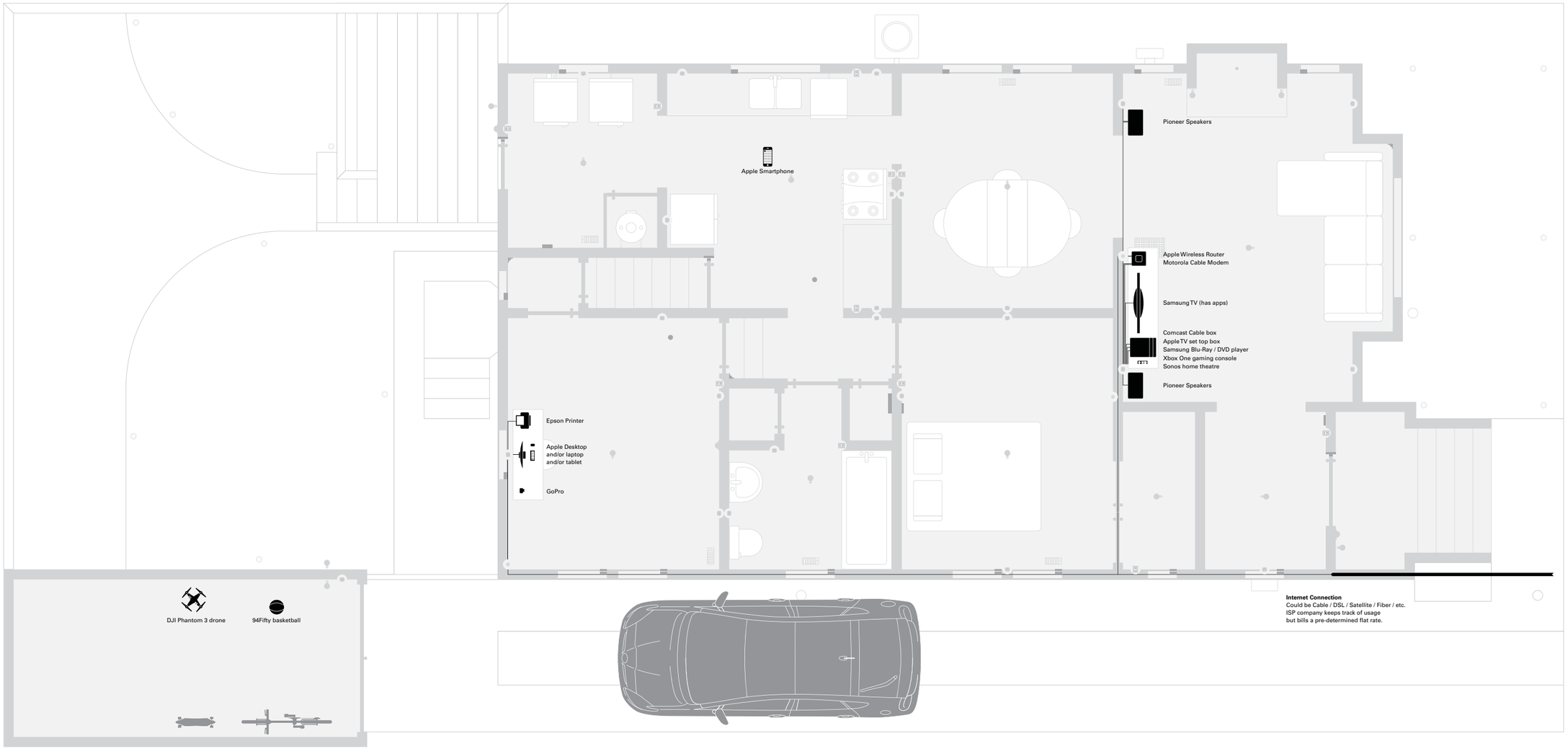
IoT devices in homes will produce and collect massive amounts of data:
HVAC



IoT devices in homes will produce and collect massive amounts of data:
Appliances

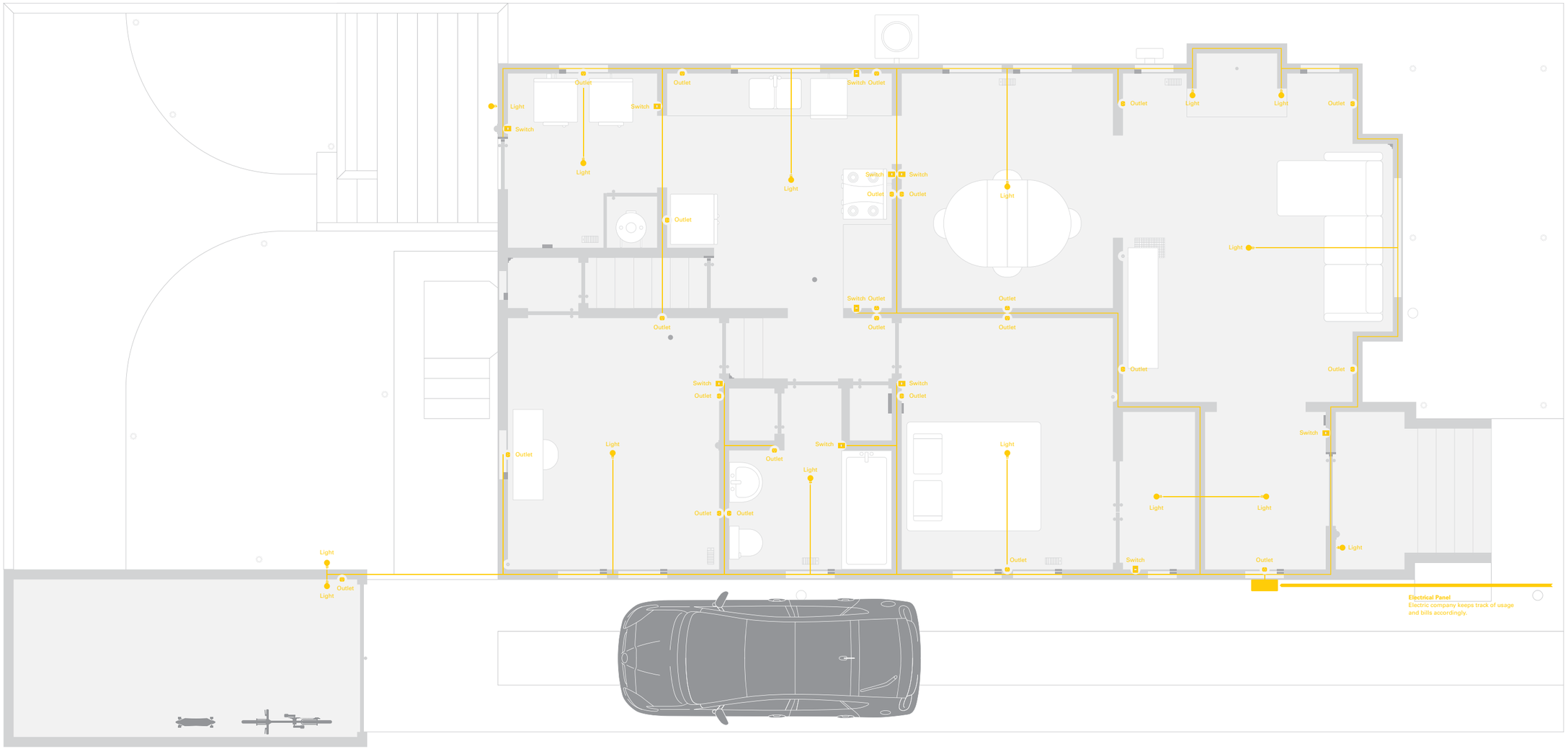


IoT devices in homes will produce and collect massive amounts of data:
Computers + Entertainment

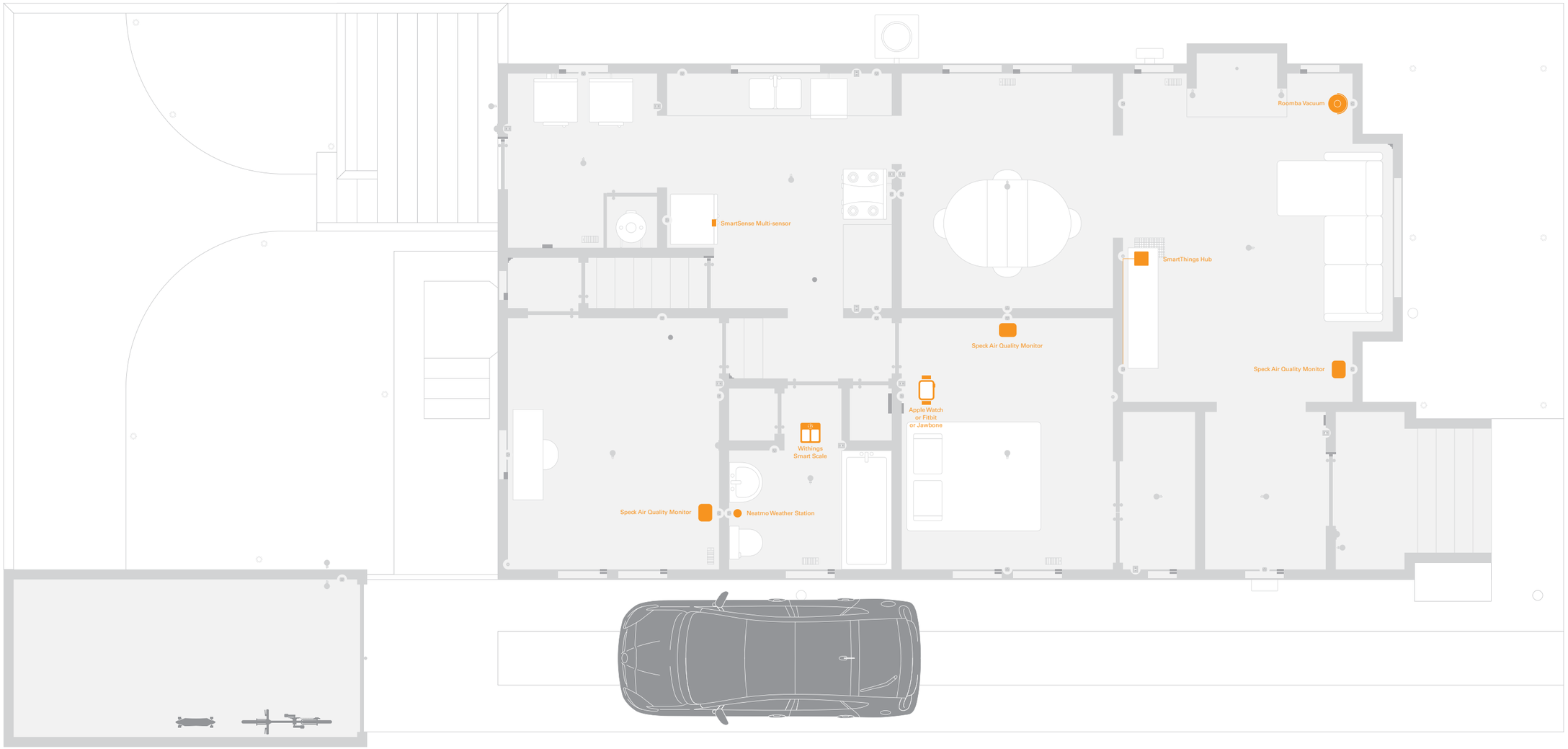


IoT devices in homes will produce and collect massive amounts of data:

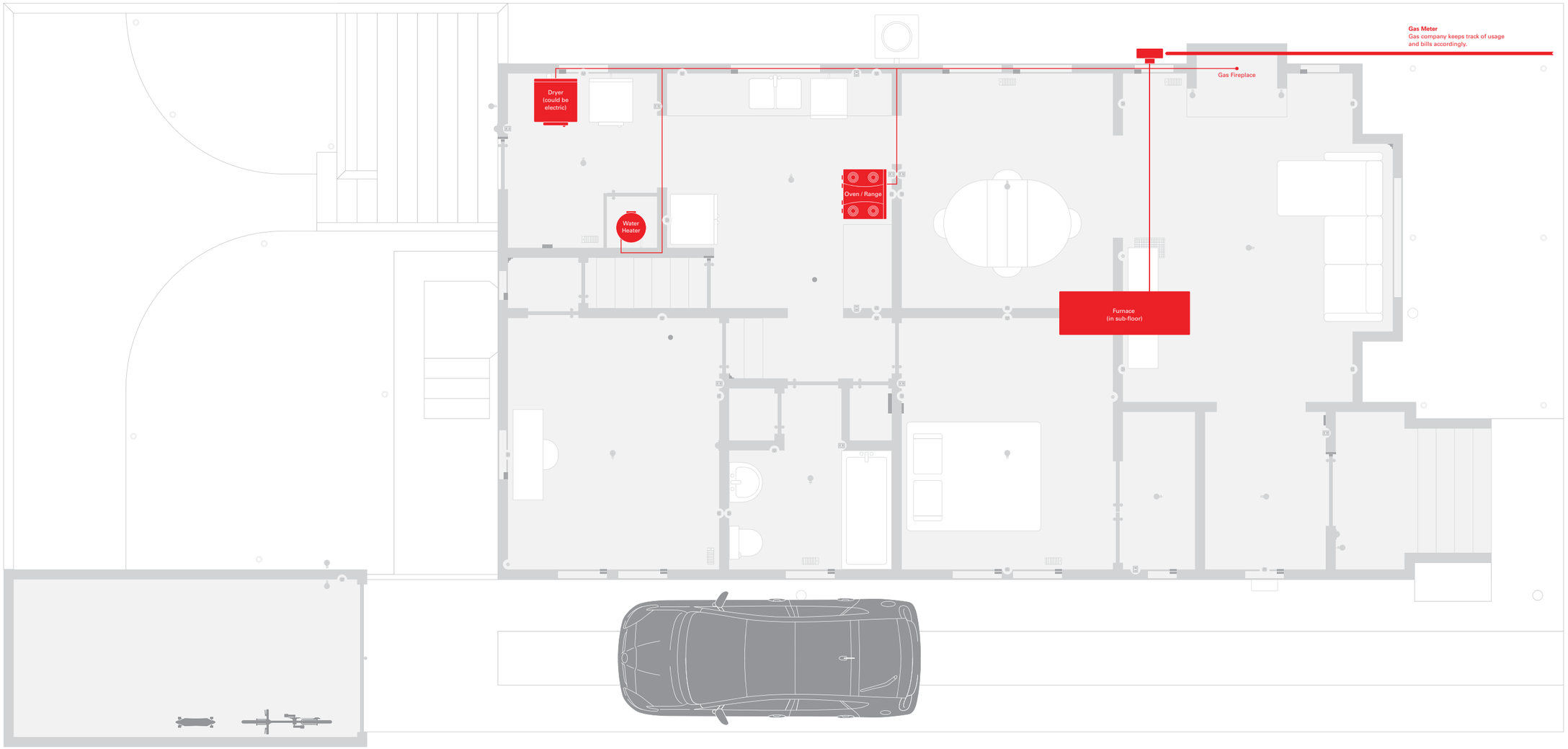
Electrical



IoT devices in homes will produce and collect massive amounts of data:
Health

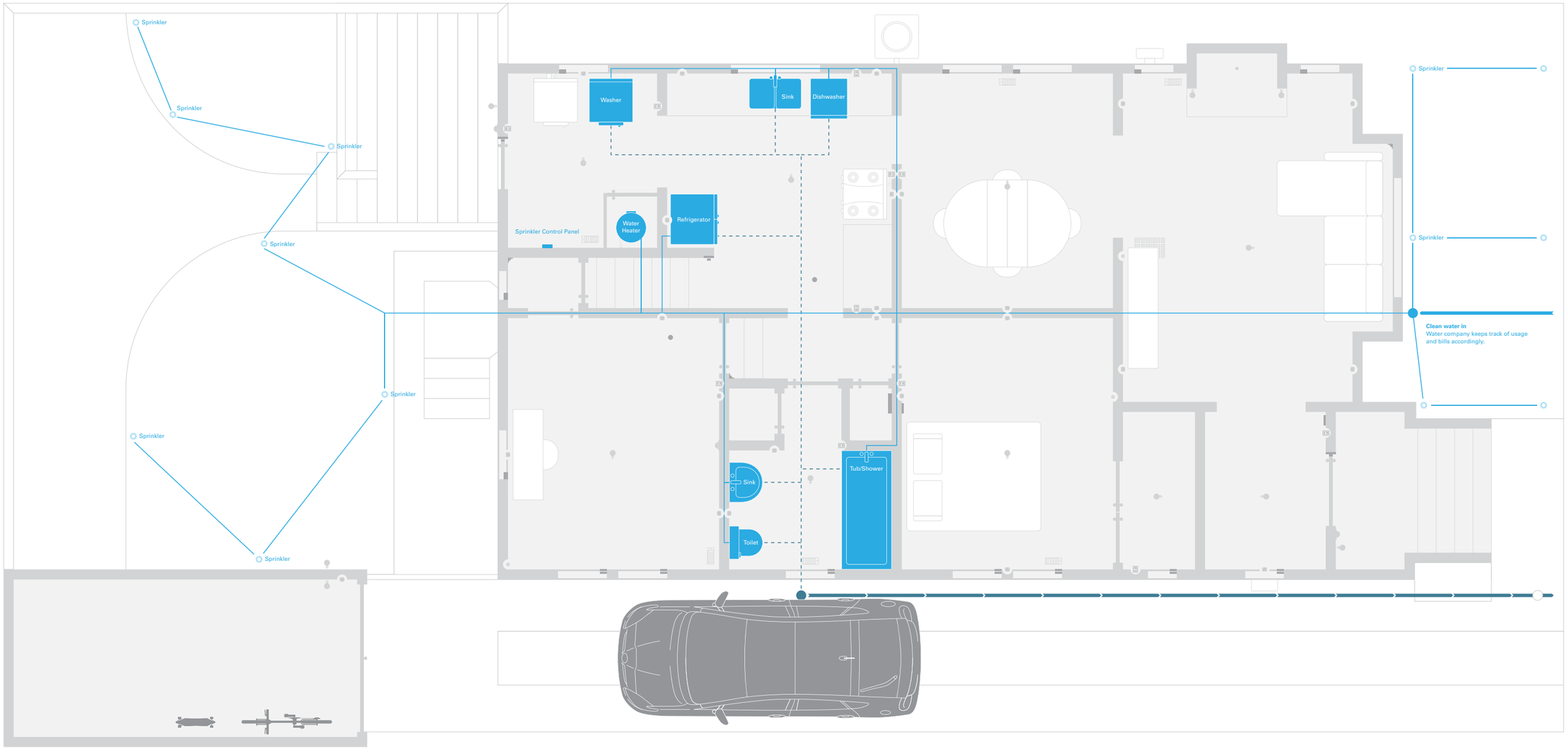


IoT devices in homes will produce and collect massive amounts of data:
Gas

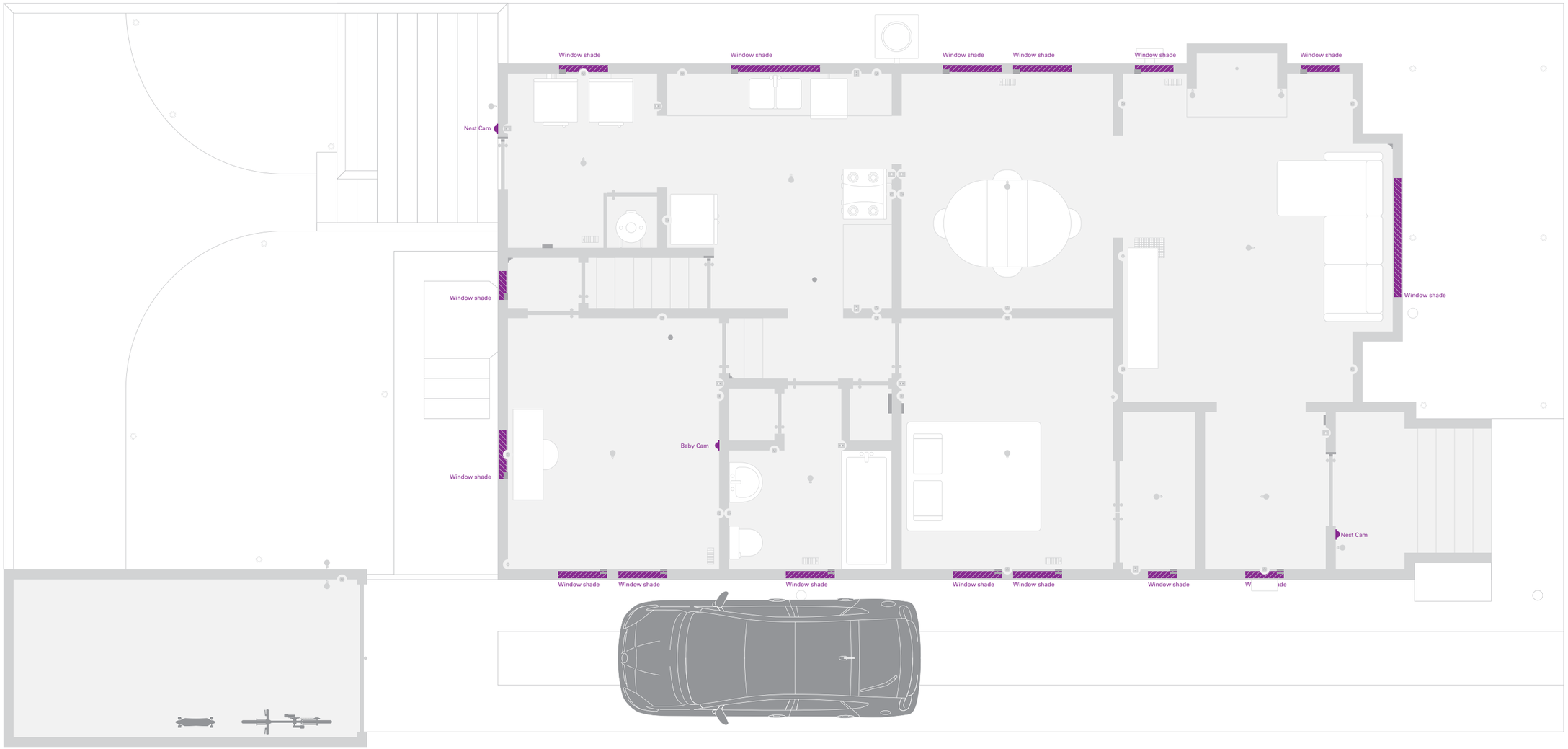


IoT devices in homes will produce and collect massive amounts of data:

Plumbing

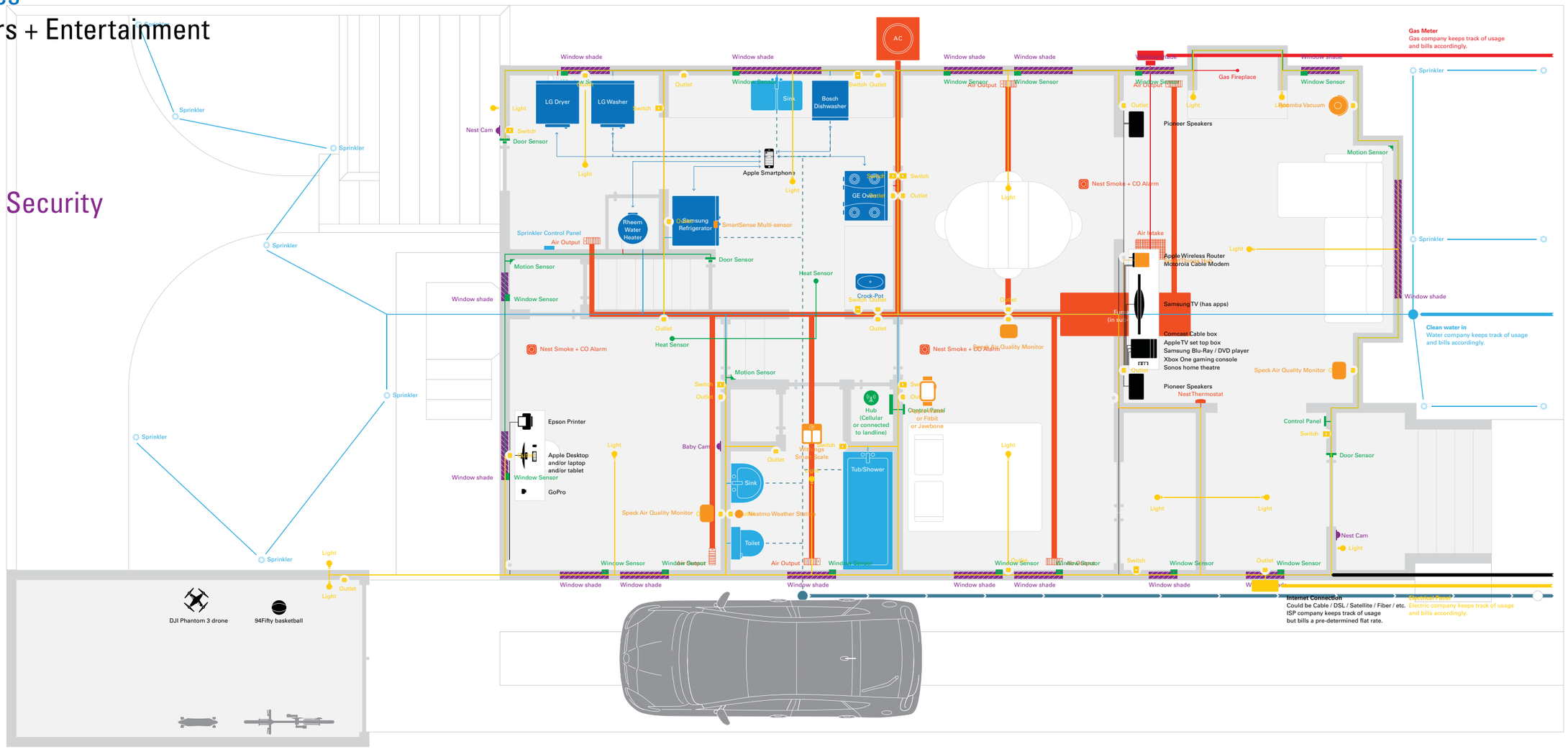


IoT devices in homes will produce and collect massive amounts of data:
Privacy + Security



IoT devices in homes will produce and collect massive amounts of data:

- HVAC
- Appliances
- Computers + Entertainment
- Electrical
- Gas
- Health
- Plumbing
- Privacy + Security



“Datafication” offers four successive levels of value,
(based on a model by Michael Porter)

Automation — enabling systems to run autonomously,
(e.g., programmed trading, self-driving cars, etc.).

Optimization — predicting changes (e.g., usage, failure, etc.),
and deploying resources accordingly (i.e., arbitrage).

Control — correcting variables that exceed thresholds,
ensuring that systems operate within bounds.

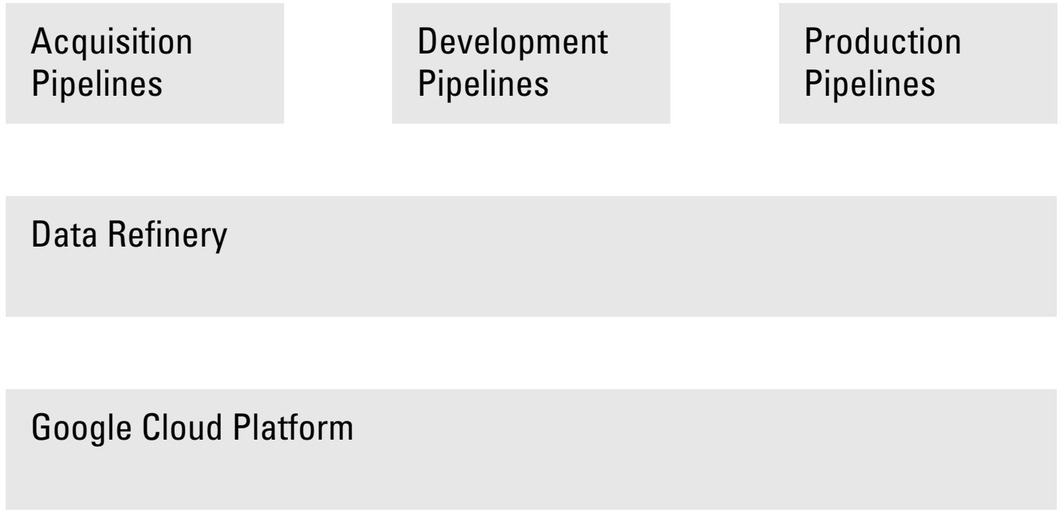
Monitoring — measuring operations;
sending alerts as variables approach thresholds.

Daikin could build or partner with a big-data refinery — on a cloud-based super-computer.

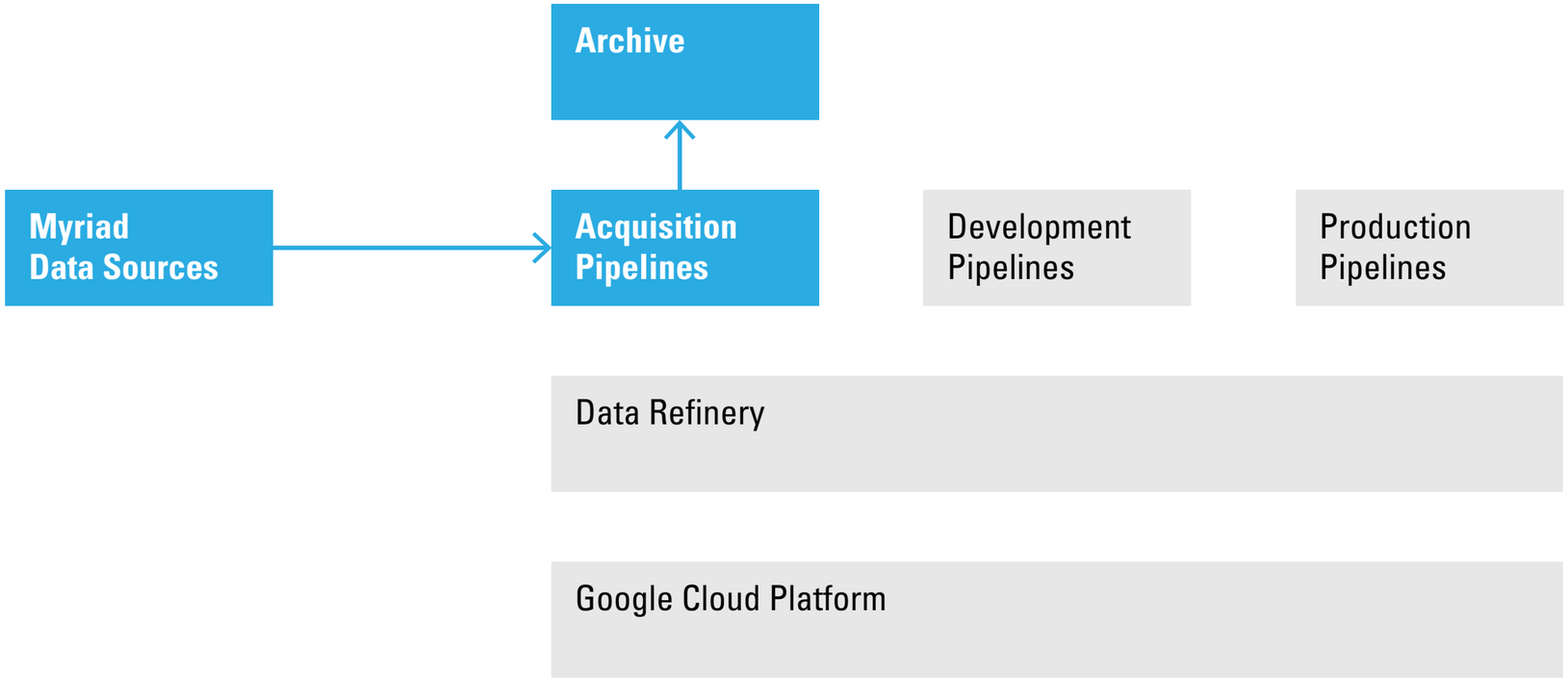
Data Refinery

Google Cloud Platform

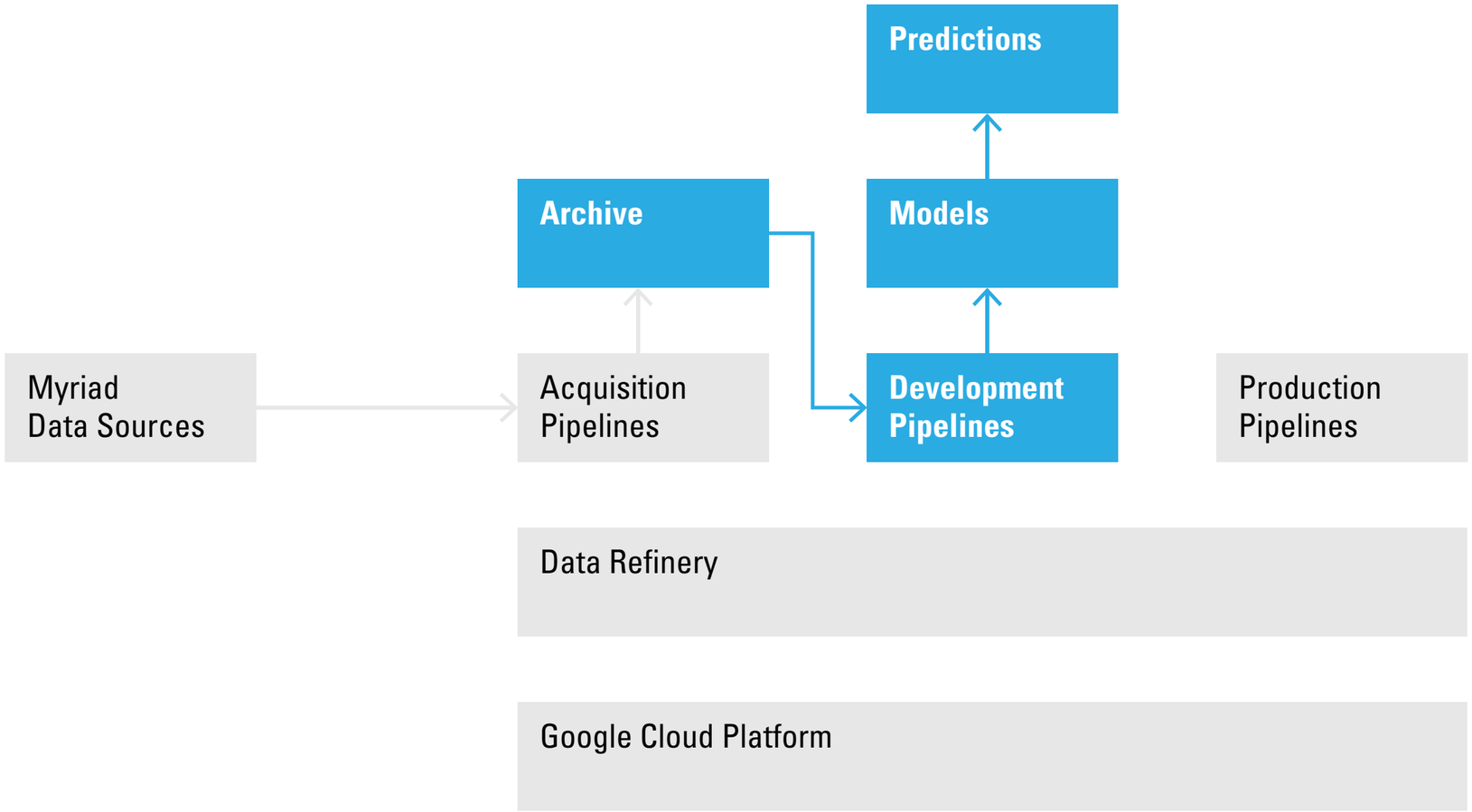
The platform could support a series of pipelines for chaining events (automated transforms) at massive scale.



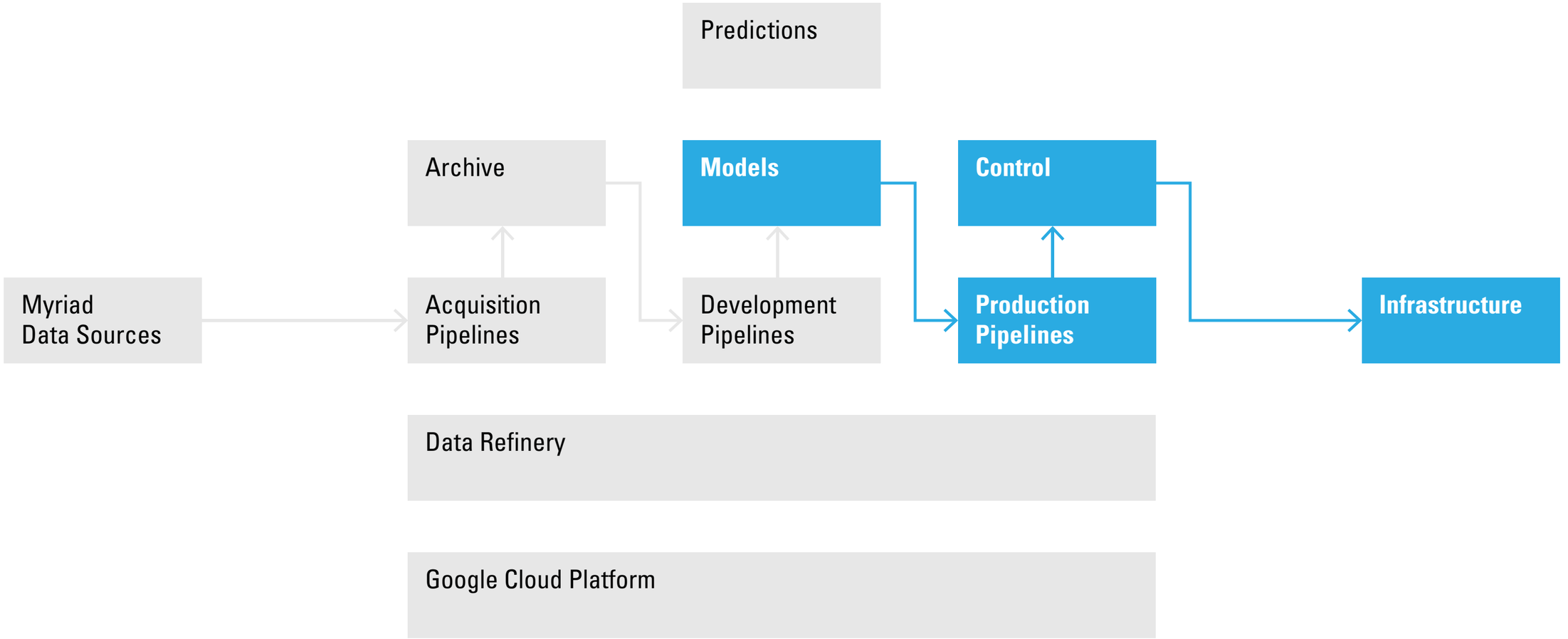
The platform could build an archive of data on consumers and equipment use.



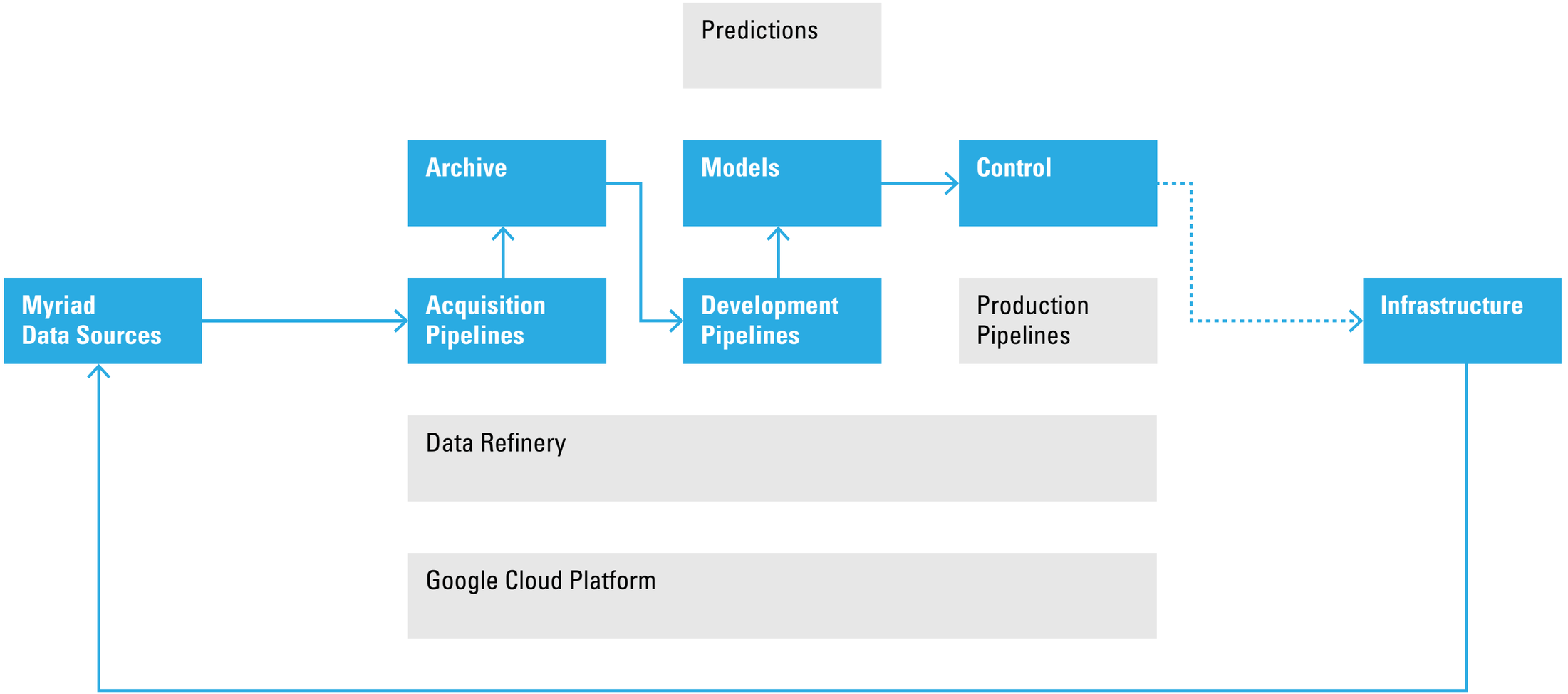
Large data sets enable development of models and predictions; the more data, the more accurate the results.



Models can be put into production, creating control systems that drive operations.



As production systems accumulate data and results, they improve their models — effectively "learning."



In the near term, Daikin has opportunities to use “datafication“ for ...

- Streamlining the commissioning process
- Ensuring consumers the process was done right
- Enabling more efficient operation
- Identifying potential failures before they happen
- Building a relationship with customers
- Collecting data to drive next generation of products

Special thanks to
Dennis Thoren
Akira Motomura

hugh@dubberly.com
@DubberlyDesign

Presentation posted at
presentations.dubberly.com/Daikin_IoT.pdf