

# Module 2: Healthcare Data

## Topics



- Healthcare Data

## Learning Objectives

- The student will be able to identify the data relevant to each of key the stakeholders in the healthcare ecosystem.
- Identify data sources, users and uses in the healthcare ecosystem.

# Data Science (statistics) in History

The statistical analysis of the data from the Crimea showed that the unsanitary living conditions led to the diseases that caused the high mortality rates, not lack of food and supplies as she had originally believed. They also discovered that the problems in the military hospitals were not limited to wartime. In peacetime, soldiers died at twice the rate as civilians aged 25 to 35.

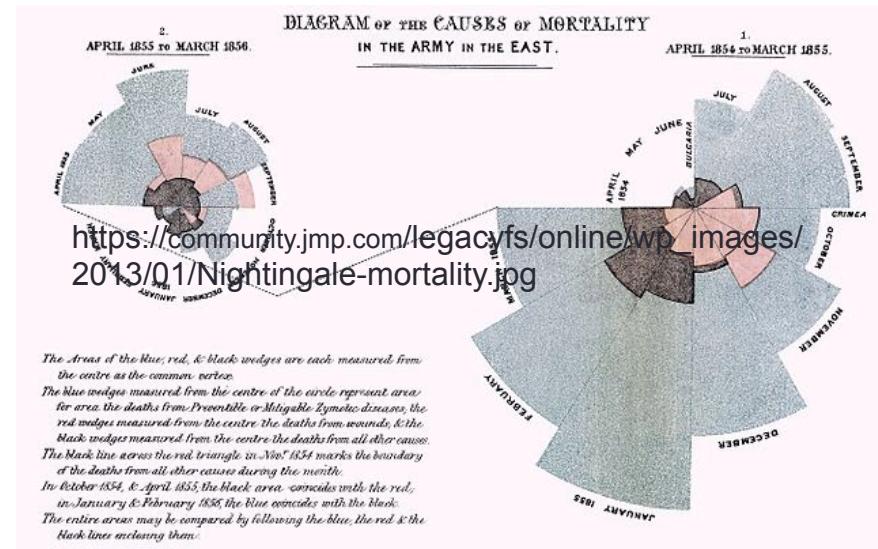
Nightingale created diagrams to persuade Queen Victoria of the need for reform by improving the sanitary conditions of military hospitals.



**Florence Nightingale** (English)  
statistician, founder of modern  
nursing.

Feb 4, 2013 Celebrating statisticians: Florence Nightingale

<https://community.jmp.com/t5/JMP-Blog/Celebrating-statisticians-Florence-Nightingale/ba-p/30247>



# Healthcare Data

**NOTE:** We will NOT be covering ALL the different kinds of Healthcare data you may encounter.

The hard drive IBM shipped in 1956:

- \* Stored 5 megabytes (MB)
  - \* Cost \$11,000 per megabyte
  - \* Was 60 inches long x 68 inches high x 29 inches deep
  - \* Weighed about 1 ton

<https://www.backblaze.com/blog/94-trillion-petabyte/>



<http://www.terc.org/ART/>

## The Healthcare Data Explosion

2012  
500  
petabytes

Worldwide  
healthcare data  
is expected  
to grow to  
**50 times**  
the current total

2020  
**25,000**  
**petabytes**

Feldman, Bonnie, Ellen M. Martin, and Tobi Skotnes. "Big Data in Healthcare Hype and Hope." October 2012. Dr. Bonnie 360 (2012).

[https://www.ghdonline.org/uploads/big-data-in-healthcare\\_B\\_Kaplan\\_2012.pdf](https://www.ghdonline.org/uploads/big-data-in-healthcare_B_Kaplan_2012.pdf)

# Healthcare Data



Figure 1 illustrates examples of sources, users, and uses of personal health data today.

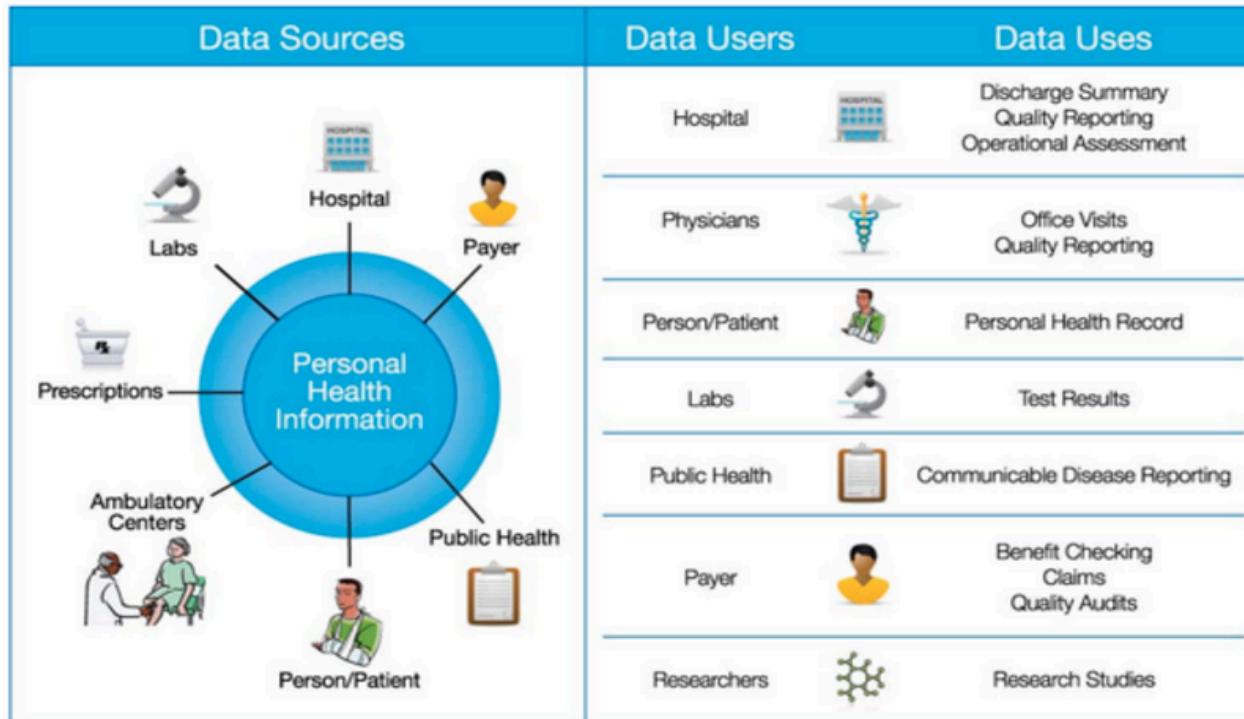


Figure 1. Data sources, users, and uses. (Modified from Foundation for eHealth Initiative, November 2004.)

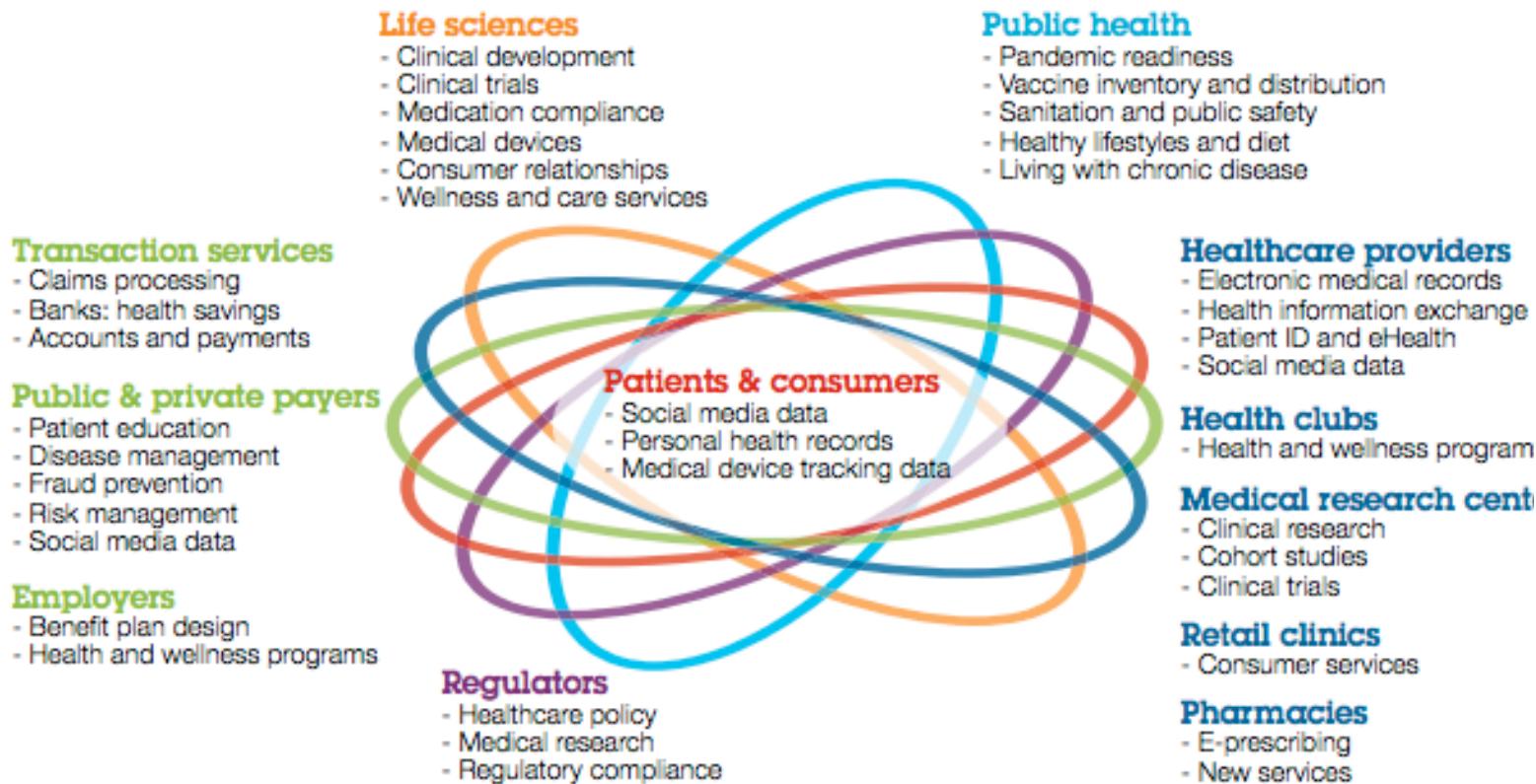
<https://storify.com/public/templates/slideshow/index.html?src=/storify.com/HealthcareWen/health-data-sources-and-users#6> (slide 6 of 206)

For more info on the adoption of EMRs see: <https://www.healthit.gov/sites/default/files/data-brief/2014HospitalAdoptionDataBrief.pdf>



## Analytics across the ecosystem - A prescription for optimizing healthcare outcomes

To optimize life sciences outcomes, organizations must drive analytics not just across the enterprise, but throughout the healthcare ecosystem IBM

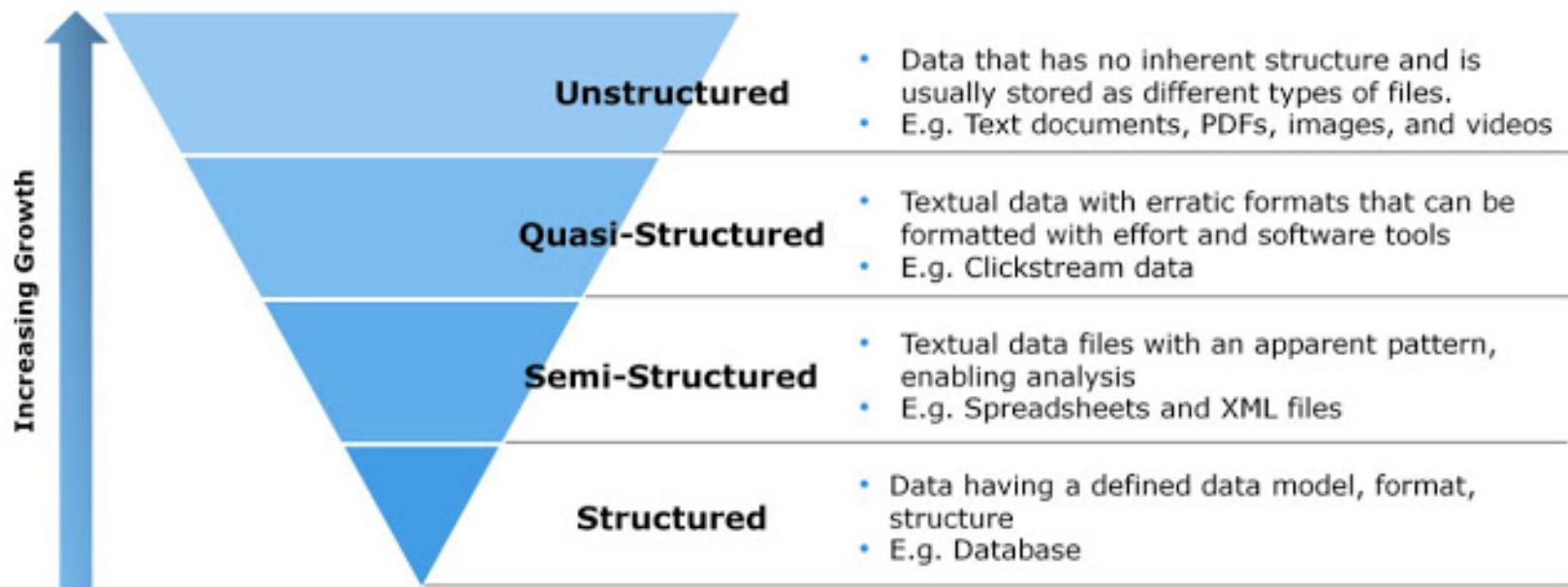


Healthcare organizations need to capture, share, integrate and analyze data across all the ecosystem stakeholders.

# Data Classifications

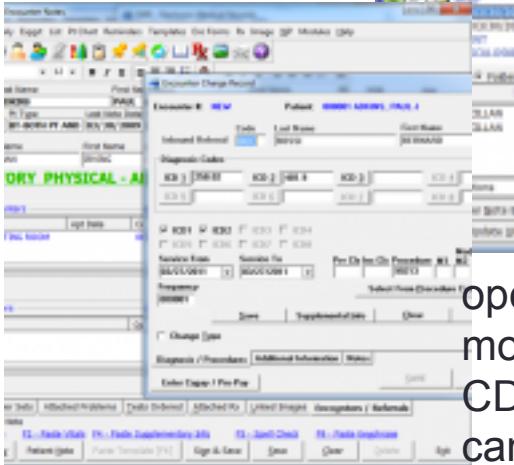
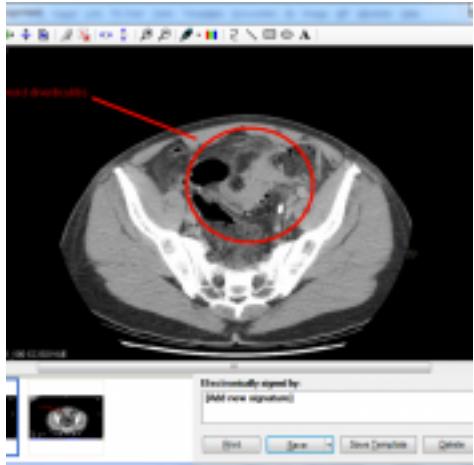


Based on how it is stored and managed, digital data can be broadly classified as either structured data or unstructured data. Structured data is organised in fixed fields within a record or file.



# Electronic Medical Records

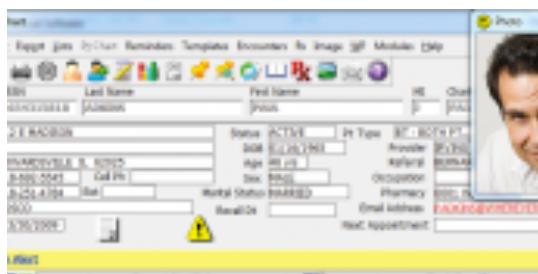
<https://youtu.be/RtianiGQR1c>



<http://americanmedical.com/software-solutions/electronic-medical-records/ele>

<http://stackoverflow.com/questions/1008897/emr-electronic-medical-record-standard-record-format>

<https://www.healthit.gov/providers-professionals/faqs/what-information-does-electronic-health-record-ehr-contain>



openEHR is a format for storage, more expressive than HL7 v2, v3 and CDA, can be converted easily.

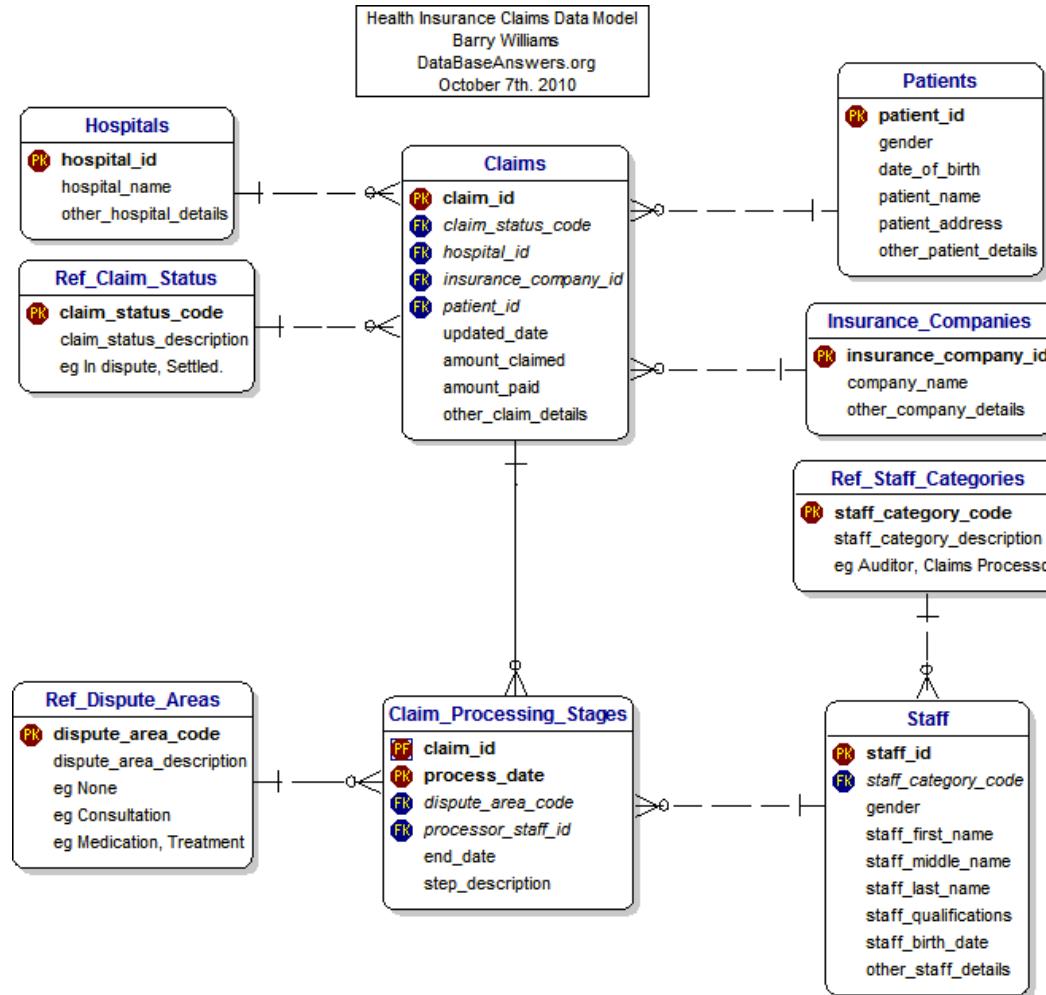
Open specifications: <http://openehr.org/programs/specification/releases/1.0.2>

For the interchange format, use HL7 v2, v3 and CDA. Also consider CCR and CCD. <http://www.aafp.org/practice-management/health-it/astm.html>



SEE THIS SITE

# Claims Data



[http://www.databaseanswers.org/data\\_models/health\\_insurance\\_claims/index.htm](http://www.databaseanswers.org/data_models/health_insurance_claims/index.htm)

# Lab Test Report

**LabCorp**  
Laboratory Corporation of America

LabCorp San Diego  
13112 Evening Creek Dr Ste 200  
San Diego, CA 92128

Phone: 858-668-3700

Patient Name		Specimen Number					
Specimen Number	Patient ID	Control Number	Date and Time Collected	Date Received	Sex	Age(Y/M/D)	Date of Birth
<b>TESTS</b>							
Potassium, Serum	4.1		mmol/L	3.5 - 5.2	01		
Chloride, Serum	100		mmol/L	97 - 108	01		
Carbon Dioxide, Total	32		mmol/L	20 - 32	01		
Calcium, Serum	9.7		mg/dL	8.5 - 10.6	01		
Protein, Total, Serum	7.7		g/dL	6.0 - 8.5	01		
Albumin, Serum	4.4		g/dL	3.5 - 5.5	01		
Globulin, Total	3.3		g/dL	1.5 - 4.5			
A/G Ratio	1.3			1.1 - 2.5			
Bilirubin, Total	0.9		mg/dL	0.1 - 1.2	01		
Alkaline Phosphatase, S	71		IU/L	25 - 150	01		
AST (SGOT)	23		IU/L	0 - 40	01		
ALT (SGPT)	28		IU/L	0 - 55	01		
Uric A+ESR-Wes+ANA+RA Qn+CR...							
Uric Acid, Serum	6.5		mg/dL	2.4 - 8.2	01		
Sedimentation Rate-Westergren			mm/hr	0 - 15	01		
Antistreptolysin O Ab	20.0		IU/mL	0.0 - 200.0	01		
C-Reactive Protein, Quant	0.7		mg/L.....	0.0 - 4.9	01		
RA Latex Turbid.	6.1		IU/mL	0.0 - 13.9	01		
Antinuclear Antibodies Direct	20		AU/mL	0 - 99	01		
			Negative	<100			
			Equivocal	100 - 120			
			Positive	>120			
Lipid Panel							
Cholesterol, Total	157		mg/dL	100 - 199	01		
Triglycerides	54		mg/dL	0 - 149	01		
HDL Cholesterol	56		mg/dL	>39	01		
Comment				01			
According to ATP-III Guidelines, HDL-C >59 mg/dL is considered a negative risk factor for CHD.							
VLDL Cholesterol Cal	11		mg/dL	5 - 40			
LDL Cholesterol Calc	90		mg/dL	0 - 99			
HAVAb+HBcAb+HBsAg+BCVAb							
Hep A Ab, Total			Negative		Negative	01	
HBsAg Screen			Negative		Negative	01	
Hep B Core Ab, Tot			Negative		Negative	01	
HCV Ab	<0.1		s/co ratio	0.0 - 0.9	01		
Negative							
Not infected with HCV, unless recent infection is suspected or other evidence exists to indicate HCV infection.							
Hgb A1c with MPG Estimation							
Hemoglobin A1c	5.4		%	<7.0	01		
			Diabetic Adult	<7.0			
			Healthy Adult	4.8 - 5.9			
KIM, YONG HO							

FINAL REPORT

Page 2 of 3

©2004-09 Laboratory Corporation of America Holdings  
All Rights Reserved  
DOC1 Ver:1.39



Most common lab tests:

Complete Blood Count (CBC)  
Basic Metabolic Profile (BMP)  
or Comprehensive Metabolic Profile (CMP)  
Urinalysis (UA)

Many more, such as:

Coagulation profiles  
Cardiac enzymes  
Thyroid function tests  
Liver function tests  
Lipid profiles

[https://usercontent2.hubstatic.com/6987327\\_f520.jpg](https://usercontent2.hubstatic.com/6987327_f520.jpg)

# Data Resources



## HealthData.gov

<https://www.healthdata.gov/search/type/dataset>

This screenshot shows the "Data Resources in the Health Sciences" guide from the University of Washington Health Sciences Library. The page includes a navigation bar with links like "Library Guides", "HSL", "Data Resources in the Health Sciences", "Ask Us!", "Tools", "Services", and "About Us". A search bar at the top right has the placeholder "Search this Guide". The main content area features a circular graphic titled "UK Data Archive interactive graphic on managing data" with sections for "Find", "Preserve/Store", "Describe", "Analyze/Visualize", and "Share". Below the graphic, there's a detailed section on "Find Data" with sub-links for "Find Clinical Data" and "Find Scientific Data". A sidebar on the left contains sections for "Overview", "Find Data", "Preserve/Store Data", "Describe Data", "Analyze/Visualize Data", "Share Data", and "Cite Data". Another sidebar on the far left provides information about "Learn About Data", mentioning projects like OHSU Big Data to Knowledge (BD2K), Open Educational Resources (OERs) Project, and Community Data Science Workshops (CDSW). At the bottom, there's a link to "http://guides.lib.uw.edu/hsl/data".

This screenshot shows the "Health" section of the DATA.GOV website. The header includes the DATA.GOV logo, a search bar, and navigation links for "DATA", "TOPICS", "IMPACT", "APPLICATIONS", "DEVELOPERS", and "CONTACT". The main content area has a blue header with a plus sign icon and the word "HEALTH". Below it, there are links for "Updates", "Data", "HealthData.gov", and "Contact Health". A large text block reads: "Explore datasets, tools, and applications related to health and health care. These resources come from across the Federal Government with the goal of improving the health and lives of all Americans." A "HIGHLIGHTS" section features a box for "Hospital Charge Data" with a "View this Dataset" button. Below this, there's a link to "https://www.data.gov/health/".

<http://guides.lib.uw.edu/hsl/data>

<http://guides.lib.uw.edu/c.php?g=99209&p=642709#16640559>