**Application areas of DS**

* Business analytics
* Bioinformatics
* Empowers better business decision making through interpreting, modeling, and deployment.
* Social media and social network analysis.
* Smart cities
* Finance

**Industry 1.0**

The first Industrial Revolution represented the period between the 1760s and around 1840.

the First Industrial Revolution was the transition to new manufacturing processes using water and steam.

 beneficial in terms of manufacturing a larger number of various goods and creating a better standard of living for some.

The textile industry, in particular, was transformed by industrialization, as was transportation.

**Industry 2.0**

The beginning of 20th century marked the start of the second industrial revolution

the second industrial revolution picked up. Historians sometimes refer to this as “The Technological Revolution” occurring mainly in Britain, Germany and America.

During this time, new technological systems were introduced, most notably superior electrical technology which allowed for even greater production and more sophisticated machines.

**Industry 3.0**

in the last few decades of the 20th century, the first computer era.

Around 1970 the Third Industrial Revolution involved the use of electronics and IT (Information Technology) to further automation in production. Manufacturing and automation advanced considerably thanks to Internet access, connectivity and renewable energy.

Industry 3.0 introduced more automated systems onto the assembly line to perform human tasks, i.e. using Programmable Logic Controllers (PLC). Although automated systems were in place, they still relied on human input and intervention.

**Industry 4.0**

the Internet and telecommunication industry in the 1990’s revolutionized the way we connected and exchanged information.

The Fourth industrial Revolution is the era of smart machines, storage systems and production facilities that can autonomously exchange information, trigger actions and control each other without human intervention.

This exchange of information is made possible with the Industrial Internet of things (IIoT) as we know it today. Key elements of Industry 4.0 include:

Cyber-physical system — a mechanical device that is run by computer-based algorithms.

The Internet of things (IoT) — interconnected networks of machine devices and vehicles embedded with computerized sensing, scanning and monitoring capabilities.

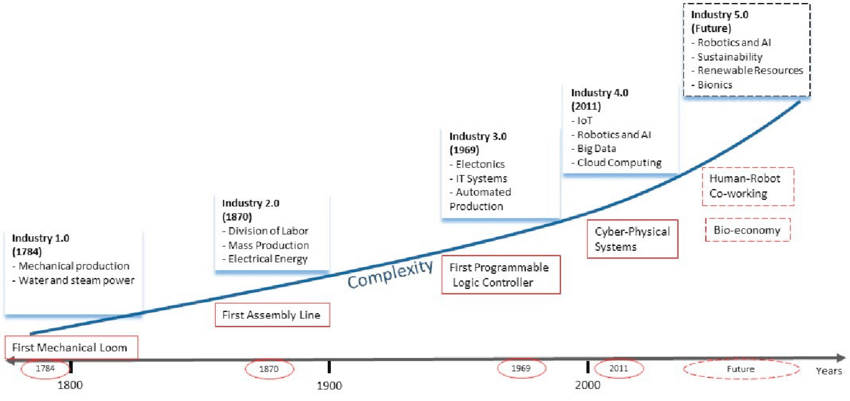
Cloud computing — offsite network hosting and data backup.

Cognitive computing — technological platforms that employ artificial intelligence.

Industry 4.0 starts to move towards Industry 5.0 when you begin to allow customers to customize what they want

**Industry 5.0**

Less than a decade has passed since talk of Industry 4.0 first surfaced in manufacturing circles, yet visionaries are already forecasting the next revolution — Industry 5.0. If the current revolution emphasizes the transformation of factories into IoT-enabled smart facilities that utilize cognitive computing and interconnect via cloud servers, Industry 5.0 is set to focus on the return of human hands and minds into the industrial framework.

Industry 5.0 is the revolution in which man and machine reconcile and find ways to work together to improve the means and efficiency of production. Funny enough, the fifth revolution could already be underway among the companies that are just now adopting the principles of Industry 4.0. Even when manufacturers start using advanced technologies, they are not instantly firing vast swaths of their workforce and becoming entirely computerized.

**Green IT (green information technology)**

Green [IT](https://searchdatacenter.techtarget.com/definition/IT) (green information technology) is the practice of environmentally sustainable computing.

The goals of green computing are similar to [green chemistry](https://en.wikipedia.org/wiki/Green_chemistry): reduce the use of hazardous materials, maximize [energy efficiency](https://en.wikipedia.org/wiki/Efficient_energy_use) during the product's lifetime, the [recyclability](https://en.wikipedia.org/wiki/Recycling) or [biodegradability](https://en.wikipedia.org/wiki/Biodegradation) of defunct products and factory waste. Green computing is important for all classes of systems, ranging from handheld systems to large-scale data centers.

Many corporate IT departments have green computing initiatives to reduce the environmental effect of their IT operations.

**Approaches**

Product longevity

LCA studies the environmental aspects and potential impacts throughout a product's life cycle (i.e. cradle-to-grave) from raw materials acquisition through production, use and disposal.

the biggest contribution to green computing usually is to prolong the equipment's lifetime. Another report from Gartner recommends to "Look for product longevity, including upgradability and modularity." For instance, manufacturing a new PC makes a far bigger [ecological footprint](https://en.wikipedia.org/wiki/Ecological_footprint) than manufacturing a new [RAM](https://en.wikipedia.org/wiki/RAM) module to upgrade an existing one.

Data center design

Data center facilities are heavy consumers of energy, accounting for between 1.1% and 1.5% of the world's total energy use in 2010 [1]. The U.S. Department of Energy estimates that data center facilities consume up to 100 to 200 times more energy than standard office buildings

The U.S. Department of Energy specifies five primary areas on which to focus energy efficient data center design best practices:[[20]](https://en.wikipedia.org/wiki/Green_computing#cite_note-20)

* Information technology (IT) systems
* Environmental conditions
* Air management
* Cooling systems
* Electrical systems

According to a [Greenpeace](https://en.wikipedia.org/wiki/Greenpeace) study, data centers represent 21% of the electricity consumed by the IT sector, which is about 382 billion kWh a year.[[29]](https://en.wikipedia.org/wiki/Green_computing#cite_note-29)

Software and deployment optimization

The efficiency of algorithms affects the amount of computer resources required for any given computing function and there are many efficiency trade-offs in writing programs. Algorithm changes, such as switching from a slow [search algorithm](https://en.wikipedia.org/wiki/Search_algorithm) to a fast search algorithm can reduce resource usage for a given task from substantial to close to zero.

In 2009, a study by a physicist at [Harvard](https://en.wikipedia.org/wiki/Harvard) estimated that the average Google search released 7 grams of carbon dioxide (CO₂).

However, Google disputed this figure, arguing instead that a typical search produced only 0.2 grams of CO₂

Resource allocation

Algorithms can also be used to route data to data centers where electricity is less expensive. Researchers from MIT, Carnegie Mellon University, and Akamai have tested an energy allocation algorithm that successfully routes traffic to the location with the cheapest energy costs. The researchers project up to a 40 percent savings on energy costs if their proposed algorithm were to be deployed. However, this approach does not actually reduce the amount of energy being used; it reduces only the cost to the company using it.

Larger server centers are sometimes located where energy and land are inexpensive and readily available. Local availability of renewable energy, climate that allows outside air to be used for cooling, or locating them where the heat they produce may be used for other purposes could be factors in green siting decisions.

Approaches to actually reduce the energy consumption of network devices by proper network/device management techniques are surveyed in. The authors grouped the approaches into 4 main strategies, namely (i) Adaptive Link Rate (ALR), (ii) Interface Proxying, (iii) Energy Aware Infrastructure, and (iv) Max Energy Aware Applications.

Virtualizing

New virtual technologies, such as [operating-system-level virtualization](https://en.wikipedia.org/wiki/Operating-system-level_virtualization) can also be used to reduce energy consumption. These technologies make a more efficient use of resources, thus reducing energy consumption by design. Also, the consolidation of virtualized technologies is more efficient than the one done in [virtual machines](https://en.wikipedia.org/wiki/Virtual_machine), so more services can be deployed in the same physical machine, reducing the amount of hardware needed.

Cloud computing

Cloud computing addresses two major ICT challenges related to Green computing – energy usage and [resource consumption](https://en.wikipedia.org/wiki/Resource_consumption). [Virtualization](https://en.wikipedia.org/wiki/Virtualization), [dynamic provisioning environment](https://en.wikipedia.org/wiki/Dynamic_provisioning_environment), multi-tenancy, [green data center](https://en.wikipedia.org/wiki/Green_data_center) approaches are enabling cloud computing to lower carbon emissions and energy usage up to a great extent*. Large enterprises and small businesses can reduce their direct energy consumption and carbon emissions by up to 30% and 90% respectively by moving certain on-premises applications into the cloud*.

Power management

The [Advanced Configuration and Power Interface](https://en.wikipedia.org/wiki/Advanced_Configuration_and_Power_Interface) (ACPI), an open industry standard, allows an operating system to directly control the power-saving aspects of its underlying hardware. This allows a system to automatically turn off components such as [monitors](https://en.wikipedia.org/wiki/Computer_display) and [hard drives](https://en.wikipedia.org/wiki/Hard_drive) after set periods of inactivity. In addition, a system may [hibernate](https://en.wikipedia.org/wiki/Hibernation_(computing)), when most components (including the [CPU](https://en.wikipedia.org/wiki/Central_processing_unit) and the system RAM) are turned off.

Storage

Smaller form factor (e.g., 2.5 inch) [hard disk drives](https://en.wikipedia.org/wiki/Hard_disk_drives) often consume less power per gigabyte than physically larger drives. Unlike hard disk drives, [solid-state drives](https://en.wikipedia.org/wiki/Solid-state_drive) store data in flash memory or [DRAM](https://en.wikipedia.org/wiki/DRAM). With no moving parts, power consumption may be reduced somewhat for low-capacity flash-based devices

Display

Unlike other display technologies, [electronic paper](https://en.wikipedia.org/wiki/Electronic_paper) does not use any power while displaying an image.[[58]](https://en.wikipedia.org/wiki/Green_computing#cite_note-58) [CRT monitors](https://en.wikipedia.org/wiki/CRT_monitor) typically use more power than LCD monitors. They also contain significant amounts of lead. [LCD monitors](https://en.wikipedia.org/wiki/LCD_monitor) typically use a [cold-cathode fluorescent bulb](https://en.wikipedia.org/wiki/Cold_cathode) to provide light for the display. Some newer displays use an array of [light-emitting diodes](https://en.wikipedia.org/wiki/Light-emitting_diode) (LEDs) in place of the fluorescent bulb, which reduces the amount of electricity used by the display.[[59]](https://en.wikipedia.org/wiki/Green_computing#cite_note-59) Fluorescent back-lights also contain mercury, whereas LED back-lights do not.

A [light-on-dark color scheme](https://en.wikipedia.org/wiki/Light-on-dark_color_scheme), also called *dark mode*, is a [color scheme](https://en.wikipedia.org/wiki/Color_scheme) that requires less energy to display on new display technologies, such as [OLED](https://en.wikipedia.org/wiki/OLED).[[60]](https://en.wikipedia.org/wiki/Green_computing#cite_note-60) This positively impacts battery life and energy consumption. While an OLED will consume around 40% of the power of an LCD displaying an image that is primarily black, it can use more than three times as much power to display an image with a white background, such as a document or web site.

Materials recycling

Recycling computing equipment can keep harmful materials such as lead, mercury, and hexavalent chromium out of [landfills](https://en.wikipedia.org/wiki/Landfill), and can also replace equipment that otherwise would need to be manufactured, saving further energy and emissions. Computer systems that have outlived their particular function can be re-purposed, or donated to various charities and non-profit organizations

The recycling of old computers raises an important privacy issue. The old storage devices still hold private information, such as emails, passwords, and credit card numbers, which can be recovered simply by someone's using software available freely on the Internet. Deletion of a file does not actually remove the file from the hard drive. Before recycling a computer, users should remove the hard drive

Telecommuting

[Teleconferencing](https://en.wikipedia.org/wiki/Teleconference) and [telepresence](https://en.wikipedia.org/wiki/Telepresence) technologies are often implemented in green computing initiatives. The advantages are many; increased worker satisfaction, reduction of [greenhouse gas emissions](https://en.wikipedia.org/wiki/Greenhouse_gas_emissions) related to travel, and increased profit margins as a result of lower overhead costs for office space, heat, lighting, etc.

Telecommunication network devices energy indices

The information and communication technologies (ICTs) energy consumption, in the US and worldwide, has been estimated respectively at 9.4% and 5.3% of the total electricity produced. The energy consumption of ICTs is today significant even when compared with other industries. Some study tried to identify the key energy indices that allow a relevant comparison between different devices (network elements). This analysis was focused on how to optimize device and network consumption for carrier telecommunication by itself. The target was to allow an immediate perception of the relationship between the network technology and the environmental effect. These studies are at the start and the gap to fill in this sector is still huge and further research will be necessary.

Supercomputers

The inaugural [Green500](https://en.wikipedia.org/wiki/Green500) list was announced on November 15, 2007, at SC|07. As a complement to the TOP500, the unveiling of the Green500 ushered in a new era where supercomputers can be compared by performance-per-watt.[[85]](https://en.wikipedia.org/wiki/Green_computing#cite_note-85) As of 2019, two Japanese supercomputers topped the Green500 energy efficiency ranking with performance exceeding 16 GFLOPS/watt, and two IBM AC922 systems followed with performance exceeding 15 GFLOPS/watt.

**Social media sites generating largest data**

**Google:** 40,000 Google Web Searches Per Second.

**Facebook:** 500 Terabytes Per Day.

1.5 billion people are active on Facebook daily

There are five new Facebook profiles created every second!

More than 300 million photos get uploaded per day

Every minute there are 510,000 comments posted and 293,000 statuses updated

**Instagram:**

400 million who are active every day

Each day 95 million photos and videos are shared on Instagram

100 million people use the Instagram “stories” feature daily

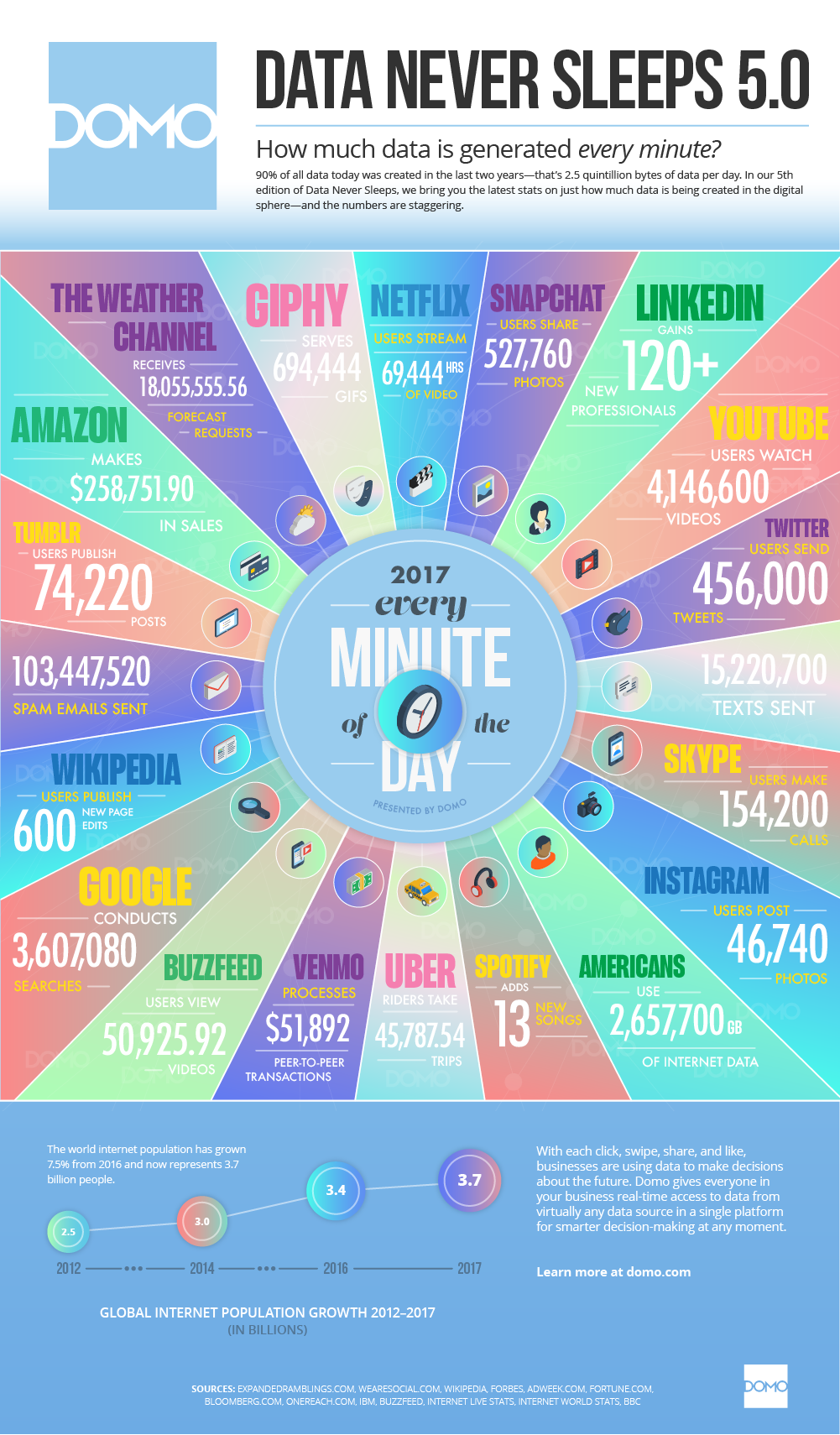
**YouTube:**

It is also the second-most popular search engine right after Google, racking up over a billion hours of views every day

Every Second. It has been 5 seconds. 318,287 videos watched. Source. 45 hours of video uploaded. Source. 63,657 hours of video watched.

**Twitter**: 12 Terabytes Per Day.

**Amazon**: $258,751.90 in Sales Per Minute.



**Sigma Technologies:**

Ben Runkle, CEO, Sigma Technologies, is trying to resolve a huge problem. The

company is consistently losing long-time customers. He does not know why they

are leaving, but he must do something fast. He is convinced that in order to reduce

his churn, he must create new products and features, and consolidate existing

technologies. he calls in his chief data scientist, Dr. Jessie Hughan.

She shows Runkle the most recent transcripts and finds something surprising:

".... Not sure how to export this; are you?"

"Where is the button that makes a new list?"

"Wait, do you even know where the slider is?"

"If I can't figure this out today, it's a real problem..."

It is clear that customers were having problems with the existing UI/UX, and

weren't upset due to a lack of features. Runkle and Hughan organized a mass UI/UX

overhaul and their sales have never been better.

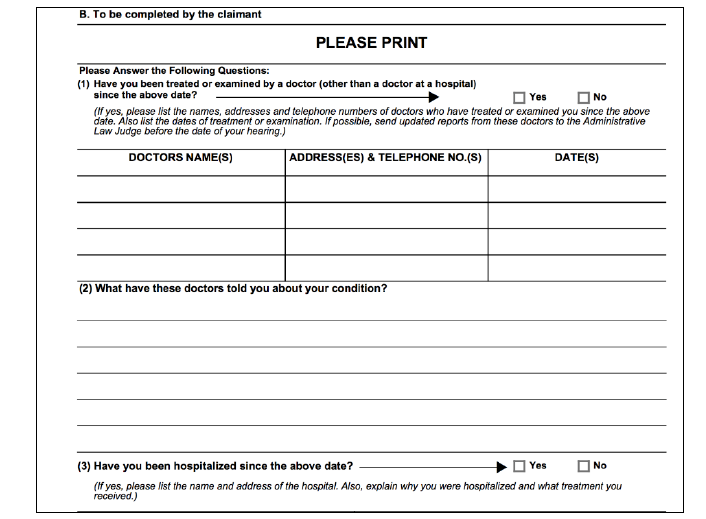
**Data Science case Study - Problem Solving statement:**

Case study – automating government paper pushing

Social security claims are known to be a major hassle for both the agent reading

it and for the person who wrote the claim. Some claims take over 2 years to get

resolved in their entirety, and that's absurd! Let's look at what goes into a claim:



Not bad. It's mostly just text, though. Fill this in, then that, then this, and so on. You

can see how it would be difficult for an agent to read these all day, form after form.

There must be a better way!

Well, there is. Elder Research Inc. parsed this unorganized data and was able to

automate 20% of all disability social security forms. This means that a computer

could look at 20% of these written forms and give its opinion on the approval.

Not only that, the third-party company that is hired to rate the approvals of the

forms actually gave the machine-graded forms a higher grade than the human forms.

So, not only did the computer handle 20% of the load, it, on average, did better than

a human.