
Glossary

Key chapters in which a term is used are indicated in square brackets.

- Accountability** Security function that ensures users are responsible for their access to and use of information based on a documented need and right to know [5].
- Accountable Care Organizations (ACOs)** An organization of health care providers that agrees to be accountable for the quality, cost, and overall care of their patients. An ACO will be reimbursed on the basis of managing the care of a population of patients and are determined by quality scores and reductions in total costs of care [14,27].
- Accountable care** A descendant of managed care, accountable care is an approach to improving care and reducing costs. See: **AccountableCare Organizations** [10].
- ACO** See: **Accountable Care Organizations** [14, 27].
- Active failures** Errors that occur in an acute situation, the effects of which are immediately felt [4].
- Active phase** The phase of a clinical research study during which investigators collect data from participants receiving an intervention or interventions under study. It is also common to monitor study participants for adverse events during this phase [26].
- Active storage** In a hierarchical data-storage scheme, the devices used to store data that have long-term validity and that must be accessed rapidly [5].
- Acute Physiology and Chronic Health Evaluation, Version III [APACHE III]** A scoring system for rating the disease severity for particular use in intensive care units [18].
- Address** An indicator of location; typically a number that refers to a specific position in a computer's memory or storage device; see also: **Internet Address** [5].
- ADE** See: **Adverse Drug Events** [19].
- Admission-discharge-transfer (ADT)** The core component of a hospital information system that maintains and updates the hospital census, including bed assignments of patients [14].
- ADT** See: **Admission-discharge-transfer** [14].
- Advanced Research Projects Agency Network (ARPANET)** A large wide-area network created in the 1960s by the U.S. Department of Defense Advanced Research Projects Agency (DARPA) for the free exchange of information among universities and research organizations; the precursor to today's Internet [1,5].
- Adverse drug events (ADEs)** Undesired patient events, whether expected or unexpected, that are attributed to administration of a drug [19].
- Aggregations** In the context of information retrieval, collections of content from a variety of content types, including bibliographic, full-text, and annotated material [21].
- AHIMA** See: **American Health Information Management Association** [21].
- Alert message** A computer-generated warning that is generated when a record meets pre-specified criteria, often referring to a potentially dangerous situation that may require action; e.g., receipt of a new laboratory test result with an abnormal value [12].
- Algorithmic process** An algorithm is a well-defined procedure or sequence of steps for solving a problem. A process that follows prescribed steps is accordingly an algorithmic process [1].
- Alphanumeric** Descriptor of data that are represented as a string of letters and numeric digits, without spaces or punctuation [21].

- Ambulatory medical record systems (AMRS)** A clinical information system designed to support all information requirements of an outpatient clinic, including registration, appointment scheduling, billing, order entry, results reporting, and clinical documentation [14].
- American Health Information Management Association (AHIMA)** Professional association devoted to the discipline of health information management (HIM) [21].
- American Medical Informatics Association (AMIA)** Professional association dedicated to biomedical and health informatics [15, 23].
- American National Standards Institute [ANSI]** A private organization that oversees voluntary consensus standards [7,23].
- American Recovery and Reinvestment Act of 2009** Public Law 111–5, commonly referred to as the Stimulus or Recovery Act, this legislation was designed to create jobs quickly and to invest in the nation’s infrastructure, education and healthcare capabilities [14].
- American Standard Code for Information Interchange (ASCII)** A 7-bit code for representing alphanumeric characters and other symbols [5].
- AMIA** See: American Medical Informatics Association [15, 23].
- AMRS** See: Ambulatory medical record systems [14].
- Analog signal** A signal that takes on a continuous range of values [5].
- Analog-to-digital conversion (ADC)** Conversion of sampled values from a continuous-valued signal to a discrete-valued digital representation [5].
- Anchoring and adjustment** A heuristic used when estimating probability, in which a person first makes a rough approximation (the anchor), then adjusts this estimate to account for additional information [3].
- Annotated content** In the context of information retrieval, content that has been annotated to describe its type, subject matter, and other attributes [21].
- Anonymize** Applied to health data and information about a unique individual, the act of de-identifying or stripping away any and all data which could be used to identify that individual [10].
- ANSI** See: American National Standards Institute [23].
- Antibiogram** Pattern of sensitivity of a microorganism to various antibiotics [19].
- APACHE III** See Acute Physiology and Chronic Health Evaluation, Version III [18].
- Apache Open source** web server software that was significant in facilitating the initial growth of the World Wide Web [6].
- Applets** Small computer programs that can be embedded in an HTML document and that will execute on the user’s computer when referenced [5,6].
- Application programming interface (API)** A specification that enables distinct software modules or components to communicate with each other [6].
- Application program** A computer program that automates routine operations that store and organize data, perform analyses, facilitate the integration and communication of information, perform bookkeeping functions, monitor patient status, aid in education [5].
- Applications (applied) research** Systematic investigation or experimentation with the goal of applying knowledge to achieve practical ends [1].
- Archival storage** In a hierarchical data-storage scheme, the devices used to store data for long-term backup, documentary, or legal purposes [5].
- Arden Syntax for Medical Logic Module A** coding scheme or language that provides a canonical means for writing rules that relate specific patient situations to appropriate actions for practitioners to follow. The Arden Syntax standard is maintained by HL7 [6, 22].
- Argument** A word or phrase that helps complete the meaning of a **predicate** [8].
- ARPANET** See Advanced Research Projects Agency Network [1,5].
- Artificial intelligence(AI)** The branch of computer science concerned with endowing computers with the ability to simulate intelligent human behavior [1].
- Artificial neural network** A computer program that performs classification by taking as input a set of findings that describe a given situation, propagating calculated weights through a network of several layers of interconnected nodes, and generating as output a set of numbers, where each output corresponds to the

- likelihood of a particular classification that could explain the findings [23].
- ASCII** See: American Standard Code for Information Interchange [5].
- Assembler** A computer program that translates assembly-language programs into machine-language instructions [5].
- Assembly language** A low-level language for writing computer programs using symbolic names and addresses within the computer's memory [5].
- Asynchronous Transfer Mode (ATM)** A network protocol designed for sending streams of small, fixed length cells of information over very high-speed, dedicated connections, often digital optical circuits [5].
- Audit trail** A chronological record of all accesses and changes to data records, often used to promote accountability for use of, and access to, medical data [5].
- Authenticated** A process for positive and unique identification of users, implemented to control system access [5].
- Authorized** Within a system, a process for limiting user activities only to actions defined as appropriate based on the user's role [5].
- Automated indexing** The most common method of full-text indexing; words in a document are stripped of common suffixes, entered as items in the index, then assigned weights based on their ability to discriminate among documents (see vector-space model) [21].
- Availability** In decision making, a heuristic method by which a person estimates the probability of an event based on the ease with which he can recall similar events [3]. In security systems, a function that ensures delivery of accurate and up-to-date information to authorized users when needed [5].
- Averaging out at chance nodes** The process by which each chance node of a decision tree is replaced in the tree by the expected value of the event that it represents [3].
- Backbone links** Sections of high-capacity trunk (backbone) network that interconnect regional and local networks [5].
- Backbone Network** A high-speed communication network that carries major traffic between smaller networks [1].
- Background question** A question that asks for general information on a topic (see also: **foreground question**) [21].
- Backward chaining** Also known as goal-directed reasoning. A form of inference used in **rule-based systems** in which the **inference engine** determines whether the premise (left-hand side) of a given rule is true by invoking other rules that can conclude the values of variables that currently are unknown and that are referenced in the premise of the given rule. The process continues recursively until all rules that can supply the required values have been considered [22].
- Bag-of-words** A language model where text is represented as a collection of words, independent of each other and disregarding word order [8].
- Bandwidth** The capacity for information transmission; the number of bits that can be transmitted per unit of time [5].
- Baseline rate:population** The prevalence of the condition under consideration in the population from which the subject was selected; **individual**: The frequency, rate, or degree of a condition before an intervention or other perturbation [2].
- Basic Local Alignment Search Tool (BLAST)** An algorithm for determining optimal genetic sequence alignments based on the observations that sections of proteins are often conserved without gaps and that there are statistical analyses of the occurrence of small subsequences within larger sequences that can be used to prune the search for matching sequences in a large database [24].
- Basic research** Systematic investigation or experimentation with the goal of discovering new knowledge, often by proposing new generalizations from the results of several experiments [1].
- Basic science** The enterprise of performing basic research [1].
- Bayes' theorem** An algebraic expression often used in clinical diagnosis for calculating post-test probability of a condition (a disease, for example) if the pretest probability (prevalence) of the condition, as well as the sensitivity and specificity of the test, are known (also called Bayes' rule). Bayes' theorem

also has broad applicability in other areas of biomedical informatics where probabilistic inference is pertinent, including the interpretation of data in bioinformatics [2,3].

Bayesian diagnosis program A computer-based system that uses Bayes' theorem to assist a user in developing and refining a differential diagnosis [22].

Before-after study (aka Historically controlled study) A study in which the evaluator attempts to draw conclusions by comparing measures made during a baseline period prior to the information resource being available and measures made after it has been implemented [11].

Behaviorism A social science framework for analyzing and modifying behavior [4,23]

Belief network A diagrammatic representation used to perform probabilistic inference; an influence diagram that has only chance nodes [3, 22].

Best of breed An information technology strategy that favors the selection of individual applications based on their specific functionality rather than a single application that integrates a variety of functions [14].

Best of cluster Best of cluster became a variant of the "best of breed" strategy by selecting a single vendor for a group of similar departmental systems, such laboratory, pharmacy and radiology [14].

Bibliographic content In information retrieval, information abstracted from the original source [21].

Bibliographic database A collection of citations or pointers to the published literature [21, 26].

Binary The condition of having only two values or alternatives [5].

Biobank A repository for biological materials that collects, processes, stores, and distributes biospecimens (usually human) for use in research [25].

Biocomputation The field encompassing the modeling and simulation of tissue, cell, and genetic behavior; see **biomedical computing** [1].

Bioinformatics The study of how information is represented and transmitted in biological systems, starting at the molecular level [1,10,24]

Biomarker A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [25].

Biomed Central An independent publishing house specializing in the publication of electronic journals in biomedicine (see www.biomedcentral.com) [21].

Biomedical computing The use of computers in biology or medicine [1].

Biomedical engineering An area of engineering concerned primarily with the research and development of biomedical instrumentation and biomedical devices [1].

Biomedical informatics The interdisciplinary field that studies and pursues the effective uses of biomedical data, information, and knowledge for scientific inquiry, problem solving, and decision making, driven by efforts to improve human health [1].

Biomedical Information Science and Technology Initiative (BISTI) An initiative launched by the NIH in 2000 to make optimal use of computer science, mathematics, and technology to address problems in biology and medicine. It includes a consortium of senior-level representatives from each of the NIH institutes and centers plus representatives of other Federal agencies concerned with biocomputing. See: <http://www.bisti.nih.gov> [1]

Biomedical taxonomy A formal system for naming entities in biomedicine [4].

Biomolecular imaging A discipline at the intersection of molecular biology and *in vivo* imaging, it enables the visualisation of cellular function and the follow-up of the molecular processes in living organisms without perturbing them [1].

Biopsychosocial model A model of medical care that emphasizes not only an understanding of disease processes, but also the psychological and social conditions of the patient that affect both the disease and its therapy [22].

Biosurveillance A public health activity that monitors a population for occurrence of a rare disease or increased occurrence of a common one [8].

Bit depth The number of bits that represent an individual pixel in an image; the more

- bits, the more intensities or colors can be represented [5].
- Bit rate** The rate of information transfer; a function of the rate at which signals can be transmitted and the efficacy with which digital information is encoded in the signal [5].
- Bit** The logical atomic element for all digital computers [5]
- BLAST** See: Basic Local Alignment Search Tool [24].
- Blinding** In the context of clinical research, blinding refers to the process of obfuscating from the participant and/or investigator what study intervention a given participant is receiving. This is commonly done to reduce study biases [26].
- Blue Button** A feature of the Veteran Administration's VistA system that exports an entire patient's record in electronic form [17].
- Body of knowledge** An information resource that encapsulates the knowledge of a field or discipline [21].
- Body** The portion of a simple electronic mail message that contains the free-text content of the message [5].
- Boolean operators** The mathematical operators and, or, and not, which are used to combine index terms in information retrieval searching [21].
- Boolean searching** A search method in which search criteria are logically combined using and, or, and not operators [21].
- Bootstrap** A small set of initial instructions that is stored in read-only memory and executed each time a computer is turned on. Execution of the bootstrap is called booting the computer. By analogy, the process of starting larger computer systems [5].
- Bottom-up** An algorithm for analyzing small pieces of a problem and building them up into larger components [8].
- Bound morpheme** A morpheme that creates a different form of a word but must always occur with another morpheme (e.g., *-ed*, *-s*) [8].
- Bridge** A device that links or routes signals from one network to another [5].
- Broadband** A data-transmission technique in which multiple signals may be transmitted simultaneously, each modulated within an assigned frequency range [5].
- Browsing** Scanning a database, a list of files, or the Internet, either for a particular item or for anything that seems to be of interest [5].
- Bundled payments** In the healthcare context, refers to the practice of reimbursing providers based on the total expected costs of a particular episode of care. Generally occupies a "middle ground" between fee-for-service and capitation mechanisms [14].
- Business logic layer** A conceptual level of system architecture that insulates the applications and processing components from the underlying data and the user interfaces that access the data [14].
- Buttons** Graphic elements within a dialog box or user-selectable areas within an HTML document that, when activated, perform a specified function (such as invoking other HTML documents and services) [5].
- C statistic** The area under an receiver operating characteristic (ROC) curve [25].
- CAD** See: Computer-aided diagnosis [20].
- Canonical form** A preferred string or name for a term or collection of names; the canonical form may be determined by a set of rules (e.g., "all capital letters with words sorted in alphabetical order") or may be simply chosen arbitrarily [21].
- Capitated payments** System of health-care reimbursement in which providers are paid a fixed amount per patient to take care of all the health-needs of a population of patients [14].
- Capitation** Payments to providers, typically on an annual basis, in return for which the clinicians provide all necessary care for the patient and do not submit additional fee-for-service bills [1].
- Cardiac output** A measure of blood volume pumped out of the left or right ventricle of the heart, expressed as liters per minute [19].
- Care coordinator** See: Case Manager [15].
- Cascading finite state automata (FSA)** A tagging method in natural language processing in which a series of finite state automata are employed such that the output of one FSA becomes the input for another [8].
- Case** Refers to the capitalization of letters in a word [8].
- Case manager** A person in charge of coordinating all aspects of a patient's care [15].

- CCD** See: Continuity of Care Document [22].
- CCOW** See: Clinical Context Object Workgroup [6].
- CDC** See Centers for Disease Control [7].
- CDE** See Common Data Element [26].
- CDR** See: Clinical data repository [8,14].
- CDSS** See: Clinical decision-support system [14, 22, 26].
- CDW** See: Clinical data warehouse [8].
- Cellular imaging** Imaging methods that visualize cells [9].
- Centering theory** A theory that attempts to explain what entities are indicated by referential expressions (such as pronouns) by noting how the center (focus of attention) of each sentence changes across the text [8].
- Centers for Disease Control and Prevention (CDC)** An agency within the US Department of Health and Human Services that provides the public with health information and promotes health through partnerships with state health departments and other organizations [7].
- Central computer system** A single system that handles all computer applications in an institution using a common set of databases and interfaces [14].
- Central processing unit (CPU)** The “brain” of the computer. The CPU executes a program stored in main memory by fetching and executing instructions in the program [5].
- Central Test Node (CTN)** DICOM software to foster cooperative demonstrations by the medical imaging vendors [20].
- Certificate** Coded authorization information that can be verified by a certification authority to grant system access [5].
- Challenge evaluation** An evaluation of information systems, often in the field of information retrieval or related areas, that provides a public test collection or gold standard data collection for various researchers to compare and analyze results [21].
- Chance node** A symbol that represents a chance event. By convention, a chance node is indicated in a decision tree by a circle [3].
- Character sets and encodings** Tables of numeric values that correspond to sets of printable or displayable characters. ASCII is one example of such an encoding [8].
- Chart parsing** A dynamic programming algorithm for structuring a sentence according to grammar by saving and reusing segments of the sentence that have been parsed [8].
- Chat** A synchronous mode of text-based communication [23].
- Check tags** In MeSH, terms that represent certain facets of medical studies, such as age, gender, human or nonhuman, and type of grant support; check tags provide additional indexing of bibliographic citations in databases such as Medline [21].
- CHI** See: Consumer health informatics [1, 10, 17, 18].
- CHIN** See: Community Health Information Network [13].
- Chunking** A natural language processing method for determining non-recursive phrases where each phrase corresponds to a specific part of speech [8].
- CINAHL (or CINHL)** See: Cumulative Index to Nursing and Allied Health Literature [21].
- CINAHL Subject Headings** A set of terms based on MeSH, with additional domain-specific terms added, used for indexing the Cumulative Index to Nursing and Allied Health Literature (CINAHL) [21].
- CIS** See: Clinical information system [14].
- Citation database** A database of citations found in scientific articles, showing the linkages among articles in the scientific literature [21].
- Classification** In image processing, the categorization of segmented regions of an image based on the values of measured parameters, such as area and intensity [7].
- CLIA certification** See: Clinical Laboratory Improvement Amendments of 1988 Certification [25].
- Client-server** Information processing interaction that distributes application processing between a local computer (the client) and a remote computer resource (the server) [5].
- Clinical and translational research** A broad spectrum of research activities involving the translation of findings from initial laboratory-based studies into early-stage clinical studies, and subsequently, from the findings of those studies in clinical and/or population-level practice. This broad area incorporates multiple Biomedical Informatics sub-domains,

including both translational bioinformatics and clinical research informatics [26].

Clinical Context Object Workgroup (CCOW)

A common protocol for single sign-on implementations in health care. It allows multiple applications to be linked together, so the end user only logs in and selects a patient in one application, and those actions propagate to the other applications [6].

Clinical data repository (CDR) Clinical database optimized for storage and retrieval for individual patients and used to support patient care and daily operations [8,14].

Clinical data warehouse (CDW) A database of clinical data obtained from primary sources such as electron health records, organized for re-use for secondary purposes [8].

Clinical datum Replaces medical datum with same definition [2]

Clinical decision support Any process that provides health-care workers and patients with situation-specific knowledge that can inform their decisions regarding health and health care [22].

Clinical decision-support system (CDSS)

A computer-based system that assists physicians in making decisions about patient care [14, 22, 26].

Clinical Document Architecture An HL7 standard for naming and structuring clinical documents, such as reports [8].

Clinical expert system A computer program designed to provide decision support for diagnosis or therapy planning at a level of sophistication that an expert physician might provide [10].

Clinical guidelines Systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances [1, 13].

Clinical informatics The application of biomedical informatics methods in the patient-care domain; a combination of computer science, information science, and clinical science designed to assist in the management and processing of clinical data, information, and knowledge to support clinical practice [1].

Clinical information system (CIS) The components of a health-care information system designed to support the delivery of patient

care, including order communications, results reporting, care planning, and clinical documentation [14].

Clinical judgment Decision making by clinicians that incorporates professional experience and social, ethical, psychological, financial, and other factors in addition to the objective medical data [10].

Clinical Laboratory Improvement Amendments of 1988 certification Clinical Laboratory Improvement Amendments of 1988, establishing laboratory testing quality standards to ensure the accuracy, reliability and timeliness of patient test results, regardless of where the test was performed [25].

Clinical modifications A published set of changes to the International Classification of Diseases (ICD) that provides additional levels of detail necessary for statistical reporting in the United States [7].

Clinical pathway Disease-specific plan that identifies clinical goals, interventions, and expected outcomes by time period [1, 14].

Clinical prediction rule A rule, derived from statistical analysis of clinical observations that is used to assign a patient to a clinical subgroup with a known probability of disease [3].

Clinical research informatics (CRI) The application of biomedical informatics methods in the clinical research domain to support all aspects of clinical research, from hypothesis generation, through study design, study execution and data collection, data analysis, and dissemination of results [26].

Clinical Research Management System (CRMS)

A clinical research management system is a technology platform that supports and enables the conduct of clinical research, including clinical trials, usually through a combination of functional modules targeting the preparatory, enrollment, active, and dissemination phases of such research programs. CRMS systems are often also referred to as Clinical Trials Management Systems (CTMS), particularly when they are used to manage only clinical trials rather than various types of clinical research [26].

Clinical research The range of studies and trials in human subjects that fall into the three sub-categories: (1) Patient-oriented research:

Research conducted with human subjects (or on material of human origin such as tissues, specimens and cognitive phenomena) for which an investigator (or colleague) directly interacts with human subjects. Patient-oriented research includes: (a) mechanisms of human disease; (b) therapeutic interventions; (c) clinical trial; and (d) development of new technologies. (2) Epidemiologic and behavioral studies. (3) Outcomes research and health services research.” [10, 26]

Clinical subgroup A subset of a population in which the members have similar characteristics and symptoms, and therefore similar likelihood of disease [3].

Clinical trials Research projects that involve the direct management of patients and are generally aimed at determining optimal modes of therapy, evaluation, or other interventions [1].

Clinical-event monitors Systems that electronically and automatically record the occurrence or changes of specific clinical events, such as blood pressure, respiratory capability, or heart rhythms [14].

Clinically relevant population The population of patients that is seen in actual practice. In the context of estimating the sensitivity and specificity of a diagnostic test, that group of patients in whom the test actually will be used [3].

Closed loop medication management system A workflow process (typically supported electronically) through which medications are ordered electronically by a physician, filled by the pharmacy, delivered to the patient, administered by a nurse, and subsequently monitored for effectiveness by the physician [14].

Closed loop Regulation of a physiological variable, such as blood pressure, by monitoring the value of the variable and altering therapy without human intervention [15].

Cloud technology or computing Cloud computing is using computing resources located in a remote location. Typically, cloud computing is provided by a separate business, and the user pays for it on per usage basis. There are variations such as private clouds,

where the ‘cloud’ is provided by the same business, but leverages methods that permit easier virtualization and expandability than traditional methods. Private clouds are popular with healthcare because of security concerns with public cloud computing [20].

Clustering algorithms A method which assigns a set of objects into groups (called clusters) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters [9].

Coaching system An intelligent tutoring system that monitors the session and intervenes only when the student requests help or makes serious mistakes [23]

Cocke-Younger-Kasami (CYK) A dynamic programming method that uses bottom-up rules for parsing grammar-free text; used only in conjunction with a grammar written in Chomsky normal form [8].

Code See Machine Language [5].

Cognitive artifacts human-made materials, devices, and systems that extend people’s abilities in perceiving objects, encoding and retrieving information from memory, and problem-solving [4].

Cognitive engineering An interdisciplinary approach to the development of principles, methods and tools to assess and guide the design of computerized systems to support human performance[4]

Cognitive heuristics Mental processes by which we learn, recall, or process information; rules of thumb [3].

Cognitive load An excess of information that competes for few cognitive resources, creating a burden on working memory [4].

Cognitive science Area of research concerned with studying the processes by which people think and behave [1, 4, 23]

Cognitive task analysis The analysis of both the information-processing demands of a task and the kinds of domain-specific knowledge required performing it, used to study human performance [4].

Cognitive walkthrough (CW) An analytic method for characterizing the cognitive processes of users performing a task. The method

- is performed by an analyst or group of analysts “walking through” the sequence of actions necessary to achieve a goal, thereby seeking to identify potential usability problems that may impede the successful completion of a task or introduce complexity in a way that may frustrate users [4].
- Collaborative workspace** A virtual environment in which multiple participants can interact, synchronously or asynchronously, to perform a collaborative task [4].
- Color resolution** A measure of the ability to distinguish among different colors (indicated in a digital image by the number of bits per pixel). Three sets of multiple bits are required to specify the intensity of red, green, and blue components of each pixel color [5].
- Commodity internet** A general-purpose connection to the Internet, not configured for any particular purpose [18].
- Common Data Elements (CDEs)** Standards for data that stipulate the methods by which the data are collected and the controlled terminologies used to represent them. Many standard sets of CDEs have been developed, often overlapping in nature [26].
- Communication** Data transmission and information exchange between computers using accepted protocols via an exchange medium such as a telephone line or fiber optic cable [5].
- Community Health Information Network (CHIN)** A computer network developed for exchange of sharable health information among independent participant organizations in a geographic area (or community) [13].
- Comparative effectiveness research** A form of clinical research that compares examines outcomes of two or more interventions to determine if one is statistically superior to another [8].
- Compiler** A program that translates a program written in a high-level programming language to a machine-language program, which can then be executed [5].
- Comprehensibility and control** Security function that ensures that data owners and data stewards have effective control over information confidentiality and access [5].
- Computational biology** The science of computer-based mathematical and statistical techniques to analyze biological systems. See also **bioinformatics** [10].
- Computed check** A procedure applied to entered data that detects errors based on whether values have the correct mathematical relationship; (e.g., white blood cell differential counts, reported as percentages, must sum to 100 [12].
- Computed tomography (CT)** An imaging modality in which X rays are projected through the body from multiple angles and the resultant absorption values are analyzed by a computer to produce cross-sectional slices [9, 20].
- Computer architecture** The basic structure of a computer, including memory organization, a scheme for encoding data and instructions, and control mechanisms for performing computing operations [5].
- Computer memories** Store programs and data that are being used actively by a CPU [5].
- Computer program** A set of instructions that tells a computer which mathematical and logical operations to perform [5].
- Computer-aided diagnosis (CAD)** Any form of diagnosis in which a computer program helps suggest or rank diagnostic considerations [20].
- Computer-based (or computerized) physician order entry (CPOE)** A clinical information system that allows physicians and other clinicians to record patient-specific orders for communication to other patient care team members and to other information systems (such as test orders to laboratory systems or medication orders to pharmacy systems). Sometimes called **provider order entry** or **practitioner order entry** to emphasize such systems’ uses by clinicians other than physicians [13,14].
- Computer-based patient records (CPRs)** An early name for **electronic health records (EHRs)** dating to the early 1990s [1].
- Concept** A unit of thought made explicit through the representation of properties of an object or a set of common objects [7]. An abstract idea generalized from specific instances of objects that occur in the world [21].
- Conceptual graph** A formal notation in which knowledge is represented through explicit relationships between concepts. Graphs can be

depicted with diagrams consisting of shapes and arrows, or in a text format [8].

Conceptual knowledge Knowledge about concepts [4].

Concordant test results Test results that reflect the true patient state (true-positive and true-negative results) [3].

Conditional probability The probability of an event, contingent on the occurrence of another event [3].

Conditionally independent Two events, A and B, are conditionally independent if the occurrence of one does not influence the probability of the occurrence of the other, when both events are conditioned on a third event C. Thus, $p[A \mid B,C] = p[A \mid C]$ and $p[B \mid A,C] = p[B \mid C]$. The conditional probability of two conditionally independent events both occurring is the product of the individual conditional probabilities: $p[A,B \mid C] = p[A \mid C] \times p[B \mid C]$. For example, two tests for a disease are conditionally independent when the probability of the result of the second test does not depend on the result of the first test, given the disease state. For the case in which disease is present, $p[\text{second test positive} \mid \text{first test positive and disease present}] = p[\text{second test positive} \mid \text{first test negative and disease present}] = p[\text{second test positive} \mid \text{disease present}]$. More succinctly, the tests are conditionally independent if the sensitivity and specificity of one test do not depend on the result of the other test (See **independent**) [3].

Conditioned event A chance event, the probability of which is affected by another chance event (the **conditioning event**) [3].

Conditioning event A chance event that affects the probability of occurrence of another chance event (the conditioned event) [3].

Confidentiality The ability of data owners and data stewards to control access to or release of private information [5].

Consistency check A procedure applied to entered data that detects errors based on internal inconsistencies; e.g., recognizing a problem with the recording of cancer of the prostate as the diagnosis for a female patient [12].

Constructivism Argues that humans generate knowledge and meaning from an interaction between their experiences and their ideas [23].

Consumer health informatics (CHI) Applications of medical informatics technologies that focus on patients or healthy individuals as the primary users [1, 10, 17, 18].

Content based image retrieval Also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision techniques to the image retrieval problem, that is, the problem of searching for digital images in large databases [9].

Content In information retrieval, media developed to communicate information or knowledge [21].

Context free grammar A mathematical model of a set of strings whose members are defined as capable of being generated from a starting symbol, using rules in which a single symbol is expanded into one or more symbols [8].

Contingency table A 2×2 table that shows the relative frequencies of true-positive, true-negative, false-positive, and false-negative results [3].

Continuity of Care Document (CCD) An HL7 standard that enables specification of the patient data that relate to one or more encounters with the healthcare system. The CCD is used for interchange of patient information (e.g., within Health Information Exchanges). The format enables all the electronic information about a patient to be aggregated within a standardized data structure that can be parsed and interpreted by a variety of information systems [22].

Continuity of care The coordination of care received by a patient over time and across multiple healthcare providers [2].

Continuum of care The full spectrum of health services provided to patients, including health maintenance, primary care, acute care, critical care, rehabilitation, home care, skilled nursing care, and hospice care [15].

Contract-management system A computer system used to support managed-care contracting by estimating the costs and payments associated with potential contract terms and by comparing actual with expected payments based on contract terms [14].

Contrast resolution A metric for how well an imaging modality can distinguish small differences in signal intensity in different regions of the image [9].

- Contrast** The difference in light intensity between dark and light areas of an image [5].
- Control intervention** In the context of clinical research, a control intervention represents the intervention (e.g. placebo, standard care, etc.) given to the group of study participants assigned to the control or comparator arm of a study. Depending on the study type, the goal is to generate data as the basis of comparison with the experimental intervention of interest in order to determine the safety, efficacy, or benefits of an experimental intervention [26].
- Controlled terminology** A finite, enumerated set of terms intended to convey information unambiguously [21, 22].
- Copyright law** Protection of written materials and intellectual property from being copied verbatim [10].
- Coreference chains** Provide a compact representation for encoding the words and phrases in a text that all refer to the same entity [8].
- Coreference resolution** In natural language processing, the assignment of specific meaning to some indirect reference [8].
- Correctional Telehealth** The application of telehealth to the care of prison inmates, where physical delivery of the patient to the practitioner is impractical [18].
- Covered entities** Under the HIPAA Privacy Rule, a covered entity is an organization or individual that handles personal health information. Covered entities include providers, health plans, and clearinghouses [27].
- CPOE** See: Computer-based (or Computerized) Physician (or Provider) Order Entry [13,14].
- CPR (or CPRS)** See: Computer-based patient records [1].
- CPU** See: Central processing unit [5].
- CRI** See: Clinical research informatics [26].
- CRMS (or CRDMS)** See: Clinical Research Management System [26].
- Cryptographic encoding** Scheme for protecting data through authentication and authorization controls based on use of keys for encrypting and decrypting information [5].
- CT (or CAT)** See: Computed tomography [9, 20].
- Cumulative Index to Nursing and Allied Health Literature (CINHL)** A non-NLM bibliographic database that covers nursing and allied health literature, including physical therapy, occupational therapy, laboratory technology, health education, physician assistants, and medical records [21].
- Curly Braces Problem** The situation that arises in **Arden Syntax** where the code used to enumerate the variables required by a **medical logic module** (MLM) cannot describe how the variables actually derive their values from data in the EHR database. Each variable definition in an MLM has {curly braces} that enclose words in natural language that indicate the meaning of the corresponding variable. The particular database query required to supply a value for the variable must be specified by the local implementer, however. The curly braces problem makes it impossible for an MLM developed at one institution to operate at another without local modification [22].
- Cursor** A blinking region of a display monitor, or a symbol such as an arrow, that indicates the currently active position on the screen [5].
- CYK** See: Cocke-Younger-Kasami [8].
- Dashboard** A user-interface element that displays data produced by several computer programs simultaneously and that allows users to interact with those programs in standardized ways [20,22].
- Data buses** An electronic pathway for transferring data—for instance, between a CPU and memory [5].
- Data capture** The process of collecting data to be stored in an information system; it includes entry by a person using a keyboard and collection of data from sensors [12].
- Data Encryption Standard (DES)** A widely-used method of for securing encryption that uses a private (secret) key for encryption and requires the same key for decryption (see also, public key cryptography) [5]
- Data independence** The insulation of applications programs from changes in data-storage structures and data-access strategies [5]
- Data-interchange standards** Adopted formats and protocols for exchange of data between independent computer systems [7]
- Data layer** A conceptual level of system architecture that isolates the data collected and stored in the enterprise from the applications and user interfaces used to access those data [14].

- Data Recording** The documentation of information for archival or future use through mechanisms such as handwritten text, drawings, machine-generated traces, or photographic images [2].
- Database management system (DBMS)** An integrated set of programs that manages access to databases [5, 21].
- Database** A collection of stored data—typically organized into fields, records, and files—and an associated description (schema) [2, 5].
- Datum** Any single observation of fact. A medical datum generally can be regarded as the value of a specific parameter (for example, *red-blood-cell count*) for a particular object (for example, a patient) at a given point in time [2].
- DBMS** See: Database Management System [5].
- DCMI** See: Dublin Core Metadata Initiative [21].
- Debugger** A system program that provides traces, memory dumps, and other tools to assist programmers in locating and eliminating errors in their programs [5].
- Decision analysis** A methodology for making decisions by identifying alternatives and assessing them with regard to both the likelihood of possible outcomes and the costs and benefits of those outcomes [22].
- Decision node** A symbol that represents a choice among actions. By convention, a decision node is represented in a decision tree by a square [3].
- Decision support** The process of assisting humans in making decisions, such as interpreting clinical information or choosing a diagnostic or therapeutic action. See; **Clinical Decision Support** [4]
- Decision tree** A diagrammatic representation of the outcomes associated with chance events and voluntary actions [3].
- Deduplication** The process that matches, links, and or merges data to eliminate redundancies [16].
- De-identified aggregate data** Data reports that are summarized or altered slightly in a way that makes the discernment of the identity of any of the individuals whose data was used for the report impossible or so difficult as to be extremely improbable. The process of de-identifying aggregate data is known as statistical disclosure control [13].
- Delta check** A procedure applied to entered data that detects large and unlikely differences between the values of a new result and of the previous observations; e.g., a recorded weight that changes by 100 lb in 2 weeks [12].
- Demonstration study** Study that establishes a relation—which may be associational or causal—between a set of measured variables [11].
- Dental informatics** The application of biomedical informatics methods and techniques to problems derived from the field of dentistry. Viewed as a subarea of clinical informatics [1].
- Deoxyribonucleic acid (DNA)** The genetic material that is the basis for heredity. DNA is a long polymer chemical made of four basic subunits. The sequence in which these subunits occur in the polymer distinguishes one DNA molecule from another and in turn directs a cell's production of proteins and all other basic cellular processes [24].
- Departmental system** A system that focus on a specific niche area in the healthcare setting, such as a laboratory, pharmacy, radiology department, etc. [14]
- Dependency grammar** A linguistic theory of syntax that is based on dependency relations between words, where one word in the sentence is independent and other words are dependent on that word. Generally, the verb of a sentence is independent and other words are directly or indirectly dependent on the verb [8].
- Dependent variable (also called outcome variable)** In a correlational or experimental study, the main variable of interest or outcome variable, which is thought to be affected by or associated with the independent variables (q.v.) [11].
- Derivational morphemes** A morpheme that changes the meaning or part of the speech of a word (e.g., *-ful* as in *painful*, converting a noun to an adjective) [8].
- DES** See: Data Encryption Standard [5].
- Descriptive study** One-group study that seeks to measure the value of a variable in a sample

- of subjects. Study with no independent variable [11].
- Design validation** A study conducted to inform the design of an information resource, e.g. a user survey [11].
- Diagnosis** The process of analyzing available data to determine the pathophysiologic explanation for a patient's symptoms [1, 10, 22].
- Diagnosis-based reimbursement** Payments to providers (typically hospitals) based on the diagnosis made by a physician at the time of admission [14].
- Diagnosis-related group (DRG)** One of almost 500 categories based on major diagnosis, length of stay, secondary diagnosis, surgical procedure, age, and types of services required. Used to determine the fixed payment per case that Medicare will reimburse hospitals for providing care to elderly patients [13, 23].
- Diagnostic decision-support system** A computer-based system that assists physicians in rendering diagnoses; a subset of clinical decision-support systems. **See clinical decision-support system** [10].
- Diagnostic process** The activity of deciding which questions to ask, which tests to order, or which procedures to perform, and determining the value of the results relative to associated risks or financial costs [22].
- DICOM** *See: Digital Image Communications in Medicine* [18,20].
- Dictionary** A set of terms representing the system of concepts of a particular subject field [7].
- Differential diagnosis** The set of active hypotheses (possible diagnoses) that a physician develops when determining the source of a patient's problem [2, 20].
- Digital computer** A computer that processes discrete values based on the binary digit or bit. Essentially all modern computers are digital, but analog computers also existed in the past [5].
- Digital Image Communications in Medicine (DICOM)** A standard for electronic exchanging digital health images, such as x-rays and CT scans [18,20].
- Digital image** An image that is stored as a grid of numbers, where each picture element (pixel) in the grid represents the intensity, and possibly color, of a small area [20].
- Digital library** Organized collections of electronic content, intended for specific communities or domains [21].
- Digital object identifier (DOI)** A system for providing unique identifiers for published digital objects, consisting of a prefix that is assigned by the International DOI Foundation to the publishing entity and a suffix that is assigned and maintained by the entity [21].
- Digital radiography (DR)** The process of producing X-ray images that are stored in digital form in computer memory, rather than on film [9, 20].
- Digital signal processing (DSP)** An integrated circuit designed for high-speed data manipulation and used in audio communications, image manipulation, and other data acquisition and control applications [5].
- Digital signal** A signal that takes on discrete values from a specified range of values [5].
- Digital subscriber line (DSL)** A digital telephone service that allows high-speed network communication using conventional (twisted pair) telephone wiring [5].
- Digital subtraction angiography (DSA)** A radiologic technique for imaging blood vessels in which a digital image acquired before injection of contrast material is subtracted pixel by pixel from an image acquired after injection. The resulting image shows only the differences in the two images, highlighting those areas where the contrast material has accumulated [9].
- Direct entry** The entry of data into a computer system by the individual who personally made the observations [12].
- Discourse** Large portions of text forming a narrative, such as paragraphs and documents [8].
- Discrete event simulation model** A modeling approach that assesses interactions between people, typically composed of patients that have attributes and that experience events [3].
- Discussion board** An on-line [23]
- Disease** Any condition in an organism that is other than the healthy state [8].
- Dissemination phase** during the dissemination phase of a clinical research study, investigators analyze and report upon the data generated during the active phase [26].

- Distributed cognition** A view of cognition that considers groups, material artifacts, and cultures and that emphasizes the inherently social and collaborative nature of cognition [4].
- Distributed computer systems** A collection of independent computers that share data, programs, and other resources [14].
- DNA** See: Deoxyribonucleic Acid [24].
- DNS** See: Domain name system [5, 12].
- Document structure** The organization of text into sections [8].
- DOI** See: Digital object identifier [21].
- Domain name system (DNS)** A hierarchical name-management system used to translate computer names to Internet protocol (IP) addresses [5, 12].
- Doppler shift** A perceived change in frequency of a signal as the signal source moves toward or away from a signal receiver [20].
- Double blind** A clinical study methodology in which neither the researchers nor the subjects know to which study group a subject has been assigned [2].
- Double-blinded study** In the context of clinical research, a double blinded study is a study in which both the investigator and participant are blinded from the assignment of an intervention. In this scenario, a trusted third party must maintain records of such study arm assignments to inform later data analyses [26].
- Draft standard for trial use** A proposal for a standard developed by HL7 that is sufficiently well defined that early adopters can use the specification in the development of HIT. Ultimately, the draft standard may be refined and put to a ballot for endorsement by the members of the organization, thus creating an official standard [22].
- DRG** See Diagnosis-Related Groups [13, 23].
- Drug repurposing** Identifying existing drugs that may be useful for indications other than those for which they were initially approved [25].
- Drug screening robots** A scientific instrument that can perform assays with potential drugs in a highly parallel and high throughput manner [24].
- DSA** See: Digital subtraction angiography [9].
- DSL** See: Digital subscriber line [5].
- DSP** See: Digital signal processing [5].
- Dublin Core Metadata Initiative (DCMI)** A standard metadata model for indexing published documents [21].
- Dynamic** A simulation program that models changes in patient state over time and in response to students' therapeutic decisions [23].
- Dynamic programming** A computationally intensive computer-science technique used, for example, to determine optimal sequence alignments in many computational biology applications [24].
- Dynamic transmission model** A model that divides a population into compartments (for example, uninfected, infected, recovered, dead), and for transitions between compartments are governed by differential or difference equations [3].
- Earley parsing** A dynamic programming method for parsing context-free grammar [8].
- EBM** See Evidence-Based Medicine [21].
- EBM database** For **Evidence-Based Medicine database**, a highly organized collection of clinical evidence to support medical decisions based on the results of controlled clinical trials [21].
- eCRF** See: Electronic Case Report Form [26].
- EDC** See: Electronic Data Capture [26].
- EDI** See: Electronic data interchange [14].
- EEG** See: Electroencephalography [9].
- EHR** See: Electronic health record [6,10,12,13,21,26].
- EIW** See: Enterprise information warehouse [14].
- Electroencephalography (EEG)** A method for measuring the electromagnetic fields generated by the electrical activity of the neurons using a large arrays of scalp sensors, the output of which are processed to localize the source of the electrical activity inside the brain [9].
- Electronic Case Report Form (eCRF)** A computational representation of paper case report forms (CRFs) used to enable EDC [26].
- Electronic Data Capture (EDC)** EDC is the process of capturing study-related data elements via computational mechanisms [26].
- Electronic data interchange (EDI)** Electronic exchange of standard data transactions, such as claims submission and electronic funds transfer [14].

- Electronic Health Record (EHR)** A repository of electronically maintained information about an individual's lifetime health status and health care, stored such that it can serve the multiple legitimate users of the record. See also **EMR** and **CPR** [6,10,12,13,21,26].
- Electronic health record system** An electronic health record and the tools used to manage the information; also referred to as a computer-based patient-record system and often shortened to electronic health record [12].
- Electronic Medical Record (EMR)** The electronic record documenting a patient's care in a provider organization such as a hospital or a physician's office. Often used interchangeably with **Electronic Health Record (EHR)**, although EHRs refer more typically to an individual's lifetime health status and care rather than the set of particular organizationally-based experiences [14].
- Electronic-long, paper-short (ELPS)** A publication method which provides on the Web site supplemental material that did not appear in the print version of the journal [21].
- ELPS** See: Electronic-long, paper-short [21].
- EMBASE** A commercial biomedical and pharmacological database from ExcerptaMedica, which provides information about medical and drug-related subjects [21].
- Emergent design** Study where the design or plan of research can and does change as the study progresses. Characteristic of subjectivist studies [11].
- EMPI** See: Enterprise master patient index [14].
- Emotion detection** A natural language technique for determining the mental state of the author of a text document [8].
- EMTREE** A hierarchically structured, controlled vocabulary used for subject indexing, used to index **EMBASE** [21].
- EMR (or EMRS)** See: Electronic Medical Record [14].
- Enabling technology** Any technology that improves organizational processes through its use rather than on its own. Computers, for example, are useless unless "enabled" by operation systems and applications or implemented in support of work flows that might not otherwise be possible [14].
- Encryption** The process of transforming information such that its meaning is hidden, with the intent of keeping it secret, such that only those who know how to decrypt it can read it; see decryption [5].
- Enrichment analysis** A statistical method to determine whether an a priori defined set of concepts shows statistically significant over-representation in descriptions of a set of items (such as genes) compared to what one would expect based on their frequency in a reference distribution [25].
- Enrollment** during enrollment of a clinical research study, potential participants are identified and research staff determine their eligibility for participation in a study, based upon the eligibility criteria described in the study protocol. If a participant is deemed eligible to participate, there are then officially "registered" for the study. It is during this phase that in some trial designs, a process of randomization and assignment to study arms occurs [26].
- Enterprise information warehouse (EIW)** A data base in which data from clinical, financial and other operational sources are collected in order to be compared and contrasted across the enterprise [14].
- Enterprise master patient index (EMPI)** An architectural component that serves as the name authority in a health-care information system composed of multiple independent systems; the EMPI provides an index of patient names and identification numbers used by the connected information systems [14].
- Entrez** A search engine from the National Center for Biotechnology Information (NCBI), at the National Library of Medicine; Entrez can be used to search a variety of life sciences databases, including PubMed [21].
- Entry term** A synonym form for a subject heading in the Medical Subject Headings (MeSH) controlled, hierarchical vocabulary [21].
- Epidemiologic** Related to the field of epidemiology [1].
- Epidemiology** The study of the patterns, causes, and effects of health and disease conditions in defined populations [16].
- Epigenetics** Heritable phenotypes that are not encoded in DNA sequence [24, 25].

- e-prescribing** The electronic process of generating, transmitting and filling a medical prescription [27].
- Error analysis** In natural language processing, a process for determining the reasons for false-positive and false-negative errors [8].
- Escrow** Use of a trusted third party to hold cryptographic keys, computer source code, or other valuable information to protect against loss or inappropriate access [5]
- Ethernet** A network standard that uses a bus or star topology and regulates communication traffic using the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) approach [20].
- Ethnography** Set of research methodologies derived primarily from social anthropology. The basis of much of the subjectivist, qualitative evaluation approaches [11].
- ETL** See: Extract, Transform, and Load [26].
- Evaluation contract** A document describing the aims of a study, the methods to be used and resources made available, usually agreed between the evaluator and key stakeholders before the study begins [11].
- Event-Condition-Action (ECA) rule** A rule that requires some *event* (such as the availability of a new data value in the database) to cause the *condition* (premise, or left-hand side) of the rule to be evaluated. If the condition is determined to be true, then some *action* is performed. Such rules are commonly found in active database systems, and form the basis of **medical logic modules** [22].
- Evidence-based guidelines (EBM)** An approach to medical practice whereby the best possible evidence from the medical literature is incorporated in decision making. Generally such evidence is derived from controlled clinical trials [1, 10, 21].
- Exabyte** 10^{18} bytes [5].
- Exome** The entire sequence of all genes within a genome, approximately 1–3 % of the entire genome [24].
- Expected value** The value that is expected on average for a specified chance event or decision [3].
- Experimental intervention** In the context of clinical research, an experimental intervention represents the treatment or other intervention delivered to a participant assigned to the experimental arm of the study in order to determine the safety, efficacy, or benefits of that intervention [26].
- Experimental science** Systematic study characterized by posing hypotheses, designing experiments, performing analyses, and interpreting results to validate or disprove hypotheses and to suggest new hypotheses for study [1].
- Extensible markup language (XML)** A subset of the **Standard Generalized Markup Language (SGML)** from the World Wide Web Consortium (W3C), designed especially for Web documents. It allows designers to create their own custom-tailored tags, enabling the definition, transmission, validation, and interpretation of data between applications and between organizations [5, 6, 8, 21, 26].
- External router** A computer that resides on multiple networks and that can forward and translate message packets sent from a local or enterprise network to a regional network beyond the bounds of the organization [5].
- External validity** In the context of clinical research, external validity refers to the ability to generalize study results into clinical care [26].
- Extract, Transform, and Load (ETL)** ETL is the process by which source data is collected and manipulated so as to adhere to the structure and semantics of a receiving data construct, such as a data warehouse [26].
- Extrinsic evaluation** An evaluation of a component of a system based on an evaluation of the performance of the entire system [8].
- F measure** A measure of overall accuracy that is a combination of precision and recall [8].
- Factual knowledge** Knowledge of facts without necessarily having any in-depth understanding of their origin or implications [4].
- False negative** A negative result that occurs in a true situation. Examples include a desired entity that is missed by a search routine or a test result that appears normal when it should be abnormal [8].
- False-negative rate (FNR)** The probability of a negative result, given that the condition under consideration is true—for example, the probability of a negative test result in a patient who has the disease under consideration [3].

- False-negative result (FN)** A negative result when the condition under consideration is true—for example, a negative test result in a patient who has the disease under consideration [3].
- False positive** A positive result that occurs in a false situation. Examples include an inappropriate entity that is returned by a search routine or a test result that appears abnormal when it should be normal [8].
- False-positive rate (FPR)** The probability of a positive result, given that the condition under consideration is false—for example, the probability of a positive test result in a patient who does not have the disease under consideration [3].
- False-positive result (FP)** A positive result when the condition under consideration is false—for example, a positive test result in a patient who does not have the disease under consideration [3].
- FDDI** See: Fiber Distributed Data Interface [20].
- Fiber Distributed Data Interface [FDDI]** A transmission standard for local area networks operating on fiberoptic cable, providing a transmission rate of 100 Mbit/s [20].
- Feedback** In a computer-based education program, system-generated responses, such as explanations, summaries, and references, provided to further a student's progress in learning [23].
- Fee-for-service** Unrestricted system of health-care reimbursement in which payers pay provider for those services the provider has deemed necessary [14].
- Fiberoptic cable** A communication medium that uses light waves to transmit information signals [5].
- Fiducial** An object used in the field of view of an imaging system which appears in the image produced, for use as a point of reference or a measure [9].
- Field function study** Study of an information resource where the system is used in the context of ongoing health care. Study of a deployed system (cf. Laboratory study) [11].
- Field** In science, the setting, which may be multiple physical locations, where the work under study is carried out [11]. In database design, the smallest named unit of data in a database. Fields are grouped together to form records [5].
- Field user effect study** A study of the actual actions or decisions of the users of the resource [11].
- File** In a database, a collection of similar records [5].
- File format** Representation of data within a file; can refer to the method for individual characters and values (for example, **ASCII** or **binary**) or their organization within the file (for example, **XML** or text) [8].
- File server** A computer that is dedicated to storing shared or private data files [5].
- File system** An organization of files within a database or on a mass storage device [5].
- Filtering algorithms** A defined procedure applied to input data to reduce the effect of noise [5].
- Finite state automaton** An abstract, computer-based representation of the state of some entity together with a set of actions that can transform the state. Collections of finite state automata can be used to model complex systems [8].
- Fire-wall** A security system intended to protect an organization's network against external threats by preventing computers in the organization's network from communicating directly with computers external to the network, and vice versa [5].
- Flash memory card** A portable electronic storage medium that uses a semiconductor chip with a standard physical interface; a convenient method for moving data between computers [5].
- Flexnerian** one of science-based acquisition of medically relevant knowledge, followed by on-the-job apprentice-style acquisition of experience, and accompanied by evolution and expansion of the curriculum to add new fields of knowledge [23].
- Floppy disk** An inexpensive magnetic disk that can be removed from the disk-drive unit and thereby used to transfer or archive files [5].
- FM** See: Frequency modulation [5].
- fMRI** See: Functional magnetic resonance imaging [9].
- FN** See: False-negative result [3].
- Force feedback** A user interface feature in which physical sensations are transmitted to the user

- to provide a tactile sensation as part of a simulated activity. See also **Haptic feedback** [18].
- Foreground question** Question that asks for general information related to a specific patient (see also **background question**) [21].
- Form factor** Typically refers to the physical dimensions of a product. With computing devices, refers to the physical size of the device, often with specific reference to the display. For example, we would observe that the form factor of a desktop monitor is significantly larger than that of a tablet or smart phone, and therefore able to display more characters and larger graphics on the screen [14].
- Formative evaluation** An assessment of a system's behavior and capabilities conducted during the development process and used to guide future development of the system [23]
- Forward chaining** Also known as data-driven reasoning. A form of inference used in **rule-based systems** in which the **inference engine** uses newly acquired (or concluded) values of variables to invoke all rules that may reference one or more of those variables in their premises (left-hand side), thereby concluding new values for variables in the conclusions (right-hand side) of those rules. The process continues recursively until all rules whose premises may reference the variables whose values become known have been considered [22].
- FP** See: False-positive result [3].
- FPR** See: False-positive rate [3].
- Frame Relay** A high-speed network protocol designed for sending digital information over shared wide-area networks using variable length packets of information [5].
- Frame** An abstract representation of a concept or entity that consists of a set of attributes, called slots, each of which can have one or more values to represent knowledge about the entity or concept [8].
- Free morpheme** A morpheme that is a word and that does not contain another morpheme (*e.g., arm, pain*) [8].
- Frequency modulation(FM)** A signal representation in which signal values are represented as changes in frequency rather than amplitude [5].
- Front-end application** A computer program that interacts with a database-management system to retrieve and save data and to accomplish user-level tasks [5].
- Full-text content** The complete textual information contained in a bibliographic source [21].
- Functional magnetic resonance imaging (fMRI)** A **magnetic resonance imaging** method that reveals changes in blood oxygenation that occur following neural activity [9].
- Functional mapping** An imaging method that relates specific sites on images to particular physiologic functions [20].
- Gateway** A computer that resides on multiple networks and that can forward and translate message packets sent between nodes in networks running different protocols [5].
- Gbps** See: Gigabits per second [5].
- GEM** See: Guideline Element Model [22].
- GenBank** A centralized repository of protein, RNA, and DNA sequences in all species, currently maintained by the National Institutes of Health [24].
- Gene expression microarray** Study the expression of large numbers of genes with one another and create multiple variations on a genetic theme to explore the implications of changes in genome function on human disease [24].
- Gene Expression Omnibus (GEO)** A centralized database of gene expression microarray datasets [24].
- Gene Ontology(GO)** A structured controlled vocabulary used for annotating genes and proteins with molecular function. The vocabulary contains three distinct ontologies, Molecular Function, Biological Process and Cellular Component [24].
- Genes** units encoded in DNA and they are transcribed into ribonucleic acid (RNA) [24].
- Genetic data** An overarching term used to label various collections of facts about the genomes of individuals, groups or species [10].
- Genome-Wide Association Studies (GWAS)** An examination of many common genetic variants in different individuals to see if any variant is associated with a given trait, *e.g.* a disease [25].
- Genomic medicine (also known as stratified-medicine)** The management of groups of

- patients with shared biological characteristics, determined through molecular diagnostic testing, to select the best therapy in order to achieve the best possible outcome for a given group [25].
- Genomics database** An organized collection of information from gene sequencing, protein characterization, and other genomic research [21].
- Genomics** The study of all of the nucleotide sequences, including structural genes, regulatory sequences, and noncoding DNA segments, in the chromosomes of an organism [24].
- Genotype** The genetic makeup, as distinguished from the physical appearance, of an organism or a group of organisms [24].
- Genotypic** Refers to the genetic makeup of an organism [1].
- GEO** See: Gene Expression Omnibus [24].
- Geographic Information System (GIS)** A system designed to capture, store, manipulate, analyze, manage, and visually present all types of location-specific data [16].
- Gigabits per second (Gbps)** A common unit of measure for data transmission over high-speed networks [5].
- Gigabyte** 2^{30} or 1,073,741,824 bytes [5].
- GIS** See: Geographic Information System [16].
- Global processing** Computations on the entire image, without regard to specific regional content [9].
- GO** See: Gene Ontology [24].
- Gold-standard test** The test or procedure whose result is used to determine the true state of the subject—for example, a pathology test such as a biopsy used to determine a patient's true disease state [3].
- Google** A commercial search engine that provides free searching of documents on the World Wide Web [21].
- GPU** See: Graphics processing unit [5].
- Grammar** A mathematical model of a potentially infinite set of strings [8].
- Graphical user interface (GUI)** A type of environment that represents programs, files, and options by means of icons, menus, and dialog boxes on the screen [5].
- Graphics processing unit (GPU)** A computer hardware component that performs graphic displays and other highly parallel computations [5].
- Graph** In computer science, a set of *nodes* or circles connected by a set of *edges* or lines [25].
- Gray scale** A scheme for representing intensity in a black-and-white image. Multiple bits per pixel are used to represent intermediate levels of gray [5].
- Guardian Angel Proposal** A proposed structure for a lifetime, patient-centered health information system [17].
- GUI** See: Graphical user interface [5].
- Guidance** In a computer-based education program, proactive feedback, help facilities, and other tools designed to assist a student in learning the covered material [23].
- Guideline Element Model (GEM)** An XML specification for marking up textual documents that describe clinical practice guidelines. The guideline-related XML tags make it possible for information systems to determine the nature of the text that has been marked up and its role in the guideline specification [22].
- GWAS** See: Genome-Wide Association Studies [25].
- Haptic feedback** A user interface feature in which physical sensations are transmitted to the user to provide a tactile sensation as part of a simulated activity [18, 20].
- Hard disk** A magnetic disk used for data storage and typically fixed in the disk-drive unit [5].
- Hardware** The physical equipment of a computer system, including the central processing unit, memory, data-storage devices, workstations, terminals, and printers [5].
- Harmonic mean** An average of a set of weighted values in which the weights are determined by the relative importance of the contribution to the average [8].
- HCI** See: Human-computer interaction [4].
- HCO** See: Healthcare organization [14].
- Head word** The key word in a multi-word phrase that conveys the central meaning of the phrase. For example, a phrase containing adjectives and a noun, the noun is typically the head word [8].
- Header (of email)** The portion of a simple electronic mail message that contains information about the date and time of the message, the

address of the sender, the addresses of the recipients, the subject, and other optional information [5].

Health Evaluation and Logical Processing [HELP] One of the first electronic health record systems, developed at LDS Hospital in Salt Lake City, Utah. Still in use today, it was innovative for its introduction of automated alerts [19].

Health informatics Used by some as a synonym for biomedical informatics, this term is increasingly used solely to refer to applied research and practice in clinical and public health informatics [1].

Health information and communication technology (HICT) The broad spectrum of hardware and software used to capture, store and transmit health information [18].

Health Information exchange (HIE) The process of moving health information electronically among disparate health care organizations for clinical care and other purposes; or an organization that is dedicated to providing health information exchange [6,12,13].

Health Information Infrastructure (HII) The set of public and private resources, including networks, databases, and policies, for collecting, storing, and transmitting health information [13, 16].

Health Information Technology (HIT) The use of computers and communications technology in healthcare and public health settings [1].

Health Information Technology for Economic and Clinical Health (HITECH) Also referred to as HITECH Act. Passed by the Congress as Title IV of the American Recovery and Reinvestment Act of 2009 (ARRA) in 2009, established four major goals that promote the use of health information technology: (1) Develop standards for the nationwide electronic exchange and use of health information; (2) Invest \$20B in incentives to encourage doctors and hospitals to use HIT to electronically exchange patients' health information; (3) Generate \$10B in savings through improvements in quality of care and care coordination, and reductions in medical errors and duplicative care and (4) Strengthen Federal privacy and security law to protect identifiable health information from misuse. Also codified the Office of the

National Coordinator for Health Information Technology (ONC) within the Department of Health and Human Services [14, 17].

Health Insurance Portability and Accountability Act (HIPAA) A law enacted in 1996 to protect health insurance coverage for workers and their families when they change or lose their jobs. An "administrative simplification" provision requires the Department of Health and Human Services to establish national standards for electronic healthcare transactions and national identifiers for providers, health plans, and employers. It also addresses the security and privacy of health data [1, 5, 8, 14, 18].

Health Level Seven (HL7) An ad hoc standards group formed to develop standards for exchange of health care data between independent computer applications; more specifically, the health care data messaging standard developed and adopted by the HL7 standards group [6, 8, 12, 14].

Health literacy A constellation of skills, including the ability to perform basic reading, math, and everyday health tasks like comprehending prescription bottles and appointment slips, required to function in the health care environment [4].

Health maintenance organization (HMO) A group practice or affiliation of independent practitioners that contracts with patients to provide comprehensive health care for a fixed periodic payment specified in advance [14].

Health on the Net[HON] A private organization establishing ethical standards for health information published on the World Wide Web [21].

Health Record Bank (HRB) An independent organization that provides a secure electronic repository for storing and maintaining an individual's lifetime health and medical records from multiple sources and assuring that the individual always has complete control over who accesses their information [13].

Healthcare organization (HCO) Any health-related organization that is involved in direct patient care [14].

Healthcare team A coordinated group of health professionals including physicians, nurses, case managers, dietitians, pharmacists, thera-

- pists, and other practitioners who collaborate in caring for a patient [2].
- HELP** See Health Evaluation and Logical Processing [19].
- HELP sector** A decision rule encoded in the HELP system, a clinical information system that was developed by researchers at LDS Hospital in Salt Lake City [22].
- Helper (plug-in)** An application that are launched by a Web browser when the browser downloads a file that the browser is not able to process itself [5].
- Heuristic evaluation (HE)** A usability inspection method, in which the system is evaluated on the basis of a small set of well-tested design principles such as visibility of system status, user control and freedom, consistency and standards, flexibility and efficiency of use [4].
- Heuristic** A mental “trick” or rule of thumb; a cognitive process used in learning or problem solving [2].
- HICT** See: Health information and communication technology [18].
- HIE** See: Health Information Exchange [6,12,13].
- Hierarchical** An arrangement between entities that conveys some superior-inferior relationship, such as parent–child, whole-part etc. [21]
- High-bandwidth** An information channel that is capable of carrying delivering data at a relatively high rate [18].
- Higher-level process** A complex process comprising multiple lower-level processes [1].
- HII** See: Health Information Infrastructure [13, 16].
- Hindsight bias** The tendency to over-estimate the prior predictability of an event, once the events has already taken place. For example, if event A occurs before event B, there may be an assumption that A predicted B [4].
- HIPAA** See: Health Insurance Portability and Accountability Act [1, 5, 8, 14, 18].
- HIS** See: Hospital information system [1,14].
- Historical control** In the context of clinical research, historical controls are subjects who represent the targeted population of interest for a study. Typically, their data are derived from existing resources in a retrospective manner and that represent targeted outcomes in a non-interventional state (often resulting from standard of care practices), so as to provide the basis for comparison to data sets derived from participants who have received an experimental intervention under study [26].
- Historically controlled study** See: **before-after study** [11].
- HIT** See: Health Information Technology [1].
- HITECH** See: Health Information Technology for Economic and Clinical Health [14, 17].
- HITECH regulations** The components of the Health Information Technology for Economic and Clinical Health Act, passed by the Congress in 2009, which authorized financial incentives to be paid to eligible physicians and hospitals for the adoption of “meaningful use” of EHRs in the United States. The law also called for the certification of EHR technology and for educational programs to enhance its dissemination and adoption [22].
- HIV** See: Human immunodeficiency virus [5].
- HL7** See: Health Level 7 [6,7,8,12,14].
- HMO** See: Health maintenance organization [14].
- Home Telehealth** The extension of telehealth services in to the home setting to support activities such as home nursing care and chronic disease management [18].
- HON** See: Health on the Net [21].
- Hospital information system (HIS)** Computer system designed to support the comprehensive information requirements of hospitals and medical centers, including patient, clinical, ancillary, and financial management [1,14].
- Hot fail over** A secondary computer system that is kept in constant synchronization with the primary system and that can take over as soon as the primary fails for any reason [12].
- Hounsfield number** The numeric information contained in each pixel of a CT image. It is related to the composition and nature of the tissue imaged and is used to represent the density of tissue [9].
- HRB** See: Health Record Bank [13].
- HTML** See HyperText [5, 8, 21].
- HTTP** See: HyperText Transfer Protocol [5, 6].
- Human-computer interaction (HCI)** Formal methods for addressing the ways in which human beings and computer programs exchange information [4].
- Human factors** The scientific discipline concerned with the understanding of interactions

among humans and other elements of a system, and the profession that applies theory, principles, data, and other methods to design in order to optimize human well-being and overall system performance [4].

Human Genome Project An international undertaking, the goal of which is to determine the complete sequence of human deoxyribonucleic acid (DNA), as it is encoded in each of the 23 chromosomes [2, 24].

Human immunodeficiency virus (HIV) A retrovirus that invades and inactivates helper T cells of the immune system and is a cause of AIDS and AIDS-related complex [5].

HyperText markup language (HTML) The document specification language used for documents on the World Wide Web [5, 8, 21].

HyperText Transfer Protocol (HTTP) The client-server protocol used to access information on the World Wide Web [5, 6].

Hypertext Text linked together in a nonsequential web of associations. Users can traverse highlighted portions of text to retrieve additional related information [5].

Hypothesis generation The process of proposing a hypothesis, usually driven by some unexplained phenomenon and the derivation of a suspected underlying mechanism [8].

Hypothetico-deductive approach A method of reasoning made up of four stages (cue acquisition, hypothesis generation, cue interpretation, and hypothesis evaluation) which is used to generate and test hypotheses. In clinical medicine, an iterative approach to diagnosis in which physicians perform sequential, staged data collection, data interpretation, and hypothesis generation to determine and refine a differential diagnosis [2,4].

Hypothetico-deductive reasoning reasoning by first generating and then testing a set of hypotheses to account for clinical data (i.e., reasoning from hypothesis to data) [4].

ICANN See: Internet Corporation for Assigned Names and Numbers [5].

ICD-9-CM See: International Classification of Diseases, 9th Edition, Clinical Modifications [7, 8, 14].

ICMP See: Internet Control Message Protocol [5].

Icon In a graphical interface, a pictorial representation of an object or function [5].

ICT See: Information and communications technology [12].

IDF See: Inverse document frequency [21].

IDN See: Integrated delivery network [10, 14, 15].

Image acquisition The process of generating images from the modality and converting them to digital form if they are not intrinsically digital [9].

Image compression A mathematical process for removing redundant or relatively unimportant information from an electronic image such that the resulting file appears the same (lossless compression) or similar (lossy compression) when compared to the original [18].

Image content representation Makes the information in images accessible to machines for processing [9].

Image database An organized collection of clinical image files, such as x-rays, photographs, and microscopic images [21].

Image enhancement The use of global processing to improve the appearance of the image either for human use or for subsequent processing by computer [9].

Image interpretation/computer reasoning The process by which the individual viewing the image renders an impression of the medical significance of the results of imaging study, potentially aided by computer methods [9].

Image management/storage Methods for storing, transmitting, displaying, retrieving, and organizing images [9]. The application of methods for storing, transmitting, displaying, retrieving, and organizing images.

Image metadata Data about images, such as the type of image (e.g., modality), patient that was imaged, date of imaging, image features (quantitative or qualitative), and other information pertaining to the image and its contents [9].

Image processing The transformation of one or more input images, either into one or more output images, or into an abstract representation of the contents of the input images [9].

Image quantitation The process of extracting useful numerical parameters or deriving calculations from the image or from ROIs in the image [9].

Image reasoning Computerized methods that use images to formulate conclusions or answer

questions that require knowledge and logical inference [9].

Image rendering/visualization A variety of techniques for creating image displays, diagrams, or animations to display images more in a different perspective from the raw images [9]

Imaging informatics A subdiscipline of medical informatics concerned with the common issues that arise in all image modalities and applications once the images are converted to digital form [1, 20].

IMIA See: International Medical Informatics Association [15].

Immersive environment A computer-based set of sensory inputs and outputs that can give the illusion of being in a different physical environment; see; **Virtual Reality** [4].

Immersive simulated environment A teaching environment in which a student manipulates tools to control simulated instruments, producing visual, pressure, and other feedback to the tool controls and instruments [23].

Immunization Information System (IIS) Confidential, population based, computerized databases that record all immunization doses administered by participating providers to persons residing within a given geopolitical area. Also known as **Immunization Registries** [16].

Immunization Registry Confidential, population based, computerized databases that record all immunization doses administered by participating providers to persons residing within a given geopolitical area. Also known as **Immunization Information Systems** [16].

Implementation science Implementation science refers to the study of socio-cultural, operational, and behavioral norms and processes surrounding the dissemination and adoption of new systems, approaches and/or knowledge [26].

Inaccessibility A property of paper records that describes the inability to access the record by more than one person or in more than one place at a time [12].

Incrementalist An approach to evaluation that tolerates ambiguity and uncertainty and allows changes from day-to-day [11].

Independent variable In a correlational or experimental study, a variable thought to

determine or be associated with the value of the dependent variable (q.v.) [11].

Independent Two events, A and B, are considered independent if the occurrence of one does not influence the probability of the occurrence of the other. Thus, $p[A | B] = p[A]$. The probability of two independent events A and B both occurring is given by the product of the individual probabilities: $p[A, B] = p[A] \times p[B]$. (See conditional independence.) [3]

Index In information retrieval, a shorthand guide to the content that allows users to find relevant content quickly [8].

Index Medicus The printed index used to catalog the medical literature. Journal articles are indexed by author name and subject heading, then aggregated in bound volumes. The Medline database was originally constructed as an online version of the Index Medicus [21].

Index test The diagnostic test whose performance is being measured [3].

Indexing In information retrieval, the assignment to each document of specific terms that indicate the subject matter of the document and that are used in searching [21].

Indirect-care Activities of health professionals that are not directly related to patient care, such as teaching and supervising students, continuing education, and attending staff meetings [15].

Inference engine A computer program that reasons about a **knowledge base**. In the case of **rule-based systems**, the inference engine may perform **forward chaining** or **backward chaining** to enable the rules to infer new information about the current situation [22].

Inflectional morpheme A morpheme that creates a different form of a word without changing the meaning of the word or the part of speech (e.g., *-ed*, *-s*, *-ing* as in *activated*, *activates*, *activating*.) [8]

Influence diagram A belief network in which explicit decision and utility nodes are also incorporated [3, 22].

Infobutton manager **Middleware** that provides a standard software interface between **infobuttons** in an EHR and the documents and other information resources that the infobuttons may display for the user [22].

- Infobutton** A context-specific link from health care application to some information resource that anticipates users' needs and provides targeted information [6,12, 22, 23].
- infoRAD** The information technology and computing oriented component of the very large exhibition hall at the annual meeting of the Radiological Society of North America [20].
- Information** Organized data from which knowledge can be derived and that accordingly provide a basis for decision making [2].
- Information and communications technology (ICT)** The use of computers and communications devices to accept, store, transmit, and manipulate data; the term is roughly a synonym for information technology, but it is used more often outside the United States [12].
- Information extraction** Methods that process text to capture and organize specific information in the text and also to capture and organize specific relations between the pieces of information [8, 21].
- Information model** A representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse. It can provide sharable, stable, and organized structure of information requirements for the domain context [9].
- Information need** In information retrieval, the searchers' expression, in their own language, of the information that they desire [21].
- Information resource** Generic term for a computer-based system that seeks to enhance health care by providing patient-specific information directly to care providers (often used equivalently with "system") [11].
- Information retrieval (IR)** Methods that efficiently and effectively search and obtain data, particularly text, from very large collections or databases. It is also the science and practice of identification and efficient use of recorded media. See also **Search** [4, 8, 21, 26].
- Information science** The field of study concerned with issues related to the management of both paper-based and electronically stored information [1].
- Information theory** The theory and mathematics underlying the processes of communication [1].
- Information visualization** The use of computer-supported, interactive, visual representations of abstract data to amplify cognition [4].
- Ink-jet printer** Output device that uses a moveable head to spray liquid ink on paper; the head moves back and forth for each line of pixels [5].
- Input and Output** Devices, such as keyboards, pointing devices, video displays, and laser printers, that facilitate user interaction and storage [5] or just **Input**: The data that represent state information, to be stored and processed to produce results – (output).
- Institute of Medicine** The health arm of the National Academy of Sciences, which provides unbiased, authoritative advice to decision makers and the public [17].
- Institutional Review Board (IRB)** A committee responsible for reviewing an institution's research projects involving human subjects in order to protect their safety, rights, and welfare [5, 10].
- Integrated circuit** A circuit of transistors, resistors, and capacitors constructed on a single chip and interconnected to perform a specific function [5, 19].
- Integrated delivery network (IDN)** A large conglomerate health-care organization developed to provide and manage comprehensive health-care services [10, 14, 15].
- Integrated Service Digital Network (ISDN)** A digital telephone service that allows high-speed network communications using conventional (twisted pair) telephone wiring [5, 18].
- Integrative model** Model for understanding a phenomenon that draws from multiple disciplines and is not necessarily based on first principles [24].
- Intellectual property** Software programs, knowledge bases, Internet pages, and other creative assets that require protection against copying and other unauthorized use [10, 21].
- Interactome** The set of all molecular interactions in a cell [25].
- Interface engine** Software that mediates the exchange of information among two or more systems. Typically, each system must know how to communicate with the interface

engine, but not need to know the information format of the other systems [14].

Intermediate effect process of continually learning, re-learning, and exercising new knowledge, punctuated by periods of apparent decrease in mastery and declines in performance, which may be necessary for learning to take place. People at intermediate levels of expertise may perform more poorly than those at lower level of expertise on some tasks, due to the challenges of assimilating new knowledge or skills over the course of the learning process [4].

Internal validity in the context of clinical research, internal validity refers to the minimization of potential biases during the design and execution of the trial [26].

International Classification of Diseases, 9th Edition, Clinical Modifications A US extension of the World Health Organization's International Classification of Diseases, 9th Edition [7, 8, 14].

International Medical Informatics Association (IMIA) An international organization dedicated to advancing biomedical and health informatics; an "organization of organizations", it's members are national informatics societies and organizations, such as AMIA [15].

International Organization for Standards (ISO) The international body for information and other standards [21].

Internet address See Internet Protocol Address [5]

Internet Protocol address A 32-bit number that uniquely identifies a computer connected to the Internet. Also called "Internet address" or "IP address" [5]

Internet Control Message Protocol (ICMP) A network-level Internet protocol that provides error correction and other information relevant to processing data packets [5].

Internet Corporation for Assigned Names and Numbers (ICANN) The organization responsible for managing Internet domain name and IP address assignments [5].

Internet protocol The protocol within TCP/IP that governs the creation and routing of data packets and their reassembly into data messages [5].

Internet service provider (ISP) A commercial communications company that supplies fee-for-service Internet connectivity to individuals and organizations [5].

Internet standards The set of conventions and protocols all Internet participants use to enable effective data communications [5].

Internet Support Group (ISG) An on-line forum for people with similar problems, challenges or conditions to share supportive resources [17].

Internet A worldwide collection of gateways and networks that communicate with each other using the TCP/IP protocol, collectively providing a range of services including electronic mail and World Wide Web access [5].

Interoperability The ability for systems to exchange data and operate in a coordinated, seamless manner [14, 21].

Interpreter A program that converts each statement in a high-level program to a machine-language representation and then executes the binary instruction(s) [5].

Interventional radiology A subspecialty of radiology that uses imaging to guide invasive diagnostic or therapeutic procedures [20].

Intrinsic evaluation An evaluation of a component of a system that focuses only on the performance of the component. See also **Extrinsic Evaluation** [8].

Intuitionist-pluralist or de-constructivist A philosophical position that holds that there is no truth and that there are as many legitimate interpretations of observed phenomena as there are observers [11].

Inverse document frequency (IDF) A measure of how infrequently a term occurs in a document collection [21].

$$IDF_i = \log \left(\frac{\text{number of documents}}{\text{number of documents with term } i} \right) + 1$$

IOM See: Institute of Medicine [17]

IP address See: Internet Protocol Address [5]

IR See: Information retrieval [4, 8, 21, 26]

IRB See: Institutional Review Board [5, 10].

ISDN See: Integrated Service Digital Network [5, 18].

ISG See: Internet Support Group [17].

- ISO** See: International Organization for Standards [21].
- ISP** See: Internet service provider [5].
- Job** A set of tasks submitted by a user for processing by a computer system [5].
- Joint Commission (JC)** An independent, not-for-profit organization, The Joint Commission accredits and certifies more than 19,000 health care organizations and programs in the United States. Joint Commission accreditation and certification is recognized nationwide as a symbol of quality that reflects an organization's commitment to meeting certain performance standards. The Joint Commission was formerly known as JCAHO (the Joint Commission for the Accreditation of Healthcare Organizations) [7,14].
- Just-in-time learning** An approach to providing necessary information to a user at the moment it is needed, usually through anticipation of the need [23].
- Kernel** The core of the operating system that resides in memory and runs in the background to supervise and control the execution of all other programs and direct operation of the hardware [5].
- Key field** A field in the record of a file that uniquely identifies the record within the file [5].
- Key Performance Indicator (KPI)** A metric defined to be an important factor in the success of an organization. Typically, several Key Performance indicators are displayed on a Dashboard [20].
- Keyboard** A data-input device used to enter alphanumeric characters through typing [5].
- Keyword** A word or phrase that conveys special meaning or to refer to information that is relevant to such a meaning (as in an index) [8].
- Kilobyte** 2¹⁰ or 1024 bytes [5].
- Knowledge acquisition** The information-elicitation and modeling process by which developers interact with subject-matter experts to create electronic **knowledge bases** [22].
- Knowledge base** A collection of stored facts, heuristics, and models that can be used for problem solving [2, 22].
- Knowledge** Relationships, facts, assumptions, heuristics, and models derived through the formal or informal analysis (or interpretation) of observations and resulting information [2].
- Knowledge-based system** A program that symbolically encodes, in a knowledge base, facts, heuristics, and models derived from experts in a field and uses that knowledge to provide problem analysis or advice that the expert might have provided if asked the same question [22].
- KPI** See: Key Performance Indicator [20].
- Laboratory function study** Study that explores important properties of an information resource in isolation from the clinical setting [11].
- Laboratory user effect study** An evaluation technique in which a user is observed when given a simulated task to perform [11].
- LAN** See: Local-area network [5, 14].
- Laser printer** Output device that uses an electromechanically controlled laser beam to generate an image on a xerographic surface, which then is used to produce paper copies [5].
- Latency** The time required for a signal to travel between two points in a network [18]
- Latent failures** Enduring systemic problems that make errors possible but are less visible or not evident for some time [4].
- Law of proximity** Principle from Gestalt psychology that states that visual entities that are close together are perceptually grouped [4].
- Law of symmetry** Principle from Gestalt psychology that states that symmetric objects are more readily perceived [4].
- LCD** See: Liquid crystal display [5].
- Lean** A management strategy that focuses only on those process that are able to contribute specific and measurable value for the end customer. The LEAN concept originated with Toyota's focus on efficient manufacturing processes [14].
- Learning health system** A proposed model for health care in which outcomes from past and current patient care provide are systematically collected, analyzed and then fed back into decision making about best practices for future patient care [26].
- Learning healthcare system** The cycle related to turning healthcare data into knowledge, translating that knowledge into practice, and creating new data, typically through the use of advanced information technology [1].

- LED** See: Light-emitting diode [5].
- Lexemes** A minimal lexical unit in a language that represents different forms of the same word [8].
- Lexical-statistical retrieval** Retrieval based on a combination of word matching and relevance ranking [21].
- Lexicon** A catalogue of the words in a language, usually containing syntactic information such as parts of speech, pluralization rules, etc. [8, 20]
- Light-emitting diode (LED)** A semiconductor device that emits a particular frequency of light when a current is passed through it; typically used for indicator lights and computer screens because low power requirement, minimal heat generated, and durability [5].
- Likelihood ratio (LR)** A measure of the discriminatory power of a test. The LR is the ratio of the probability of a result when the condition under consideration is true to the probability of a result when the condition under consideration is false (for example, the probability of a result in a diseased patient to the probability of a result in a nondiseased patient). The LR for a positive test is the ratio of true-positive rate (TPR) to false-positive rate (FPR) [3].
- Link-based** An indexing approach that gives relevance weight to web pages based on how often they are cited by other pages [21].
- Linux** An **open source** operating system based on principles of Unix and first developed by Linus Torvalds in 1991 [6].
- Liquid crystal display(LCD)** A display technology that uses rod-shaped molecules to bend light and alter contrast and viewing angle to produce images [5].
- Listserver** A distribution list for electronic mail messages [5].
- Literature reference database** See: bibliographic database [21]
- Local-area network (LAN)** A network for data communication that connects multiple nodes, all typically owned by a single institution and located within a small geographic area [5, 14].
- Logical Observations, Identifiers, Names and Codes [LOINC]** A controlled terminology created for providing coded terms for observational procedures. Originally focused on laboratory tests, it has expanded to include many other diagnostic procedures [17].
- Logical positivist** A philosophical position that holds that there is a single truth that can be inferred from the right combination of studies [11].
- Logic-based** A knowledge representation method based on the use of predicates [8].
- LOINC** See: Logical Observations, Identifiers, Names and Codes [17].
- Long-term memory** The part of memory that acquires information from short-term memory and retains it for long periods of time [4].
- Long-term storage** A medium for storing information that can persist over long periods without the need for a power supply to maintain data integrity [5].
- Lossless compression** A mathematical technique for reducing the number of bits needed to store data while still allowing for the recreation of the original data [20].
- Lossy compression** A mathematical technique for reducing the number of bits needed to store data but that results in loss of information [20].
- Low-level processes** An elementary process that has its basis in the physical world of chemistry or physics [1].
- LR** See: Likelihood ratio [3].
- Machine code** The set of primitive instructions to a computer represented in binary code (machine language) [5].
- Machine language** The set of primitive instructions represented in binary code (machine code) [5].
- Machine learning** A computing technique in which information learned from data is used to improve system performance [8,19].
- Machine translation** Automatic mapping of text written in one natural language into text of another language [8].
- Macros** A reusable set of computer instructions, generally for a repetitive task [5].
- Magnetic disk** A round, flat plate of material that can accept and store magnetic charge. Data are encoded on magnetic disk as sequences of charges on concentric tracks [5].
- Magnetic resonance imaging (MRI)** A modality that produces images by evaluating the differential response of atomic nuclei in the body

when the patient is placed in an intense magnetic field [9, 20].

Magnetic resonance spectroscopy A noninvasive technique that is similar to magnetic resonance imaging but uses a stronger field and is used to monitor body chemistry (as in metabolism or blood flow) rather than anatomical structures [9].

Magnetic tape A long ribbon of material that can accept and store magnetic charge. Data are encoded on magnetic tape as sequences of charges along longitudinal tracks [5].

Magnetoencephalography (MEG) A method for measuring the electromagnetic fields generated by the electrical activity of the neurons using a large arrays of scalp sensors, the output of which are processed in a similar way to CT in order to localize the source of the electromagnetic and metabolic shifts occurring in the brain during trauma [9].

Mailing list A set of mailing addresses used for bulk distribution of electronic or physical mail [5].

Mainframe computer system A large, expensive, multi-user computer, typically operated and maintained by professional computing personnel. Often referred to as a “mainframe” for short [5,19].

Malpractice Class of litigation in health care based on negligence theory; failure of a health professional to render proper services in keeping with the standards of the community [10].

Management The process of treating a patient (or allowing the condition to resolve on its own) once the medical diagnosis has been determined [22].

Manual indexing The process by which human indexers, usually using standardized terminology, assign indexing terms and attributes to documents, often following a specific protocol [21].

Markov cycle The period of time specified for a transition probability within a Markov model [3].

Markov model A mathematical model of a set of strings in which the probability of a given symbol occurring depends on the identity of the immediately preceding symbol or the two immediately preceding symbols. Processes modeled in this way are often called **Markov processes** [3].

Markov process A mathematical model of a set of strings in which the probability of a given symbol occurring depends on the identity of the immediately preceding symbol or the two immediately preceding symbols [8].

Markup language A document specification language that identifies and labels the components of the document’s contents [24].

Master patient index (MPI) A database that is used across a healthcare organization to maintain consistent, accurate, and current demographic and essential clinical data on the patients seen and managed within its various departments [14].

Mean average precision (MAP) A method for measuring overall retrieval precision in which precision is measured at every point at which a relevant document is obtained, and the MAP measure is found by averaging these points for the whole query [21].

Mean time between failures (MTBF) The average predicted time interval between anticipated operational malfunctions of a system, based on long-term observations [13].

Meaningful use The set of standards defined by the Centers for Medicare & Medicaid Services (CMS) Incentive Programs that governs the use of electronic health records and allows eligible providers and hospitals to earn incentive payments by meeting specific criteria. The term refers to the belief that health care providers using electronic health records in a meaningful, or effective, way will be able to improve health care quality and efficiency [1, 6, 12, 14, 16, 17, 22, 27].

Measurement study Study to determine the extent and nature of the errors with which a measurement is made using a specific instrument (cf. Demonstration study) [11].

Measures of concordance Measures of agreement in test performance: the true-positive and true-negative rates [3].

MedBiquitous A healthcare-specific standards consortium led by Johns Hopkins Medicine [23].

Medical computer science The subdivision of computer science that applies the methods of computing to medical topics [1].

Medical computing The application of methods of computing to medical topics (see **medical computer science**) [1].

- Medical entities dictionary (MED)** A compendium of terms found in electronic medical record systems. Among the best known MEDs is that developed and maintained by the Columbia University Medical Center and Columbia University. Contains in excess of 100,000 terms [14].
- Medical errors** Errors or mistakes, committed by health professionals, that hold the potential to result in harm to the patient [4].
- Medical home** A primary care practice that will maintain a comprehensive problem list to make fully informed decisions in coordinating their care [15].
- Medical informatics** An earlier term for the **biomedical informatics** discipline, medical informatics is now viewed as the subfield of clinical informatics that deals with the management of disease and the role of physicians [1].
- Medical Information Bus (MIB)** A data-communication system that supports data acquisition from a variety of independent devices [19].
- Medical information science** The field of study concerned with issues related to the management and use of biomedical information (see also **biomedical informatics**) [1].
- Medical Literature Analysis and Retrieval System (MEDLARS)** The initial electronic version of Index Medicus developed by the National Library of Medicine [21].
- Medical Logic Module (MLM)** A single chunk of medical reasoning or decision rule, typically encoded using the **Arden Syntax** [6, 22].
- Medical record committees** An institutional panel charged with ensuring appropriate use of medical records within the organization [10].
- Medical Subject Headings (MeSH)** Some 18,000 terms used to identify the subject content of the biomedical literature. The National Library of Medicine's MeSH vocabulary has emerged as the de facto standard for biomedical indexing [21].
- Medication** A substance used for medical treatment, typically a medicine or drug [8].
- MEDLARS Online (MEDLINE)** The National Library of Medicine's electronic catalog of the biomedical literature, which includes information abstracted from journal articles, including author names, article title, journal source, publication date, abstract, and medical subject headings [21].
- Medline Plus** An online resource from the National Library of Medicine that contains health topics, drug information, medical dictionaries, directories, and other resources, organized for use by health care consumers [21, 23].
- Megabits per second (Mbps)** A common unit of measure for specifying a rate of data transmission [5].
- Megabyte** 2^{20} or 1,048,576 bytes [5].
- Member checking** In subjectivist research, the process of reflecting preliminary findings back to individuals in the setting under study, one way of confirming that the findings are truthful [11].
- Memorandum of understanding** A document describing a bilateral or multilateral agreement between two or more parties. It expresses a convergence of will between the parties, indicating an intended common line of action [11].
- Memory sticks** A portable device that typically plugs into a computer's USB port and is capable of storing data. Also called a "thumb drive" or a "USB drive" [14].
- Memory** Areas that are used to store programs and data. The computer's working memory comprises read-only memory (ROM) and random-access memory (RAM) [5].
- Mental images** A form of internal representation that captures perceptual information recovered from the environment [4].
- Mental models** A construct for describing how individuals form internal models of systems. They are designed to answer questions such as "how does it work?" or "what will happen if I take the following action?" [4]
- Mental representations** internal cognitive states that have a certain correspondence with the external world [4].
- Menu** In a user interface, a displayed list of valid commands or options from which a user may choose [5].
- Merck Medicus** An aggregated set of resources, including Harrison's Online, MDCConsult, and DXplain [21].
- Meta-analysis** A summary study that combines quantitatively the estimates from individual studies [3].

- Metabolomic** Pertaining to the study of small-molecule metabolites created as the end-products of specific cellular processes [25].
- Metadata** Literally, data about data, describing the format and meaning of a set of data [5, 21, 25].
- Metagenomics** Using DNA sequencing technology to characterize complex samples derived from an environmental sample, e.g., microbial populations. For example, the gut ‘microbiome’ can be characterized by applying next generation sequencing of stool samples [24, 25].
- Metathesaurus** One component of the Unified Medical Language System, the Metathesaurus contains linkages between terms in Medical Subject Headings (MeSH) and in dozens of controlled vocabularies [14, 21].
- MIB** See Medical Information Bus [7].
- Microarray chips** A microchip that holds DNA probes that can recognize DNA from samples being tested [24].
- Microprocessor** An integrated circuit that contains all the functions of a central processing unit of a computer [19].
- Microsimulation models** Individual-level health state transition models that provide a means to model very complex events flexibly over time [3].
- MIMIC II Database** See Multiparameter Intelligent Monitoring in Intensive Care [19].
- Minicomputers** A class of computers that were introduced in the 1960s as a smaller alternative to mainframe computers. Minicomputers enabled smaller companies and departments within organizations (like HCOs) to implement software applications at significantly less cost than was required by mainframe computers [14, 19].
- Mistake** Occurs when an inappropriate course of action reflects erroneous judgment or inference [4].
- Mixed-initiative dialog** A mode of interaction with a computer system in which the computer may pose questions for the user to answer, and vice versa [22].
- Mixed-initiative systems** An educational program in which user and program share control of the interaction. Usually, the program guides the interaction, but the student can assume control and digress when new questions arise during a study session [23].
- Mobile health** The practice of medicine and public health supported by mobile devices. Also referred to as mHealth or m-health [18].
- Model organism databases** Organized reference databases the combine bibliographic databases, full text, and databases of sequences, structure, and function for organisms whose genomic data has been highly characterized, such as the mouse, fruit fly, and *Sarcomyces* yeast [21].
- Modem** A device used to modulate and demodulate digital signals for transmission to a remote computer over telephone lines; converts digital data to audible analog signals, and vice versa [5].
- Modifiers of interest** In natural language processing, a term that is used to describe or otherwise modify a named-entity that has been recognized [8].
- Molecular imaging** A technique for capturing images at the cellular and subcellular level by marking particular chemicals in ways that can be detected with image or radiodetection [9].
- Monitoring tool** The application of logical rules and conditions (e.g., range-checking, enforcement of data completion, etc.) to ensure the completeness and quality of research-related data [26].
- Monotonic** Describes a function that consistently increases or decreases, rather than oscillates [4].
- Morpheme** The smallest unit in the grammar of a language which has a meaning or a linguistic function; it can be a root of a word (e.g., *-arm*), a prefix (e.g., *re-*), or a suffix (e.g., *-it is*) [8].
- Morphology** The study of meaningful units in language and how they combine to form words [8].
- Morphometrics** The quantitative study of growth and development, a research area that depends on the use of imaging methods [20].
- Mosaic** The first graphical web browser credited with popularizing the World Wide Web and developed at the National Center for Supercomputing Applications (NCSA) at the University of Illinois [6].
- Motion artifact** Visual interference caused by the difference between the frame rate of an

- imaging device and the motion of the object being imaged [18].
- Mouse** A small boxlike device that is moved on a flat surface to position a cursor on the screen of a display monitor. A user can select and mark data for entry by depressing buttons on the mouse [5].
- Multi-axial** A terminology system composed of several distinct, mutually exclusive term subsets that are combined to support **postcoordination** [7].
- Multimodal interface** A design concept which allows users to interact with computers using multiple modes of communication or tools, including speaking, clicking, or touchscreen input [4].
- Multiparameter Intelligent Monitoring in Intensive Care (MIMIC-II)** A publicly and freely available research database that encompasses a diverse and very large population of ICU patients. It contains high temporal resolution data including lab results, electronic documentation, and bedside monitor trends and waveforms [19].
- Multiprocessing** The use of multiple processors in a single computer system to increase the power of the system (see parallel processing) [5].
- Multiprogramming** A scheme by which multiple programs simultaneously reside in the main memory of a single central processing unit [5].
- Multiprotocol label switching (MPLS)** A mechanism in high-performance telecommunications networks that directs data from one network node to the next based on short path labels rather than long network addresses, avoiding complex lookups in a routing table [18].
- Multiuser system** A computer system that shares its resources among multiple simultaneous users [5].
- Mutually exclusive** State in which one, and only one, of the possible conditions is true; for example, either A or not A is true, and one of the statements is false. When using Bayes' theorem to perform medical diagnosis, we generally assume that diseases are mutually exclusive, meaning that the patient has exactly one of the diseases under consideration and not more [3].
- Myocardial ischemia** Reversible damage to cardiac muscle caused by decreased blood flow and resulting poor oxygenation. Such ischemia may cause chest pain or other symptoms [19].
- Naïve Bayesian model** The use of **Bayes Theorem** in a way that assumes conditional independence of variables that may in fact be linked statistically [22].
- Name** Designation of an object by a linguistic expression [7].
- Name authority** An entity or mechanism for controlling the identification and formulation of unique identifiers for names. In the Internet, a name authority is required to associate common domain names with their IP addresses [14].
- Named-entity normalization** The natural language processing method, after finding a named entity in a document, for linking (normalizing) that mention to appropriate database identifiers [8].
- Named-entity recognition** In language processing, a subtask of information extraction that seeks to locate and classify atomic elements in text into predefined categories [8].
- Name-server** In networked environments such as the Internet, a computer that converts a host name into an IP address before the message is placed on the network [5].
- National Center for Biotechnology Information (NCBI)** Established in 1988 as a national resource for molecular biology information, the NCBI is a component of the National Library of Medicine that creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information [21].
- National Guidelines Clearinghouse** A public resource, coordinated by the Agency for Health Research and Quality, that collects and distributes evidence-based clinical practice guidelines (see www.guideline.gov) [21].
- National Health Information Infra-structure (NHII)** A comprehensive knowledge-based network of interoperable systems of clinical, public health, and personal health information that is intended to improve decision-making by making health information available when and where it is needed [17].

- National Information Standards Organization (NISO)** A non-profit association accredited by the American National Standards Institute (ANSI), that identifies, develops, maintains, and publishes technical standards to manage information (see www.niso.org) [21].
- National institute for Standards and Technology (NIST)** A non-regulatory federal agency within the U.S. Commerce Department's Technology Administration; its mission is to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life (see www.nist.gov) [21].
- National Quality Forum** A not-for-profit organization that develops and implements national strategies for health care quality measurement and reporting [7].
- Nationwide Health Information Network (NwHIN)** A set of standards, services, and policies that have been shepherded by the Office of the National Coordinator of Health Information Technology to enable secure health information exchange over the Internet [12].
- Natural language processing (NLP)** facilitates tasks by enabling use of automated methods that represent the relevant information in the text with high validity and reliability [8, 22].
- Natural language query** A question expressed in unconstrained text, from which meaning must somehow be extracted or inferred so that a suitable response can be generated [21].
- Natural language** Unfettered spoken or written language. Free text [8].
- Naturalistic** Describes a study in which little if anything is done by the evaluator to alter the setting in which the study is carried out [11].
- Needs assessment** A study carried out to help understand the users, their context and their needs and skills, to inform the design of the information resource [11, 23].
- Negative dictionary** A list of **stop words** used in information retrieval [21].
- Negative predictive value (PV⁻)** The probability that the condition of interest is absent if the result is negative—for example, the probability that specific a disease is absent given a negative test result [3].
- Negligence theory** A concept from tort law that states that providers of goods and services are expected to uphold the standards of the community, thereby facing claims of negligence if individuals are harmed by substandard goods or services [10].
- Nested structures** In natural language processing, a phrase or phrases that are used in place of simple words within other phrases [8].
- Net reclassification improvement (NRI)** in classification methods, a measure of the net fraction of reclassifications made in the correct direction, using one method over another method without the designated improvement [25].
- Network access provider** A company that builds and maintains high speed networks to which customers can connect, generally to access the Internet (see also Internet service provider) [5].
- Network Operations Center (NOC)** A centralized monitoring facility for physically distributed computer and/or telecommunications facilities that allows continuous real-time reporting of the status of the connected components [13].
- Network protocol** The set of rules or conventions that specifies how data are prepared and transmitted over a network and that governs data communication among the nodes of a network [5].
- Network stack** The method within a single machine by which the responsibilities for network communications are divided into different levels, with clear interfaces between the levels, thereby making network software more modular [5].
- Neuroinformatics** An emerging subarea of applied biomedical informatics in which the discipline's methods are applied to the management of neurological data sets and the modeling of neural structures and function [9].
- Next Generation Internet Initiative** A federally funded research program in the late 1990s and early in the current decade that sought to provide technical enhancements to the Internet to support future applications that currently are infeasible or are incapable of scaling for routine use [17].

Next generation sequencing methods

Technologies for performing high throughput sequencing of large quantities of DNA or RNA. Typically, these technologies determine the sequences of many millions of short segments of DNA that need to be reassembled and interpreted using bioinformatics [24].

Noise The component of acquired data that is attributable to factors other than the underlying phenomenon being measured (for example, electromagnetic interference, inaccuracy in sensors, or poor contact between sensor and source) [5].

Nomenclature A system of terms used in a scientific discipline to denote classifications and relationships among objects and processes [2, 4, 7].

Nosocomial hospital-acquired infection An infection acquired by a patient after admission to a hospital for a different reason [19].

NQF See: National Quality Forum [7].

Nuclear magnetic resonance (NMR) spectroscopy A spectral technique used in chemistry to characterize chemical compounds by measuring magnetic characteristics of their atomic nuclei [9, 24].

Nuclear medicine imaging A modality for producing images by measuring the radiation emitted by a radioactive isotope that has been attached to a biologically active compound and injected into the body [9].

Nursing informatics The application of biomedical informatics methods and techniques to problems derived from the field of nursing. Viewed as a subarea of clinical informatics [1].

NwHIN Direct A set of standards and services to enable the simple, direct, and secure transport of health information between pairs of health care providers; it is a component of the Nationwide Health Information Network and it complements the Network's more sophisticated components.

Nyquist frequency The minimum sampling rate necessary to achieve reasonable signal quality. In general, it is twice the frequency of the highest-frequency component of interest in a signal [5].

Object Any part of the perceivable or conceivable world [5].

Object Constraint Language (OCL) A textual language for describing rules that apply to the elements a model created in the Uniform Modeling Language. OLC specifies constraints on allowable values in the model. OCL also supports queries of UML models (and of models constructed in similar languages). OCL is a standard of the Object Modeling Group (OMG), and forms the basis of the GELLO query language that may be used in conjunction with the **Arden Syntax** [22].

Objectivist approaches Class of evaluation approaches that make use of experimental designs and statistical analyses of quantitative data [11].

Object-oriented database A database that is structured around individual objects (concepts) that generally include relationships among those objects and, in some cases, executable code that is relevant to the management and or understanding of that object [24].

Odds-ratio form An algebraic expression for calculating the posttest odds of a disease, or other condition of interest, if the pretest odds and **likelihood ratio** are known (an alternative formulation of **Bayes' theorem**, also called the odds-likelihood form) [3].

Office of the National Coordinator for Health Information Technology (ONC) An agency within the US Department of Health and Human Services that is charged with supporting the adoption of health information technology and promoting nationwide health information exchange to improve health care [1, 12].

Omics A set of areas of study in biology that use the suffix “-ome”, used to connote breadth or completeness of the objects being studied, for example genomics or proteomics [25].

-omics technologies High throughput experimentation that exhaustively queries a certain biochemical aspect of the state of an organism. Such technologies include proteomics (protein), genomics (gene expression), metabolomics (metabolites), etc. [24]

On line analytic processing (OLAP) A system that focuses on querying across multiple patients simultaneously, typically by few users for infrequent, but very complex queries, often research [5].

- On line transaction processing (OLTP)** A system designed for use by thousands of simultaneous users doing repetitive queries [5].
- Ontology** A description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. In biomedicine, such ontologies typically specify the meanings and hierarchical relationships among terms and concepts in a domain [8, 22, 25].
- Open consent model** A legal mechanism by which an individual can disclose their own private health information or genetic information for research use. This mechanism is used by the Personal Genome Project to enable release of entire genomes of identified individuals [24].
- Open source** An approach to software development in which programmers can read, redistribute, and modify the source code for a piece of software, resulting in community development of a shared product [6, 12].
- Open standards development policy** In standards group, a policy that allows anyone to become involved in discussing and defining the standard [7].
- Operating system (OS)** A program that allocates computer hardware resources to user programs and that supervises and controls the execution of all other programs [5].
- Optical Character Recognition (OCR)** The conversion of typed text within scanned documents to computer understandable text [12].
- Optical coherence tomography (OCT)** An optical signal acquisition and processing method. It captures micrometer-resolution, three-dimensional images from within optical scattering media (e.g., biological tissue) [9].
- Optical disk** A round, flat plate of plastic or metal that is used to store information. Data are encoded through the use of a laser that marks the surface of the disc [5].
- Order entry** The use of a computer system for entering treatments, requests for lab tests or radiologic studies, or other interventions that the attending clinician wishes to have performed for the benefit of a patient [14].
- Orienting issues/questions** The initial questions or issues that evaluators seek to answer in a subjectivist study, the answers to which often in turn prompt further questions [11].
- Outcome data** Formal information regarding the results of interventions [10].
- Outcome measurements** Using metrics that assess the end result of an intervention rather than an intervening process. For example, remembering to check a patient's Hemoglobin A1C is a process measure, whereas reducing the complications of diabetes is an outcome measure [27].
- Outcome variable** Similar to "dependent variable," a variable that captures the end result of a health care or educational process; for example, long-term operative complication rate or mastery of a subject area [11].
- Output** The results produced when a process is applied to input. Some forms of output are hardcopy documents, images displayed on video display terminals, and calculated values of variables [5].
- P4 medicine** P4 medicine: a term coined by Dr. Leroy Hood for healthcare that strives to be personalized, predictive, preventive and participatory [25].
- Packets** In networking, a variable-length message containing data plus the network addresses of the sending and receiving nodes, and other control information [5].
- PageRank (PR) algorithm** In indexing for information retrieval on the Internet, an algorithmic scheme for giving more weight to a Web page when a large number of other pages link to it [21].
- Pager** One of the first mobile devices for electronic communication between a base station (typically a telephone, but later a computer) and an individual person. Initially restricted to receiving only numeric data (e.g., a telephone number), pagers later incorporated the ability to transmit a response (referred to as "two way pagers") as well as alpha characters so that a message of limited length could be transmitted from a small keyboard. Pagers have been gradually replaced by cellular phones because of their greater flexibility and broader geographical coverage [14].
- Page** A partitioned component of a computer users' programs and data that can be kept in temporary storage and brought into main memory by the operating system as needed [5].

- Parallel processing** The use of multiple processing units running in parallel to solve a single problem (see multiprocessing) [5].
- Parse tree** The representation of structural relationships that results when using a grammar (usually context free) to analyze a given sentence [8].
- Partial parsing** The analysis of structural relationships that results when using a grammar to analyze a segment of a given sentence [8].
- Partial-match searching** An approach to information retrieval that recognizes the inexact nature of both indexing and retrieval, and attempts to return the user content ranked by how close it comes to the user's query [21].
- Participant calendaring** Participant calendaring refers to the capability of a CRMS to support the tracking of participant compliance with a study schema, usually represented as a calendar of temporal events [26].
- Participant screening and registration** participant screening and registration refers to the capability of a CTMS to support the enrollment phase of a clinical study [26].
- Participants** The people or organizations who provide data for the study. According to the role of the information resource, these may include patients, friends and family, formal and informal carers, the general public, health professionals, system developers, guideline developers, students, health service managers, etc. [11]
- Part-of-speech tags** Assignment of syntactic classes to a given sequence of words, e.g., determiner, adjective, noun and verb [8].
- Parts of speech** The categories to which words in a sentence are assigned in accordance with their syntactic function [8].
- Patent** A specific legal approach for protecting methods used in implementing or instantiating ideas (see **intellectual property**) [10].
- Pathognomonic** Distinctively characteristic, and thus, uniquely identifying a condition or object (100 % specific) [2].
- Patient centered care** Clinical care that is based on personal characteristics of the patient in addition to his or her disease. Such characteristics include cultural traditions, preferences and values, family situations and lifestyles [15].
- Patient centered medical home** A team-based health care delivery model led by a physician, physician's assistant, or nurse practitioner that provides comprehensive, coordinated, and continuous medical care to patients with the goal of obtaining maximized health outcomes [15, 27].
- Patient portal** An online application that allows individuals to view health information and otherwise interact with their physicians and hospitals [6, 17, 18].
- Patient record** The collection of information traditionally kept by a health care provider or organization about an individual's health status and health care; also referred to as the patient's chart, medical record, or health record, and originally called the "unit record." [12]
- Patient safety** The reduction in the risk of unnecessary harm associated with health care to an acceptable minimum; also the name of a movement and specific research area [4].
- Patient triage** The process of allocating patients to different levels or urgency of care depending upon the complaints or symptoms displayed [14].
- Patient-tracking applications** Monitor patient movement in multistep processes [14].
- Pattern check** A procedure applied to entered data to verify that the entered data have a required pattern; e.g., the three digits, hyphen, and four digits of a local telephone number [12]
- Pay for performance** Payments to providers that are based on meeting pre-defined expectations for quality [14].
- Per diem** Payments to providers (typically hospitals) based on a single day of care [14].
- Perimeter definition** Specification of the boundaries of trusted access to an information system, both physically and logically [5].
- Personal clinical electronic communication** Web-based messaging solutions that avoid the limitations of email by keeping all interactions within a secure, online environment [18].
- Personal computers** A small, relatively inexpensive, single-user computer [5].
- Personal health application** Software for computers, tablet computers, or smart phones

that are intended to allow individual patients to monitor their own health or to stimulate their own personal health activities [17].

Personal health record (PHR) A collection of information about an individual's health status and health care that is maintained by the individual (rather than by a health care provider); the data may be entered directly by the patient, captured from a sensing device, or transferred from a laboratory or health care provider. It may include medical information from several independent provider organizations, and may also have health and well-being information [10, 12, 15, 17, 20, 27].

Personal Internetworked Notary and Guardian (PING) An early personally controlled health record, later known as Indivo [17].

Personalized medicine Also often call individualized medicine, refers to a medical model in which decisions are custom-tailored to the patient based on that individual's genomic data, preferences, or other considerations. Such decisions may involve diagnosis, treatment, or assessments of prognosis. Also known as **precision medicine** [2, 22, 25].

Personally controlled health record (PCHR) Similar to a **PHR**, the PCHR differs in the nature of the control offered to the patient, with such features as semantic tags on data elements that can be used to determine the subsets of information that can be shared with specific providers [17].

Petabyte A unit of information equal to 1000 terabytes or 10^{15} bytes [5].

Pharmacodynamics program (PD) The study of how a drug works, its mechanism of action and pathway of achieving its affect, or "what the drug does to the body." [25]

Pharmacogenetics The study of drug-gene relationships that are dominated by a single gene [25].

Pharmacogenomics The study of how genes and genetic variation influence drug response [1,25].

Pharmacokinetic program Pharmacokinetics or PK is the study of how a drug is absorbed, distributed, metabolized and excreted by the body, or "what the body does to the drug." [25]

Pharmacovigilance The pharmacologicalscience relating to the collection, detection, assessment,

monitoring, and prevention of adverse effects with pharmaceutical products [8].

Phase I (clinical trial) Investigators evaluate a novel therapy in a small group of participants in order to assess overall safety. This safety assessment includes dosing levels in the case of non-interventional therapeutic trials, and potential side effects or adverse effects of the therapy. Often, Phase I trials of non-interventional therapies involve the use of normal volunteers who do not have the disease state targeted by the novel therapy [25, 26].

Phase II (clinical trial) Investigators evaluate a novel therapy in a larger group of participants in order to assess the efficacy of the treatment in the targeted disease state. During this phase, assessment of overall safety is continued [26].

Phase III (clinical trial) Investigators evaluate a novel therapy in an even larger group of participants and compare its performance to a reference standard which is usually the current standard of care for the targeted disease state. This phase typically employs anrandomized controlled design, and often a multi-center RCT given the numbers of variation of subjects that must be recruited to adequately test the hypothesis. In general, this is the final study phase to be performed before seeking regulatory approval for the novel therapy and broader use in standard-of-care environments [26].

Phase IV (clinical trial) Investigators study the performance and safety of a novel therapy after it has been approved and marketed. This type of study is performed in order to detect long-term outcomes and effects of the therapy. It is often called "post-market surveillance" and is, in fact, not an RCT at all, but a less formal, observational study [26].

Phase In the context of clinical research, study phases are used to indicate the scientific aim of a given clinical trial. There are 4 phases (**Phase I**, **Phase II**, **Phase III**, and **Phase IV**) [26].

Phenome characterization Identification of the individual traits of an organism that characterize its phenotype [20].

Phenome-wide association scan A study that derives case and controls populations using the EMR to define clinical phenotypes and then

examines the association of those phenotypes with specific genotypes [25].

Phenotype definition The process of determining the set of observable descriptors that characterize an organism's phenotype [8].

Phenotype The observable physical characteristics of an organism, produced by the interaction of genotype with environment [24]

Phenotypic Refers to the physical characteristics or appearance of an organism [1].

Picture Archive and Communication Systems (PACS) An integrated computer system that acquires, stores, retrieves, and displays digital images [9, 18, 20].

Pixel One of the small picture elements that makes up a digital image. The number of pixels per square inch determines the spatial resolution. Pixels can be associated with a single bit to indicate black and white or with multiple bits to indicate color or gray scale [5, 9].

Placebo In the context of clinical research, a placebo is a false intervention (e.g. a mock intervention given to a participant that resembles the intervention experienced by individuals receiving the experimental intervention, except that it has no anticipated impact on the individual's health or other indicated status), usually used in the context of a control group or intervention [26].

Plain old telephone service (POTS) The standard low speed, analog telephone service that is still used by many homes and businesses [18].

Plug-in A software component that is added to web browsers or other programs to allow them a special functionality, such as an ability to deal with certain kinds of media (e.g., video or audio) [5].

Pointing device A manual device, such as a mouse, light pen, or joy stick, that can be used to specify an area of interest on a computer screen [5].

Population management Health care practices that assist with a large group of people, including preventive medicine and immunization, screening for disease, and prioritization of interventions based on community needs [22].

Positive predictive value (PV+) The probability that the condition of interest is true if the result

is positive—for example, the probability that a disease is present given a positive test result [3].

Positron emission tomography A tomographic imaging method that measures the uptake of various metabolic products (generally a combination of a positron-emitting tracer with a chemical such as glucose), e.g., by the functioning brain, heart, or lung [9].

Postcoordination The combination of two or more terms from one or more terminologies to create a phrase used for coding data; for example, “Acute Inflammation” and “Appendix” combined to code a patient with appendicitis. See also, **precoordination** [7].

Posterior probability The updated probability that the condition of interest is present after additional information has been acquired [3].

Postgenomic database A database that combines molecular and genetic information with data of clinical importance or relevance. *Online Mendelian Inheritance in Man* (OMIM) is a frequently cited example of such a database [24].

Post-test probability The updated probability that the disease or other condition under consideration is present after the test result is known (more generally, the **posterior probability**) [3].

Practice management system The software used by physicians for scheduling, registration, billing, and receivables management in their offices. May increasingly be linked to an EHR [14].

Pragmatics The study of how contextual information affects the interpretation of the underlying meaning of the language [8].

Precision Medicine The application of specific diagnostic and therapeutic methods matched to an individual based on highly unique information about the individual, such as their genetic profile or properties of their tumor [20].

Precision The degree of accuracy with which the value of a sampled observation matches the value of the underlying condition, or the exactness with which an operation is performed. In information retrieval, a measure of a system's performance in retrieving relevant information (expressed as the fraction of relevant records among total records retrieved in a search) [5, 8, 21].

- Precoordination** A complex phrase in a terminology that can be constructed from multiple terms but is, itself, assigned a unique identifier within the terminology; for example, “Acute Inflammation of the Appendix”. See also, **postcoordination** [7].
- Predicate** The part of a sentence or clause containing a verb and stating something about the subject [8].
- Predicate logic** In mathematical logic, the generic term for symbolic formal systems like first-order logic, second-order logic, etc. [8]
- Predictive value (PV)** The posttest probability that a condition is present based on the results of a test (see positive predictive value and negative predictive value) [2].
- Preparatory phase** In the preparatory phase of a clinical research study, investigators are involved in the initial design and documentation of a study (developing a protocol document), prior to the identification and enrollment of study participants [26].
- Pretest probability** The probability that the disease or other condition under consideration is present before the test result is known (more generally, the **prior probability**) [3].
- Prevalence** The frequency of the condition under consideration in the population. For example, we calculate the prevalence of disease by dividing the number of diseased individuals by the number of individuals in the population. Prevalence is the prior probability of a specific condition (or diagnosis), before any other information is available [2, 3].
- Primary knowledge-based information** The original source of knowledge, generally in a peer reviewed journal article that reports on a research project’s results [21].
- Prior probability** The probability that the condition of interest is present before additional information has been acquired. In a population, the prior probability also is called the prevalence [3].
- Privacy** A concept that applies to people, rather than documents, in which there is a presumed right to protect that individual from unauthorized divulging of personal data of any kind [5].
- Probabilistic context free grammar** A context free grammar in which the possible ways to expand a given symbol have varying probabilities rather than equal weight [8].
- Probabilistic relationship** Exists when the occurrence of one chance event affects the probability of the occurrence of another chance event [3].
- Probabilistic sensitivity analysis** An approach for understanding how the uncertainty in all (or a large number of) model parameters affects the conclusion of a decision analysis [3].
- Probability** Informally, a means of expressing belief in the likelihood of an event. Probability is more precisely defined mathematically in terms of its essential properties [3].
- Probabilistic causal network** Also known as a Bayesian network, a statistical model built of directed acyclic graph structures (nodes) that are connected through relationships (edges). The strength of each of the relationships is defined through conditional probabilities [25].
- Problem impact study** A study carried out in the field with real users as participants and real tasks to assess the impact of the information resource on the original problem it was designed to resolve [11].
- Problem space** The range of possible solutions to a problem [4].
- Problem-based learning** small groups of students, supported by a facilitator, learned through discussion of individual case scenarios [23].
- Procedural knowledge** Knowledge of how to perform a task (as opposed to factual knowledge about the world) [4].
- Procedure** An action or intervention undertaken during the management of a patient (e.g., starting an IV line, performing surgery). Procedures may also be cognitive [8].
- Process integration** An organizational analysis methodology in which a series of tasks are reviewed in terms of their impact on each other rather than being reviewed separately. In a hospital setting, for example, a process integration view would look at patient registration and scheduling as an integrated workflow rather than as separate task areas. The goal is to achieve greater efficiency and effectiveness by focusing on how tasks can better work together rather than optimizing specific areas [14].

- Product** An object that goes through the processes of design, manufacture, distribution, and sale [10].
- Prognostic scoring system** An approach to prediction of patient outcomes based on formal analysis of current variables, generally through methods that compare the patient in some way with large numbers of similar patients from the past [10].
- Progressive caution** The idea that reason, caution and attention to ethical issues must govern research and expanding applications in the field of biomedical informatics [10].
- Propositions** An expression, generally in language or other symbolic form, that can be believed, doubted, or denied or is either true or false [4].
- Prospective study** An experiment in which researchers, before collecting data for analysis, define study questions and hypotheses, the study population, and data to be collected [2,12].
- Prosthesis** A device that replaces a body part—e.g., artificial hip or heart [1].
- Protected memory** An segment of computer memory that cannot be over-written by the usual means [5].
- Protein Data Bank (PDB)** A centralized repository of experimentally determined three dimensional protein and nucleic acid structures [24].
- Proteomics** The study of the protein products produced by genes in the genome [24].
- Protocol analysis** In cognitive psychology, methods for gathering and interpreting data that are presumed to reveal the mental processes used during problem solving (e.g., analysis of “think-aloud” protocols) [4].
- Protocol authoring tools** A software product used by researchers to construct a description of a study’s rationale, guidelines, endpoints, and the like. Such descriptions may be structured formally so that they can be manipulated by trial management software [26].
- Protocol management** Protocol management refers to the capability of a CRMS to support the preparatory phase of a clinical study [26].
- Protocol** A standardized method or approach [5].
- Provider-profiling system** Software that utilizes available data sources to report on patterns of care by one or several providers [14].
- Pseudo-identifier** A unique identifier substituted for the real identifier to mask the identify but can under certain circumstances allow linking back to the original person identifier if needed [16].
- Public health informatics** An application area of biomedical informatics in which the field’s methods and techniques are applied to problems drawn from the domain of public health [1].
- Public health** The field that deals with monitoring and influencing trends in habits and disease in an effort to protect or enhance the health of a population, from small communities to entire countries [10, 16].
- Public Library of Science (PLOS)** A family of scientific journals that is published under the open-access model [21].
- Publication type** One of several classes of articles or books into which a new publication will fall (e.g., review articles, case reports, original research, textbook, etc.) [21].
- Public-key cryptography** In data encryption, a method whereby two keys are used, one to encrypt the information and a second to decrypt it. Because two keys are involved, only one needs be kept secret [5].
- Public-private keys** A pair of sequences of characters or digits used in data encryption in which one is kept private and the other is made public. A message encrypted with the public key can only be opened by the holder of the private key, and a message signed with the private key can be verified as authentic by anyone with the public key [12].
- PubMed** A software environment for searching the Medline database, developed as part of the suite of search packages, known as **Entrez**, by the NLM’s **National Center for Biotechnology Information (NCBI)** [21].
- PubMed Central (PMC)** An effort by the National Library of Medicine to gather the full-text of scientific articles in a freely accessible database, enhancing the value of Medline by providing the full articles in addition to titles, authors, and abstracts [21].
- QRS wave** In an electrocardiogram (ECG), the portion of the wave form that represents the time it takes for depolarization of the ventricles [5].
- Quality assurance** A means for monitoring and maintaining the goodness of a service, product, or process [22].

- Quality management** A specific effort to let quality of care be the goal that determines changes in processes, staffing, or investments [15].
- Quality measurements** Numeric metrics that assess the quality of health care services. Examples of quality measures include the portion of a physician's patients who are screened for breast cancer and 30-day hospital readmission rates. These measurements have traditionally been derived from administrative claims data or paper charts but there is increasing interest in using clinical data from electronic sources [27].
- Quality-adjusted life year (QALY)** A measure of the value of a health outcome that reflects both longevity and morbidity; it is the expected length of life in years, adjusted to account for diminished quality of life due to physical or mental disability, pain, and so on [3].
- Quasi-experiments** A quasi-experiment is a non-randomized, observational study design in which conclusions are drawn from the evaluation of naturally occurring and non-controlled events or cases [26].
- Query and Reporting Tool** Software that supports both the planned and ad-hoc extraction and aggregation of data sets from multiple data forms or equivalent data capture instruments used within a clinical trials management system [26].
- Query** The ability to extract information from an EHR based on a set of criteria; e.g., one could query for all patients with diabetes who have missed their follow-up appointments [12].
- Query-response cycle** For a database system, the process of submitting a single request for information and receiving the results [13].
- Question answering (QA)** A computer-based process whereby a user submits a natural language question that is then automatically answered by returning a specific response (as opposed to returning documents) [8, 21].
- Question understanding** A form of natural language understanding that supports computer-based question answering [8].
- Radiology Information System (RIS)** Computer-based information system that supports radiology department operations; includes management of the film library, scheduling of patient examinations, reporting of results, and billing [20].
- Radiology** The medical field that deals with the definition of health conditions through the use of visual images that reflect information from within the human body [20].
- Random-access memory (RAM)** The portion of a computer's working memory that can be both read and written into. It is used to store the results of intermediate computation, and the programs and data that are currently in use (also called variable memory or core memory) [5].
- Randomized clinical trial (RCT)** A prospective experiment in which subjects are randomly assigned to study subgroups to compare the effects of alternate treatments [2].
- Randomly** Without bias [2].
- Range check** A procedure applied to entered data that detects or prevents entry of values that are out of range; e.g., a serum potassium level of 50.0 mmol/L—the normal range for healthy individuals is 3.5–5.0 mol/L [12].
- Read-only memory (ROM)** The portion of a computer's working memory that can be read, but not written into [5].
- Really simple syndication (RSS)** A form of XML that publishes a list of headlines, article titles or events encoded in a way that can be easily read by another program called a news aggregator or news reader [8].
- Real-time acquisition** The continuous measurement and recording of electronic signals through a direct connection with the signal source [5].
- Recall** In information retrieval, the ability of a system to retrieve relevant information (expressed as the ratio of relevant records retrieved to all relevant records in the database) [8,21].
- Receiver** In data interchange, the program or system that receives a transmitted message [7].
- Receiver operating characteristic (ROC)** A graphical plot that depicts the performance of a binary classifier system as its discrimination threshold is varied [3, 25].
- Records** In a data file, a group of data fields that collectively represent information about a single entity [5].
- Reductionist approaches** An attempt to explain phenomena by reducing them to common, and often simple, first principles [24].

- Reductionist biomedical model** A model of medical care that emphasizes pathophysiology and biological principles. The model assumes that diseases can be understood purely in terms of the component biological processes that are altered as a consequence of illness [22].
- Reference Information Model (RIM)** The data model for HL7 Version 3.0. The RIM describes the kinds of information that may be transmitted within health-care organizations, and includes *acts* that may take place (procedures, observations, interventions, and so on), relationships among acts, the manner in which health-care personnel, patients, and other entities may participate in such acts, and the roles that can be assumed by the participants (patient, provider, specimen, and so on) [7, 22].
- Reference resolution** In NLP, recognizing that two mentions in two different textual locations refer to the same entity [8].
- Reference standard** See gold standard test [8].
- Referential expression** A sequence of one or more words that refers to a particular person, object or event, e.g., “she,” “Dr. Jones,” “or “that procedure.”[8]
- Referral bias** In evaluation studies, a bias that is introduced when the patients entering a study are in some way atypical of the total population, generally because they have been referred to the study based on criteria that reflect some kind of bias by the referring physicians [3].
- Region of interest (ROI)** A selected subset of pixels within an image identified for a particular purpose [9].
- Regional Extension Centers (RECs)** In the context of health information technology, the 60+ state and local organizations (initially funded by ONC) to help primary care providers in their designated area adopt and use EHRs through outreach, education, and technical assistance [13].
- Regional Health Information Organization (RHIO)** A community-wide, multi-stakeholder organization that utilizes information technology to make more complete patient information and decision support available to authorized users when and where needed [13].
- Regional network** A network that provides regional access from local organizations and individuals to the major backbone networks that interconnect regions [5].
- Registers** In a computer, a group of electronic switches used to store and manipulate numbers or text [5].
- Regular expression** A mathematical model of a set of strings, defined using characters of an alphabet and the operators concatenation, union and closure (zero or more occurrences of an expression) [8].
- Regulated Clinical Research Information Management (RCRIM)** An HL7 workgroup that is developing standards to improve information management for preclinical and clinical research [26].
- Relations among named entities** The characterization of two entities in NLP with respect to the semantic nature of the relationship between them [8].
- Relative recall** An approach to measuring recall when it is unrealistic to enumerate all the relevant documents in a database. Thus the denominator in the calculation of recall is redefined to represent the number of relevant documents identified by multiple searches on the query topic [21].
- Relevance judgment** In the context of information retrieval, a judgment of which documents should be retrieved by which topics in a test collection [21].
- Relevance ranking** The degree to which the results are relevant to the information need specified in a query [21].
- Reminder message** A computer-generated warning that is generated when a record meets prespecified criteria, often referring to an action that is expected but is frequently forgotten; e.g., a message that a patient is due for an immunization [12].
- Remote access** Access to a system or to information therein, typically by telephone or communications network, by a user who is physically removed from the system [5].
- Remote Intensive Care** Use of networked communications methods to monitor patients in an intensive care unit from a distance far removed from the patients themselves. See **remote monitoring** [18].

- Remote interpretation** Evaluating tests (especially imaging studies) by having them delivered digitally to a location that may be far removed from the patient [18].
- Remote monitoring** The use of electronic devices to monitor the condition of a patient from a distant location. Typically used to refer to the ability to record and review patient data (such as vital signs) by a physician located in his/her office or a hospital while the patient remains at home. See also **remote intensive care** [14, 18].
- Remote-presence health care** The use of video teleconferencing, image transmission, and other technologies that allow clinicians to evaluate and treat patients in other than face-to-face situations [10].
- Report generation** A mechanism by which users specify their data requests on the input screen of a program that then produces the actual query, using information stored in a database schema, often at predetermined intervals [5].
- Representation** A level of medical data encoding, the process by which as much detail as possible is coded [7].
- Representational effect** The phenomenon by which different representations of a common abstract structure can have a significant effect on reasoning and decision making [4].
- Representational state** A particular configuration of an information-bearing structure, such as a monitor display, a verbal utterance, or a printed label, that plays some functional role in a process within the system [4].
- Representativeness** A heuristic by which a person judges the chance that a condition is true based on the degree of similarity between the current situation and the stereotypical situation in which the condition is true. For example, a physician might estimate the probability that a patient has a particular disease based on the degree to which the patient's symptoms matches the classic disease profile [3].
- Request for Proposals** A formal notification of a funding opportunity, requiring application through submission of a grant proposal [20].
- Research protocol** In clinical research, a prescribed plan for managing subjects that describes what actions to take under specific conditions [2].
- Resource Description Framework (RDF)** An emerging standard for cataloging metadata about information resources (such as Web pages) using the Extensible Markup Language (XML) [21].
- Results reporting** A software system or subsystem used to allow clinicians to access the results of laboratory, radiology, and other tests for a patient [14].
- Retrieval** A process by which queries are compared against an index to create results for the user who specified the query [21].
- Retrospective chart review** The use of past data from clinical charts (classically paper records) of selected patients in order to perform research regarding a clinical question. See also **retrospective study** [2].
- Retrospective study** A research study performed by analyzing data that were previously gathered for another purpose, such as patient care. See also **retrospective chart review** [2, 12].
- Return on investment** A metric for the benefits of an investment, equal to the net benefits of an investment divided by its cost [22].
- Review of systems** The component of a typical history and physical examination in which the physician asks general questions about each of the body's major organ systems to discover problems that may not have been suggested by the patient's chief complaint [2].
- RFP** See: Request for Proposals [20].
- Ribonucleic acid (RNA)** Ribonucleic acid, a nucleic acid present in all living cells. Its principal role is to act as a messenger carrying instructions from DNA in the production of proteins [24].
- Rich text format (RTF)** A format developed to allow the transfer of graphics and formatted text between different applications and operating systems [8].
- RIM** See Reference Information Model [7].
- Risk attitude** A person's willingness to take risks [3].
- Risk-neutral** Having the characteristic of being indifferent between the expected value of a gamble and the gamble itself [3].
- Role-limited access** The mechanism by which an individual's access to information in a database, such as a medical record, is limited depending upon that user's job characteristics and their need to have access to the information [5].

- Router/switch** In networking, a device that sits on the network, receives messages, and forwards them accordingly to their intended destination [5].
- RS-232** A commonly used standard for serial data communication that defines the number and type of the wire connections, the voltage, and the characteristics of the signal, and thus allows data communication among electronic devices produced by different manufacturers [19].
- RSS feed** A bibliographic message stream that provides content from Internet sources [21].
- Rule-based system** A kind of **knowledge-based system** that performs inference using production rules [22].
- Sampling rate** The rate at which the continuously varying values of an analog signal are measured and recorded [5].
- Schema** In a database-management system, a machine-readable definition of the contents and organization of a database [5].
- Schemata** higher-level kinds of knowledge structures [4].
- Script** In software systems, a keystroke-by-keystroke record of the actions performed for later reuse [5].
- SDO** See: Standards development organizations [7].
- Search engine** A computer system that returns content from a search statement entered by a user [21].
- Search** See Information retrieval [21].
- Secondary knowledge-based information** Writing that reviews, condenses, and/or synthesizes the primary literature (see **primary knowledge-based information**) [21].
- Secret-key cryptography** In data encryption, a method whereby the same key is used to encrypt and to decrypt information. Thus, the key must be kept secret, known to only the sender and intended receiver of information [5].
- Secure Sockets Layer (SSL)** A protocol for transmitting private documents via the Internet. It has been replaced by Transport Layer Security. By convention, URLs that require an SSL connection start with https: instead of http: [5, 12]
- Security** The process of protecting information from destruction or misuse, including both physical and computer-based mechanisms [5].
- Segmentation** In image processing, the extraction of selected regions of interest from an image using automated or manual techniques [20].
- Selectivity** In data collection and recording, the process that accounts for individual styles, reflecting an ongoing decision-making process, and often reflecting marked distinctions among clinicians [2].
- Semantic analysis** The study of how symbols or signs are used to designate the meaning of words and the study of how words combine to form or fail to form meaning [8].
- Semantic class** in NLP, a broad class that is associated with a specific domain and includes many instances [8].
- Semantic grammar** A mathematical model of a set of sentences based on patterns of semantic categories, e.g., patient, doctor, medication, treatment, and diagnosis [8].
- Semantic patterns** The study of the patterns formed by the co-occurrence of individual words in a phrase of the co-occurrence of the associated semantic types of the words [8].
- Semantic relations** A classification of the meaning of a linguistic relationship, e.g., “treated in 1995” signifies time while “treated in ER” signifies location [8].
- Semantic sense** In NLP, the distinction between individual word meaning of terms that may be in the same **semantic class** [8].
- Semantic types** The categorization of words into semantic classes according to meaning. Usually, the classes that are formed are relevant to specific domains [8].
- Semantic Web** A future view which envisions the Internet not only as a source of content but also as a source of intelligently linked, agent-driven, structured collections of machine-readable information [21].
- Semantics** The meaning of individual words and the meaning of phrases or sentences consisting of combinations of words [5].
- Semi structured interview** Where the investigator specifies in advance a set of topics that he would like to address but is flexible about the order in which these topics are addressed, and is open to discussion of topics not on the pre-specified list [11].
- Sender** In data interchange, the program or system that sends a transmitted message [7].

- Sensitivity (of a test)** The probability of a positive result, given that the condition under consideration is present—for example, the probability of a positive test result in a person who has the disease under consideration (also called the **true-positive rate**) [2, 3].
- Sentence boundary** In NLP, distinguishing the end of one sentence and the beginning of the next [8].
- Sentiment analysis** The study of how symbols or signs are used to designate the meaning of words and the study of how words combine to form or fail to form meaning [8].
- Sequence alignment** An arrangement of two or more sequences (usually of DNA or RNA), highlighting their similarity. The sequences are padded with gaps (usually denoted by dashes) so that wherever possible, columns contain identical or similar characters from the sequences involved [24].
- Sequence database** A database that stores the nucleotide or amino acid sequences of genes (or genetic markers) and proteins respectively [24].
- Sequence information** Information from a database that captures the sequence of component elements in a biological structure (e.g., the sequence of amino acids in a protein or of nucleotides in a DNA segment) [24].
- Sequential Bayes** A reasoning method based on a **naïve Bayesian model**, where Bayes' rule is applied sequentially for each new piece of evidence that is provided to the system. With each application of Bayes' rule, the **posterior probability** of each diagnostic possibility is used as the new **prior probability** for that diagnosis the next time Bayes' rule is invoked [22].
- Server** A computer that shares its resources with other computers and supports the activities of many users simultaneously within an enterprise [5].
- Service oriented architectures (SOA)** A software design framework that allows specific processing or information functions (services) to run on an independent computing platform that can be called by simple messages from another computer application. Often considered to be more flexible and efficient than more traditional data base architectures. Best known example is the Internet which is based largely on SOA design principles [6, 14, 22].
- Service** An intangible activity provided to consumers, generally at a price, by a (presumably) qualified individual or system [10].
- Set-based searching** Constraining a search to include only documents in a given class or set (e.g., from a given institution or journal) [21].
- Shallow parsing** See partial parsing [8].
- Shielding** In cabling, refers to an outer layer of insulation covering an innerlayer of conducting material. Shielded cable is used to reduce electronic noise and voltage spikes [5].
- Short-term/working memory** An emergent property of interaction with the environment; refers to the resources needed to maintain information active during cognitive activity [4].
- Signal processing** An area of systems engineering, electrical engineering and applied mathematics that deals with operations on or analysis of signals, or measurements of time-varying or spatially-varying physical quantities [9].
- Simple Mail Transport Protocol (SMTP)** The standard protocol used by networked systems, including the Internet, for packaging and distributing email so that it can be processed by a wide variety of software systems [5, 17].
- Simple Object Access Protocol (SOAP)** A protocol for information exchange through the HTTP/HTTPS or SMTP transport protocol using web services and utilizing Extensible Markup Language (XML) as the format for messages [6].
- Simulation** A system that behaves according to a model of a process or another system; for example, simulation of a patient's response to therapeutic interventions allows a student to learn which techniques are effective without risking human life [23].
- Simultaneous access** Access to shared, computer-stored information by multiple concurrent users [5].
- Simultaneous controls** Use of participants in a comparative study who are not exposed to the information resource. They can be randomly allocated to access to the information resource or in some other way [11].
- Single nucleotide polymorphism (SNP)** A DNA sequence variation, occurring when a single nucleotide in the genome is altered. For example, a SNP might change the nucleotide sequence AAGCCTA to AAGCTTA. A variation must

occur in at least 1 % of the population to be considered a SNP [24].

Single-photon emission computed tomography

A nuclear medicine tomographic imaging technique using gamma rays. It is very similar to conventional nuclear medicine planar imaging using a gamma camera. However, it is able to provide true 3D information. This information is typically presented as cross-sectional slices through the patient, but can be freely reformatted or manipulated as required [9].

Single-user systems Computers designed for use by single individuals, such as personal computers, as opposed to servers or other resources that are designed to be shared by multiple people at the same time [5].

Six sigma A management strategy that seeks to improve the quality of work processes by identifying and removing the causes of defects and minimizing the variability of those processes. Statistically, a six sigma process is one that is free of defects or errors 99.99966 %, which equates to operating a process that fits six standard deviations between the mean value of the process and the specification limit of that process [14].

Slip A type of medical error that occurs when the actor selects the appropriate course of action, but it was executed inappropriately [4]

Slots In a frame-based representation, the elements that are used to define the semantic characteristics of the frame [8].

Smart phones A mobile telephone that typically integrates voice calls with access to the Internet to enable both access to web sites and the ability to download email and applications that then reside on the device [14].

SMS messaging The sending of messages using the text communication service component of phone, web or mobile communication systems—Short Message Service [18]

SNOMED Systematized Nomenclature of Medicine - A set of standardized medical terms that can be processed electronically; useful for enhancing the standardized use of medical terms in clinical systems [14].

SNOMED-CT The result of the merger of an earlier version of SNOMED with the Read Clinical Terms [8].

SNP See Single nucleotide polymorphism [25].

Social networking The use of a dedicated Web site to communicate informally (on the site, by email, or via SMS messages) with other members of the site, typically by posting messages, photographs, etc. [17]

Software development life cycle (SDLC) or software development process A framework imposed over software development in order to better ensure a repeatable, predictable process that controls cost and improves quality of a software product [6].

Software oversight committee A groups within organizations that is constituted to oversee computer programs and to assess their safety and efficacy in the local setting [10].

Software psychology A behavioral approach to understanding and furthering software design, specifically studying human beings' interactions with systems and software. It is the intellectual predecessor to the discipline of Human-Computer interaction [4].

Software Computer programs that direct the hardware how to carry out specific automated processes [5].

Solid state drive (SSD) A data storage device using integrated circuit assemblies as memory to store data persistently. SSDs have no moving mechanical components, which distinguish them from traditional electromechanical magnetic disks such as hard disk drives (HDDs) or floppy disks, which contain spinning disks and movable read/write heads [5].

Spamming The process of sending unsolicited email to large numbers of unwilling recipients, typically to sell a product or make a political statement [5].

Spatial resolution A measure of the ability to distinguish among points that are close to each other (indicated in a digital image by the number of **pixels** per square inch) [5, 9].

Specialist Lexicon One of three UMLS Knowledge Sources, this lexicon is intended to be a general English lexicon that includes many biomedical terms and supports natural language processing [21].

Specificity (of a test) The probability of a negative result, given that the condition under consideration is absent—for example, the probability of a negative test result in a person who does not have a disease under consideration (also called the **true-negative rate**) [2, 3].

- Spectrum bias** Systematic error in the estimate of a study parameter that results when the study population includes only selected subgroups of the clinically relevant population—for example, the systematic error in the estimates of sensitivity and specificity that results when test performance is measured in a study population consisting of only healthy volunteers and patients with advanced disease [3].
- Speech recognition** Translation by computer of voice input, spoken using a natural vocabulary and cadence, into appropriate natural language text, codes, and commands [5, 12].
- Spelling check** A procedure that checks the spelling of individual words in entered data [12].
- Spirometer** An instrument for measuring the air capacity of the lungs [18].
- Standard of care** The community-accepted norm for management of a specified clinical problem [10].
- Standard order sets** Predefined lists of steps that should be taken to deal with certain recurring situations in the care of patients, typically in hospitals; e.g., orders to be followed routinely when a patient is in the post-surgical recovery room [1].
- Standard-gamble** A technique for utility assessment that enables an analyst to determine the utility of an outcome by comparing an individual's preference for a chance event when compared with a situation of certain outcome [3].
- Standards development organizations** An organization charged with developing a standard that is accepted by the community of affected individuals [7].
- Static** In patient simulations, a program that presents a predefined case in detail but which does not vary in its response depending on the actions taken by the learner [23].
- Stemming** The process of converting a word to its root form by removing common suffixes from the end [21].
- Stop words** In full-text indexing, a list of words that are low in semantic content (e.g., “the”, “a”, “an”) and are generally not useful as mechanisms for retrieving documents [21].
- Storage devices** A piece of computer equipment on which information can be stored [5].
- Store-and-forward** A telecommunications technique in which information is sent to an intermediate station where it is kept and sent at a later time to the final destination or to another intermediate station [18].
- Strict product liability** The principle that states that a product must not be harmful [10].
- Structural alignment** The study of methods for organizing and managing diverse sources of information about the physical organization of the body and other physical structures [24].
- Structural informatics** The study of methods for organizing and managing diverse sources of information about the physical organization of the body and other physical structures. Often used synonymously with “imaging informatics.” [1]
- Structure validation** A study carried out to help understand the needs for an information resource, and demonstrate that its proposed structure makes sense to key stakeholders [11].
- Structured data entry** A method of human-computer interaction in which users fill in missing values by making selections from predefined menus. The approach discretizes user input and makes it possible for a computer system to reason directly with the data that are provided [22].
- Structured encounter form** A form for collecting and recording specific information during a patient visit [12].
- Structured interview** An interview with a schedule of questions that are always presented in the same words and in the same order [11].
- Structured Query Language (SQL)** A commonly used syntax for retrieving information from relational databases [5]
- Study arm** in the context of clinical research, a study arm represents a specific modality of an experimental intervention to which a participant is assigned, usually through a process of randomization (e.g., random assigned in a balanced manner to such an arm). Arms are used in clinical study designs where multiple variants of a given experimental intervention are under study, for example, varying the timing or dose of a given medication between arms to determine an optimal therapeutic strategy [26].
- Study population** The population of subjects—usually a subset of the clinically relevant population—in whom experimental outcomes

(for example, the performance of a diagnostic test) are measured [3].

Subheadings In MeSH, qualifiers of subject headings that narrow the focus of a term [21].

Subjectivist approaches Class of approaches to evaluation that rely primarily on qualitative data derived from observation, interview, and analysis of documents and other artifacts. Studies under this rubric focus on description and explanation; they tend to evolve rather than be prescribed in advance [11].

Sublanguage Language of a specialized domain, such as medicine, biology, or law [8].

Summarization A computer system that attempts to automatically summarize a larger body of content [21]

Summary ROC curve A composite **ROC curve** developed by using estimates from many studies [3]

Summative evaluation after the product is in use, is valuable both to justify the completed project and to learn from one's mistakes [23]

Supervised learning technique A method for determining how data values may suggest classifications, where the possible classifications are enumerated in advance, and the performance of a system is enhanced by evaluating how well the system classifies a training set of data. Statistical regression, neural networks, and support vector machines are forms supervised learning [22]

Supervised learning An approach to machine learning in which an algorithm uses a set of inputs and corresponding outputs to try to learn a model that will enable prediction of an output when faced with a previously unseen input [25]

Surveillance The ongoing collection, analysis, interpretation, and dissemination of data on health conditions (e.g., breast cancer) and threats to health (e.g., smoking prevalence). In a computer-based medical record system, systematic review of patients' clinical data to detect and flag conditions that merit attention. [12, 13, 16]

Symbolic-programming language A programming language in which the program can treat itself, or material like itself, as data. Such programs can write programs (not just as charac-

ter strings or texts, but as the actual data structures that the program is made of). The best known and most influential of these languages is LISP [5].

Syndromic surveillance An ongoing process for monitoring of clinical data, generally from public health, hospital, or outpatient resources, whereby the goal is early identification of outbreaks, epidemics, new diseases, or, in recent years, bioterrorist events [8, 10].

Synonyms Multiple ways of expressing the same concept [21].

Syntax The grammatical structure of language describing the relations among words in a sentence [5, 8].

System programs The operating system, compilers, and other software that are included with a computer system and that allow users to operate the hardware [5].

Systematic review A type of journal article that reviews the literature related to a specific clinical question, analyzing the data in accordance with formal methods to assure that data are suitably compared and pooled [21].

Systems biology Research on biological networks or biochemical pathways. Often, systems biology analyses take a comprehensive approach to model biological function by taking the interactions (physical, regulatory, similarity, etc.) of a set of genes as a whole [24].

Tablet Generally refers to a personal computing device that resembles a paper tablet in size and incorporates features such as a touch screen to facilitate data entry [14].

Tactile feedback In virtual or telepresence environments, the process of providing (through technology) a sensation of touching an object that is imaginary or otherwise beyond the user's reach. (see also haptic feedback) [5].

TCP/IP Transmission Control Protocol/Internet Protocol – A set of standard communications protocols used for the Internet and for networks within organizations as well [14, 20].

Teleconsultation The use of telemedicine techniques to support the interaction between two (or more) clinicians where one is providing advice to the other, typically about a specific patient's care [18].

- Telegraphic** In NLP, describes language that does not follow the usual rules of grammar but is compact and efficient. Clinical notes written by hand often demonstrate a “telegraphic style” [8].
- Telehealth** The use of electronic information and telecommunications technologies to support long-distance clinical health care, patient and professional health-related education, public health and health administration. See **telemedicine** [18, 27].
- Telehome care** The use of communications and information technology to deliver health services and to exchange health information to and from the home (or community) when distance separates the participants [18].
- Tele-ICU** See remote intensive care [19].
- Telemedicine** A broad term used to describe the delivery of health care at a distance, increasingly but not exclusively by means of the Internet [1, 10, 18].
- Teleophthalmology** The use of telemedicine methods to deliver ophthalmology services [18].
- Telepresence** A technique of telemedicine in which a viewer can be physically removed from an actual surgery, viewing the abnormality through a video monitor that displays the operative field and allows the observer to participate in the procedure [18, 20].
- Telepsychiatry** The use of telemedicine methods to deliver psychiatric services [18].
- Teleradiology** The provision of remote interpretations, increasing as a mode of delivery of radiology services [1, 18,20].
- Telesurgery** The use of advanced telemedicine methods to allow a doctor to perform surgery on a patient even though he or she is not physically in the operating room [18].
- Temporal resolution** A metric for how well an imaging modality can distinguish points in time that are very close together [9].
- Terabyte** A unit of information equal to one million million (10^{12}) or strictly, 2^{40} bytes [5].
- Term frequency (TF)** In information retrieval, a measurement of how frequently a term occurs in a document [21].
- Term weighting** The assignment of metrics to terms so as to help specify their utility in retrieving documents well matched to a query [21].
- Term** Designation of a defined concept by a linguistic expression in a special language [7]. In information retrieval, a word or phrase which forms part of the basis for a search request [21].
- Terminal** A simple device that has no processing capability of its own but allows a user to access a server [5].
- Terminology** A set of terms representing the system of concepts of a particular subject field [7].
- Terminology authority** An entity or mechanism that determines the acceptable term to use for a specific entity, descriptor, or other concept [14].
- Terminology services** Software methods, typically based on computer-based dictionaries or language systems, that allow other systems to determine the locally acceptable term to use for a given purpose [14].
- Test collection** In the context of information retrieval, a collection of real-world content, a sampling of user queries, and relevance judgments that allow system-based evaluation of search systems [21].
- Test-interpretation bias** Systematic error in the estimates of sensitivity and specificity that results when the index and gold-standard test are not interpreted independently [3].
- Test-referral bias** Systematic error in the estimates of sensitivity and specificity that results when subjects with a positive index test are more likely to receive the gold-standard test [3].
- Tethered personal health record** An EHR portal that is provided to patients by an institution and can typically be used to manage information only from that provider organization [15].
- Text-comprehension** A process in which text can be described at multiple levels of realization from surface codes (e.g., words and syntax) to deeper level of semantics [4].
- Text generation** Methods that create coherent natural language text from structured data or from textual documents in order to satisfy a communication goal [8].
- Text mining** The use of large text collections (e.g., medical histories, consultation reports, articles from the literature, web-based resources) and natural language processing to allow inferences to be drawn, often in the form

- of associations or knowledge that were not previously apparent. See also **data mining** [21].
- Text processing** The analysis of text by computer [8].
- Text readability assessment and simplification** An application of **NLP** in which computational methods are used to assess the clarity of writing for a certain audience or to revise the exposition using simpler terminology and sentence construction [8].
- Text REtrieval Conference (TREC)** Organized by NIST, an annual conference on text retrieval that has provided a testbed for evaluation and a forum for presentation of results. (see trec.nist.gov) [21].
- Text summarization** Takes one or several documents as input and produces a single, coherent text that synthesizes the main points of the input documents [8].
- TF*IDF weighting** A specific approach to term weighting which combines **the inverse document frequency (IDF)** and **term frequency (TF)** [21].
- Thesaurus** A set of subject headings or descriptors, usually with a cross-reference system for use in the organization of a collection of documents for reference and retrieval [21].
- Thick-client** A computer node in a network or client-server architecture that provides rich functionality independent of the central server. See also **thin client** [6].
- Thin client** A program on a local computer system that mostly provides connectivity to a larger resource over a computer network, thereby providing access to computational power that is not provided by the machine, which is local to the user [6, 14].
- Think-aloud protocol** In cognitive science, the generation of a description of what a person is thinking or considering as they solve a problem [4].
- Thread** The smallest sequence of programmed instructions that can be managed independently by an operating systemscheduler [5].
- Three-dimensional structure information** In a biological database, information regarding the three-dimensional relationships among elements in a molecular structure [24].
- Time-sharing networks** An historical term describing some of the earliest computer networks allowing remote access to systems. (*BH: might we want to exclude this term?*) [21]
- Time-trade-off** A common approach to utility assessment, comparing a better state of health lasting a shorter time, with a lesser state of health lasting a longer time. The time-tradeoff technique provides a convenient method for valuing outcomes that accounts for gains (or losses) in both length and quality of life [3].
- Tokenization** The process of breaking an unstructured sequence of characters into larger units called “token,” e.g., words, numbers, dates and punctuation [8].
- Tokens** In language processing, the composite entities constructed from individual characters, typically words, numbers, dates, or punctuation [8].
- Top-down** In search or analysis, the breaking down of a system to gain insight into its compositional sub-systems [8].
- Topology** In networking, the overall connectivity of the nodes in a network [5].
- Touch screen** A display screen that allows users to select items by touching them on the screen [5].
- Track pad** A computer input device for controlling the pointer on a display screen by sliding the finger along a *touch*-sensitive surface: used chiefly in laptop computers. Also called a touchpad [5].
- Transaction set** In data transfer, the full set of information exchanged between a sender and a receiver [7].
- Transcription** The conversion of a recording of dictated notes into electronic text by a typist [12].
- Transition matrix** A table of numbers giving the probability of moving from one state in a **Markov model** into another state or the state that is reached in a finite-state machine depending on the current character of the alphabet [8].
- Transition probability** The probability that a person will transit from one health state to another during a specified time period [3].
- Translational Bioinformatics (TBI)** According to the AMIA: the development of storage, analytic, and interpretive methods to optimize the transformation of increasingly voluminous biomedical data, and genomic data, into proactive, predictive, preventive, and participatory health [25].

- Translational medicine** Translational medicine: the process of transferring scientific discoveries into preventive practice and clinical care [25].
- Transmission control protocol/internet protocol (TCP/IP)** The standard protocols used for data transmission on the Internet and other common local and wide-area networks [5].
- Transport Layer Security (TLS)** A protocol that ensures the privacy of data transmitted over the Internet. It grew out of Secure Sockets Layer [12].
- Treatment threshold probability** The probability of disease at which the expected values of withholding or giving treatment are equal. Above the threshold treatment is recommended; below the threshold, treatment is not recommended and further testing may be warranted [3].
- Trigger event** In monitoring, events that cause a set of transactions to be generated [7].
- True negative** In assessing a situation, an instance that is classified negatively and is subsequently shown to have been correctly classified [8].
- True positive** In assessing a situation, an instances that is classified positively and is subsequently shown to have been correctly classified [8].
- True-negative rate (TNR)** The probability of a negative result, given that the condition under consideration is false—for example, the probability of a negative test result in a patient who does not have the disease under consideration (also called **specificity**) [3].
- True-negative result (TN)** A negative result when the condition under consideration is false—for example, a negative test result in a patient who does not have the disease under consideration [3].
- True-positive rate (TPR)** The probability of a positive result, given that the condition under consideration is true—for example, the probability of a positive test result in a patient who has the disease under consideration (also called **sensitivity**) [3].
- True-positive result (TP)** A positive result when the condition under consideration is true—for example, a positive test result in a patient who has the disease under consideration [3].
- Turn-around-time** The period for completing a process cycle, commonly expressed as an average of previous such periods [20].
- Tutoring** A computer program designed to provide self-directed education to a student or trainee [23].
- Twisted-pair wires** The typical copper wiring used for routine telephone service but adaptable for newer communication technologies [5].
- Type-checking** In computer programming, the act of checking that the types of values, such as integers, decimal numbers, and strings of characters, match throughout their use [5].
- Typology** A way of classifying things to make sense of them, for a certain purpose [11].
- Ubiquitous computing** A form of computing and human-computerinteraction that seeks to embed computing power invisibly in all facets of life [4].
- Ultrasound** A common energy source derived from high-frequency sound waves [9].
- UMLS** See: Unified Medical Language System [21].
- UMLS Knowledge Sources** Components of the **Unified Medical Language System** that support its use and semantic breadth [21].
- UMLS Semantic Network** A knowledge source in the UMLS that provides a consistent categorization of all concepts represented in the Metathesaurus. Each Metathesaurus concept is assigned at least one semantic type from the Semantic Network [21].
- Unicode** Represents characters needed for foreign languages using up to 16 bits [5].
- Unified Medical Language System (UMLS) Project** A terminology system, developed under the direction of the National Library of Medicine, to produce a common structure that ties together the various vocabularies that have been created for biomedical domains [8, 21].
- Unified Modeling Language (UML)** A standardized general-purpose modeling language developed for object-oriented software engineering that provides a set of graphic notation techniques to create visual models that depict the relationships between actors and activities in the program or process being modeled [26].
- Uniform resource identifier (URI)** The combination of a URN and URL, intended to provide persistent access to digital objects [21].
- Uniform resource locator (URL)** The address of an information resource on the World Wide Web [5, 21].
- Uniform resource name (URN)** A name for a Web page, intended to be more persistent than a

- URL, which often changes over time as domains evolve or Web sites are reorganized [21].
- Unique health identifier (UHI)** A government-provided number that is assigned to an individual for purposes of keeping track of their health information [27].
- Universal Serial Bus(USB)** A connection technology for attaching peripheral devices to a computer, providing fast data exchange [19].
- Unobtrusive measures** Measures made using the records accrued as part of the routine use of the information resource, including, for example, user log files [11].
- Unstructured interview** an interview where there are no predetermined questions [11].
- URAC** An organization that accredits the quality of information from various sources, including health-related Web sites [21].
- Usability** The quality of being able to provide good service to one who wishes to use a product [4].
- Usability testing** A class of methods for collecting empirical data of representative users performing representative tasks; considered the gold standard in usability evaluation methods [4, 11].
- User authentication** The process of identifying a user of an information resource, and verifying that the user is allowed to access the services of that resource. A standard user authentication method is to collect and verify a username and password [6].
- User-interface layer** The architectural layer of a software environment that handles the interface with users [14].
- Utility** In decision making, a number that represents the value of a specific outcome to a decision maker (see, for example, quality-adjusted life-years) [3, 22].
- Validity check** A set of procedures applied to data entered into an EHR intended to detect or prevent the entry of erroneous data; e.g., range checks and pattern checks [12].
- Variable** Quantity measured in a study. Variables can be measured at the nominal, ordinal, interval, or ratio levels [11].
- Vector mathematics** In the context of information retrieval, mathematical systems for measuring and comparing vector representations of documents and their contents [21].
- Vector-space model** A method of full-text indexing in which documents can be conceptualized as vectors of terms, with retrieval based on the cosine similarity of the angle between the query and document vectors [21].
- Vertically integrated** Refers to an organizational structure in which a variety of products or services are offered within a single chain of command; contrasted with horizontal integration in which a single type of product is offered in different geographical markets. A hospital which offers a variety of services from obstetrics to geriatrics would be “vertically integrated”. A diagnostic imaging organization with multiple sites would be “horizontally integrated” [14].
- Veterinary informatics** The application of biomedical informatics methods and techniques to problems derived from the field of veterinary medicine. Viewed as a subarea of clinical informatics [1].
- Video-display terminal (VDT)** A device for displaying input signals as characters on a screen, typically a computer monitor [14].
- View schemas** An application-specific description of a view that supports that program’s activities with respect to some general database for which there are multiple views [5].
- View** In a database-management system, a logical submodel of the contents and structure of a database used to support one or a subset of applications [5].
- Virtual address** A technique in memory management such that each address referenced by the CPU goes through an address mapping from the virtual address of the program to a physical address in main memory [5].
- Virtual memory** A scheme by which users can access information stored in auxiliary memory as though it were in main memory. Virtual memory addresses are automatically translated into actual addresses by the hardware [5].
- Virtual Private Network (VPN)** A private communications network, usually used within a company or organization, or by several different companies or organizations, communicating over a public network. VPN message traffic is carried on public networking infrastructure (e.g., the Internet) using standard (often insecure) protocols [5, 20].
- Virtual reality** A collection of interface methods that simulate reality more closely than does the standard display monitor, generally with a

- response to user maneuvers that heighten the sense of being connected to the simulation [4].
- Virus/worm** A software program that is written for malicious purposes to spread from one machine to another and to do some kind of damage. Such programs are generally self-replicating, which has led to the comparison with biological viruses [5].
- Visual-analog scale** A method for valuing health outcomes, wherein a person simply rates the quality of life with a health outcome on a scale from 0 to 100 [3].
- Vocabulary** A dictionary containing the terminology of a subject field [4, 7, 21].
- Volatile** A characteristic of a computer's memory, in that contents are changed when the next program runs and are not retained when power is turned off [5].
- Volume rendering** A method whereby a computer program projects a two-dimensional image directly from a three-dimensional **voxel** array by casting rays from the eye of the observer through the volume array to the image plane [9].
- vonNeuman machine** A computer architecture that comprises a single processing unit, computer memory, and a memory bus [5]
- Voxel** A volume element, or small cubic area of a three-dimensional digital image (see pixel) [9]
- WashingtonDC Principles for Free Access to Science** An organization of non-profit publishers that aims to balance wide access with the need to maintain sustainable revenue models [21].
- Web browser** A computer program used to access and display information resources on the World Wide Web [5].
- Web catalog** Web pages containing mainly links to other Web pages and sites [21].
- Web Services Discovery Language (WSDL)** An XML-based language used to describe the attributes of a web service, such as a SOAP service [6].
- Web-based technologies** Computer capabilities that rely on the architecture principles of the Internet for accessing data from remote servers [14].
- Weblogs/blogs** A type of Web site that provides discussion or information on various topics [21].
- WebMD** An American company that provides web-based health information services [23].
- Whole Slide Digitization** The process of capturing an entire specimen on a slide into a digital image. Compared with capturing images of a single field of view from a microscope, this captures the entire specimen, and can be millions of pixels on a side. This allows subsequent or remote review of the specimen without requiring capture of individual fields [20].
- Wide-area networks (WANs)** A network that connects computers owned by independent institutions and distributed over long distances [5].
- Word senses** The possible meanings of a term [8].
- Word size** The number of bits that define a word in a given computer [5].
- Word** In computer memory, a sequence of bits that can be accessed as a unit [5].
- Word sense disambiguation (WSD)** The process of determining the correct sense of a word in a given context [8].
- Workstation** A powerful desktop computer system designed to support a single user. Workstations provide specialized hardware and software to facilitate the problem-solving and information-processing tasks of professionals in their domains of expertise [5].
- World Intellectual Property Organization (WIPO)** An international organization, headquartered in Geneva and dedicated to promoting the use and protection of intellectual property [21].
- World Wide Web (WWW or Web)** An application implemented on the Internet in which multimedia information resources are made accessible by any of a number of protocols, the most common of which is the **HyperText Transfer Protocol (HTTP)** [1, 5, 21].
- Worm** A self-replicating computer program, similar to a computer virus; a worm is self-contained and does not need to be part of another program to propagate itself [5].
- XML** A metalanguage that allows users to define their own customized markup languages. See **Extensible Markup Language** [14].
- X-ray crystallography** A technique in crystallography in which the pattern produced by the diffraction of x-rays through the closely spaced lattice of atoms in a crystal is recorded and then analyzed to reveal the nature of that lattice, generally leading to an understanding of the material and molecular structure of a substance [24].