Principia Medicinae Digitalis Sotoniensis

Essays on the Evolution of the UHS Clinical Data Estate 1980 -2024

Section 2 Essay 1

The Principles of Clinical Data Visualisation and Pattern Recognition in Information

Author: David Rew

Date of posting: 27.10.2024

Publication Plan

The essays which comprise this series will be made available in the first instance on my professional website, https://www.wessexsurgical.co.uk as downloadable PDF documents for review, comment and as a basis for further contributions. They will be amended, updated and supplementary as necessary and as any new material becomes available. All with knowledge and participation in the University Hospital Southampton (UHS) digital programme are welcome to contribute, by writing to me at dr1@soton.ac.uk

Once the project is as complete as is achievable with the available contributions, final copies of each of the essays will be submitted to the University of Southampton ePrint server for formal publication.

Copyright

Section 2

Essay 2:1: The Principles of Clinical Data Visualisation

Essay 2:2: UHS Lifelines Version 1 2009-2011

Essay 2:3: UHS Lifelines Version3 2014-2016

Essay 2:4: UHS Lifelines Version 3 2017-2018

Essay 2:5: UHS Lifelines Version 3 2019-2020

Essay 2:6: UHS Lifelines Version 4 2020-2024

Essay 2:7: Reflections and the future of UHS Lifelines

This is the tenth essay in this series, in which I have sought to record the history of the University Hospital Southampton (UHS) Clinical Data Estate (CDE) from the earliest beginnings in the late 1980 to the present time (2024).

This is also the first essay in the second section of the collection, which describes the history and development of the UHS Lifelines system from first principles to a prime function in the UHS CDE and Electronic Patient Record (EPR).

Essay 2:1 describes the principles of dynamic data visualisation and of pattern display and recognition in electronic data. The principles are the foundations of the UHS Lifelines system

Essay 2:2 describes the development of the first version of UHS Lifelines in Microsoft ASP code to a fully working system through 2009 to 2010.
The second version of Lifelines was intimately linked to the history of the Southampton Breast Cancer Data System (SBCDS). The development of this version is covered in Section 3 of this essay collection, which describes in detail the history of SBCDS through the period 2011 to 2015. Version 2 of Lifelines was also developed and written in ASP code..

Essay 2.3 describes the development of the third version of Lifelines, from 2014 to 2016. This was rewritten in Microsoft DotNet code. The system was launched into the UHS CHARTS Electronic Patient Record (EPR) in 2016.

Essay 2.4 describes the evolution of UHS Lifelines through the period 2017 to 2018.

Essay 2.5 addresses the period 2019-2020 and the culmination of the Lifelines Version 3 development programme.

Essay 2.6 describes the evolution of Version 3 into Version 4 of UHS Lifelines, and its full integration and advanced functionality in the UHS Clinical Data Estate as of 2024.

Essay 2.7 summarises the project and explores the future of the UHS Lifelines concept.

For clarity, I have separated the accounts of the history and development of UHS Lifelines, SBCDS and the Enhanced Somerset Cancer Register (SCR+) into separate sections in essays. In practice, the evolution of the three systems were entirely intertwined.

In later Sections, I describe the application of the data visualisation principles which were developed in the serial versions of UHS Lifelines to the creation and implementation of two further exemplar data systems; SBCDS and the Enhanced Somerset Cancer Registry, SCR+.

These systems as yet (2024) unique to UHS, but we hope that they will act as templates more widely for future EPR systems across the NHS and beyond, are more in tune with the practical daily needs of health professionals in primary, secondary and specialist care.

Essay Contents

Introduction

Information is collected and manipulated to aid human decision making. Data Visualisation is the art and science of presenting alpha-numerical and geographical information to the human eye and brain in an optimal format to convey the message and the intelligence in a data set. This permits the most effective and efficient interpretation of that information to the decision maker.

The Representation of Patterns in Information

Information visualisation techniques exploit the extraordinary visual powers of the human eye and the computing power of the human brain to capture, recognise and interpret images and patterns on a continuous basis when we are awake. These powers of pattern interpretation are refined continuously through life and experience. Of course, these powers can be fooled, but by and large they serve us very well in daily life.

The fundamental principles of information visualisation long predate the computer age. They are captured in every image based representation of the world, from cave paintings to digital photography, and in the graphical representation of data of any type, including commonplace formats such as histograms, bar, pie and TPR charts, architectural and engineering drawings, maps, road signs, and astronomical charts.

The purpose of information visualisation techniques may be summarised as follows:

- to organise and present a data set with the greatest simplicity, clarity, accuracy, visual impact and intuitiveness, so as to convey the content to the observer/s as efficiently and effectively as possible, so as to help him/her/them to understand the message in the data or to make the most effective decisions from the data.

- To do so in ways which recognise the way in which the eye and brain ingest and process information;

Data visualisation is thus emphatically not about presenting information in the prettiest or most intricate patterns to impress the observer/s appreciation the quality of the presenter's artistic creativity or ability to cram information onto the page or screen.

It is absolutely focussed on the interaction between the information, the creative interface and the communication with the higher functions of the visual brain to aid automaticity, speed and accuracy in the interpretation and hence the quality of the resulting decision making. Of course, computers substantially extend the tools of data visualisation through the powers, speed and capacity for data manipulation, data interaction, graphical representation and user experimentation.

The science of data visualisation nevertheless incorporates the art of graphical design, because it is focussed upon the operational aspects of data presentation to the observer.

At its best, excellence in data visualisation science renders complex, diverse and heterogenous data sets into elegant, intuitive and easily navigable formats, to aid effective decision making at least risk of omission of key themes and actionable evidence.

A Personal Journey Through Data Visualisation

My interest in the science of data visualisation started in the physiology laboratories at the University of Cambridge as an undergraduate medical student and supervisee of Dr Colin Blakemore in 1976. Colin was later to be knighted, and his glittering career included the Chairmanship of the Medical Research Council. He died in June 2022, immediately prior to which I had briefly corresponded with him about our current work in Southampton on data visualisation principles and his early influence on it.

As a post-doctoral researcher In Cambridge, Colin was seeking a better understanding of the working of the mammalian visual system and of the links between the eyes and the occipital or visual cortex in the hindbrain. His supervision periods were therefore heavily weighted to his own interests in developmental neurophysiology and in particular those of vision.

The information processing power of the visual brain

I learned from Colin that in simplistic terms of visual information processing and for the purposes of this essay, there are two critical components to the human brain; the hindbrain for visual processing, and the forebrain for “higher” processing functions (language and computation) (Figure 1).

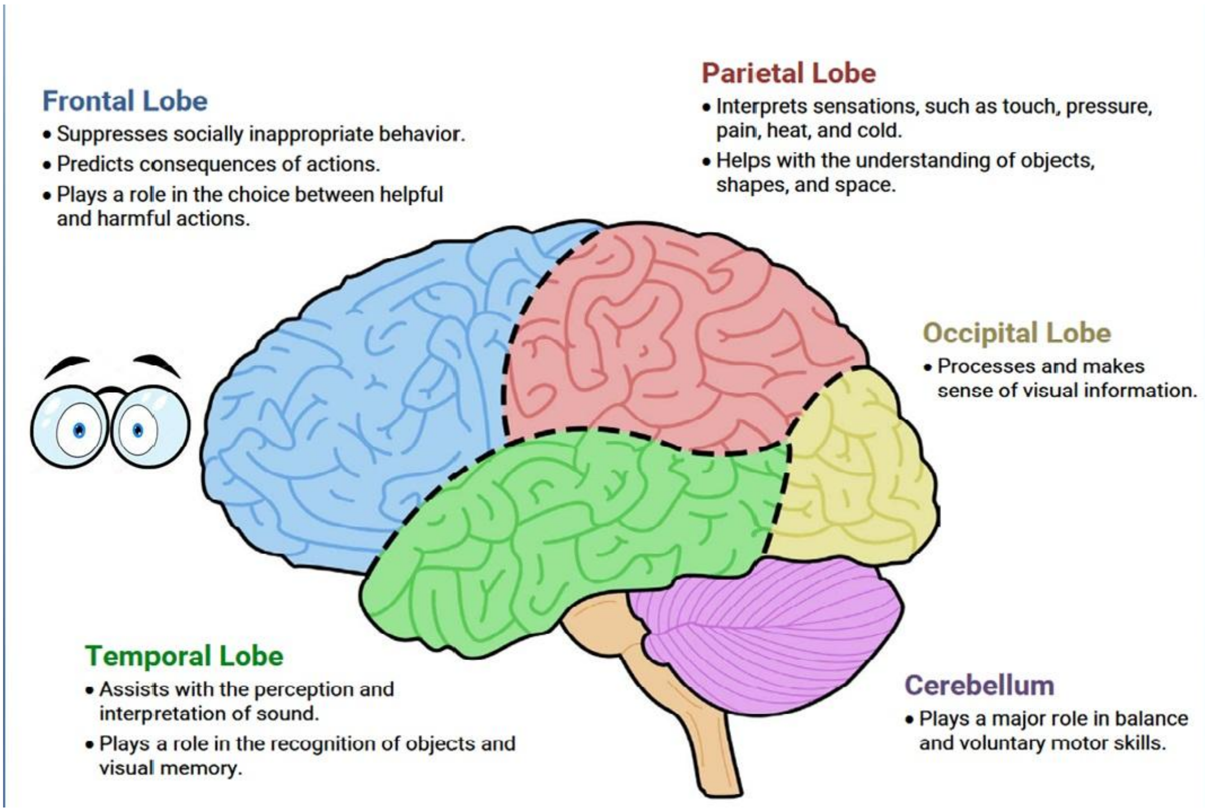

Figure 1. Adapted from The Human Brain Diagram, Therapist Aid 2015
https://www.therapistaid.com/therapy-worksheet/the-human-brain-diagram/

The visual cortex is highly evolved for us to process and interpret the hugely complex information sets which are presented to us through every second of our lives, in ways which are immediately intelligible and highly intuitive. In effect, this key system within the mammalian and human brain continuously processes and interprets huge volumes of “digital” information from our eyes at lightning fast speeds. Through these functions, we recognise shape, colour, movement, depth, patterns and objects as we observe, respond to and move around our world in seeking to stay alive. This visual ability helps us to thrive in complex environments. It is inherent from birth. It is nearly instantaneous and virtually effortless.

In parallel, the higher centres of the human brain are trained to recognise, name, interpret those images, in terms of landscape, language, facial recognition and so on. Our cortical functions in the frontal brain of reading and computation are conducted at a slower speed, and the computing effort is prone both to error and fatiguability.

My interest in data visualisation in clinical medicine

My graduation from Cambridge and undergraduate medical training at Kings College Hospital in 1981 led to an intense period of 13 years of surgical training posts around London and Wessex. I also spent two years of research in Professor Irving Taylor's University Surgical Unit in Southampton into the measurement of tumour cell proliferation characteristics of human solid tumours in living systems through the stimulation of cell and tissue samples using laser light at precisely controlled wavelengths. This work led to a Master of Surgery Thesis of the University of Cambridge.

In 1994 I moved to Leicester for a five year stint as a Senior Lecturer in Sir Peter Bell's University Surgical Unit, with a personal research interest in the further applications of laser cytometry to the study of human tumour biology, At Glenfield Hospital in Leicester, I led a programme on the development of an laser flow and scanning cytometry laboratory for the further use of light to extract useful information at the cellular and subcellular level for the study of tumour function and therapeutics. This secured doctoral Theses for Louise Reeve and Davinder Kaur under the co-supervision with Professor Alison Goodall.

I returned to Southampton in 1999 as an NHS Consultant Surgeon and as an Honorary Senior Lecturer in the Faculty of Medicine at the University of Southampton. The surgeon's brain is necessarily very visual, and my own academic interest in the use of light in the science and art of data visualisation was further stimulated by the historical work of French Engineer Charles Minard. This has been set out in full in a book by Sandra Rendgen on his complete works, published by Princeton Architectural Press.

In the most famous exemplar of his work in 1869, using an evolved graphical technique, Minard described Napoleon's disastrous campaign in Russia through 1812-1813 in a brutal realisation of the route of his army. Seven dimensions of data were drawn from a range of historical resources, as seen in Figure 2. Minard's map has been described as one of the best statistical graphics ever drawn. Napoleon crossed the River Vistula in 1812 with 400,000 men. Ravaged in battles and by the Russian winter, Napoleon returned the following year with only 10,000 survivors (Figure 3). A copy of the map still takes pride of place in my study.

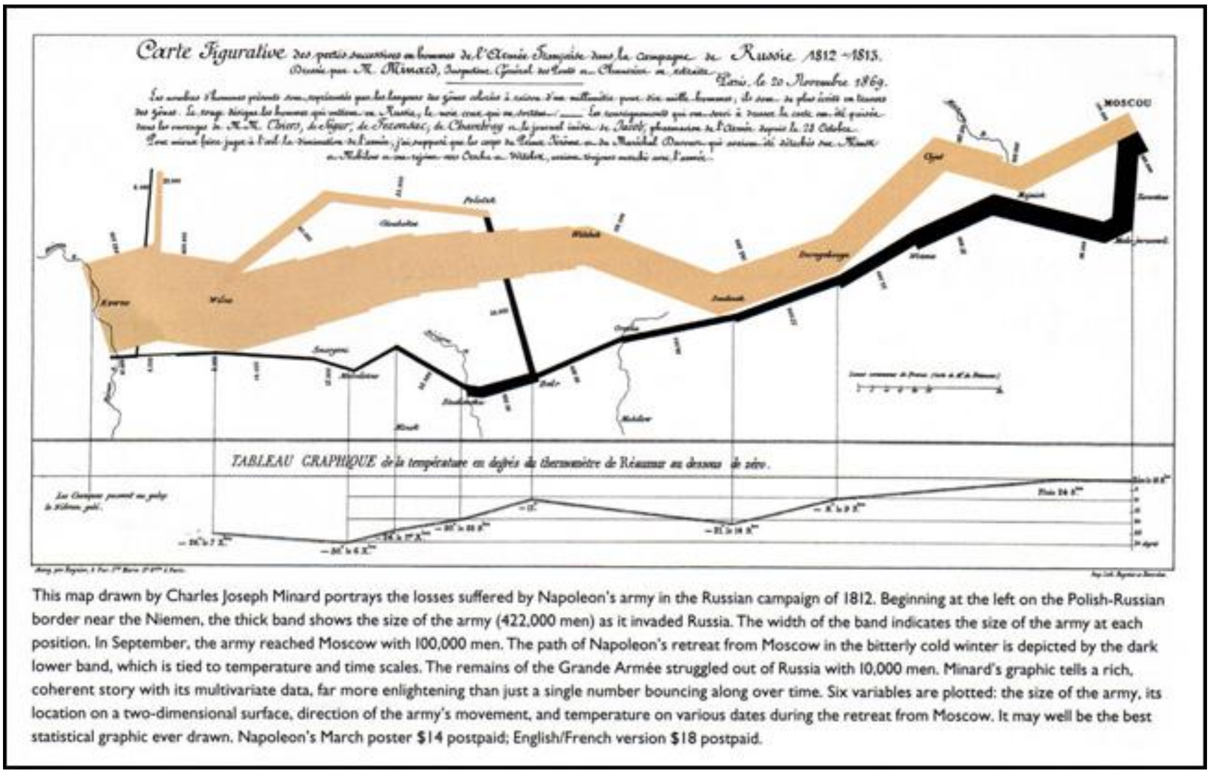

Figure 2. Charles Minard’s graphic of Napoleon’s Moscow Campaign, 1869

Figure 3. Napoleon’s retreat from Moscow: Oil on Canvas: Adolph Norten 1851

Edward Tufte's textbook on "The Visual Display of Quantitative Information" is a major and authoritative resource on the subject. https://www.edwardtufte.com/tufte/books_vdqi.

The visual representation of complex clinical data

The differences in functionality in the visual and frontal cerebral cortices are captured in the aphorism that "a picture is worth a thousand words" (Figure 3). Visually rich and pre-assembled information in a pattern or an image will therefore be processed and interpreted by the brain through the visual cortex far more quickly and efficiently than will text and tables through the frontal cortex.

Human conflict has been a particular stimulus to innovation,. Florence Nightingale was a pioneer in graphical clinical data visualisation. She famously recorded the deaths of the soldiery from battle wounds and preventable infections and other diseases in powerful data visualisations through her pie charts from the Crimean War (Figure 4).

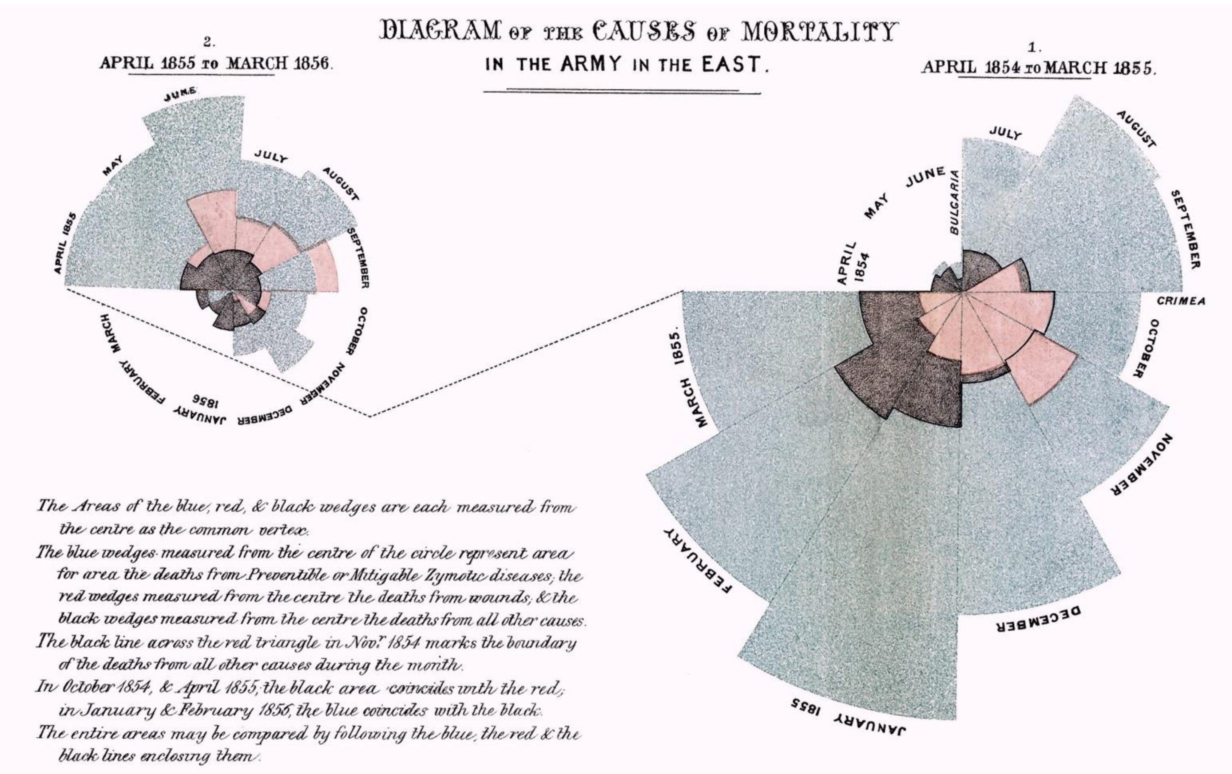

Figure 4. Diagram of the causes of mortality in the Army of the East, 1855-56 by Florence Nightingale: https://commons.wikimedia.org/wiki/File:Nightingale-mortality.jpg

The use of graphical data visualisation was also put to powerful use during the 19th century by public health officers, as the following images illustrate.

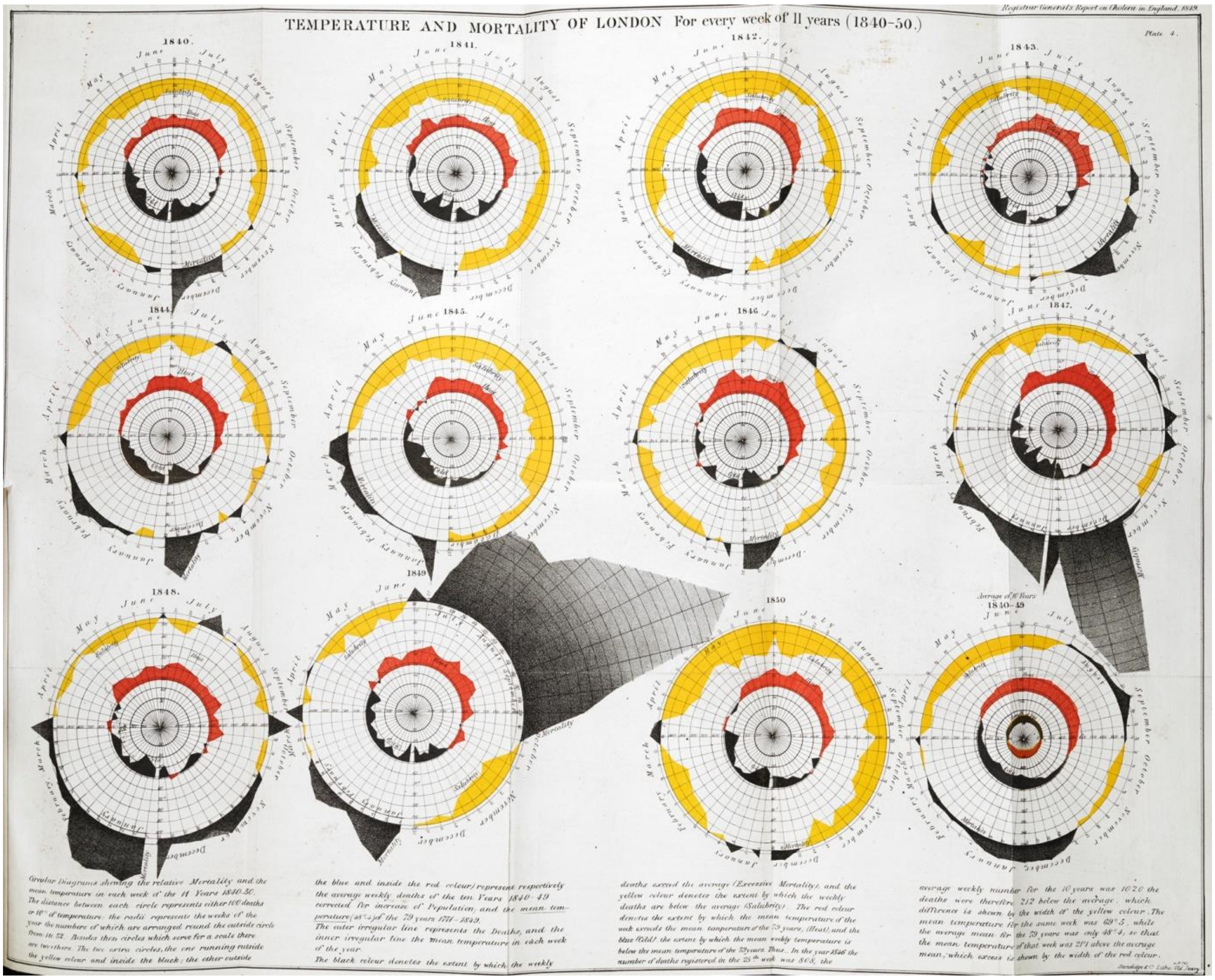

Figure 5: Charts showing the Temperature and mortality of London for every week of 11 years (1840-1850) by William Farr CB., drawn and published in his Report on the Mortality of Cholera in England 1848-49. [Registrar general's report on cholera in England, 1849]. London, 1852. Source: B.S.34/12 plate 4.

The shaded black colour denotes the extent by which the weekly deaths exceed the average; the yellow colour denotes the extent by which the weekly deaths are below the average; the red colour denotes the extent by which the mean temperature of the week exceeds the mean temperature of the years 1771-1849; the solid black denotes the extent by which the mean weekly temperature is below the mean temperature of the years 1771-1849.

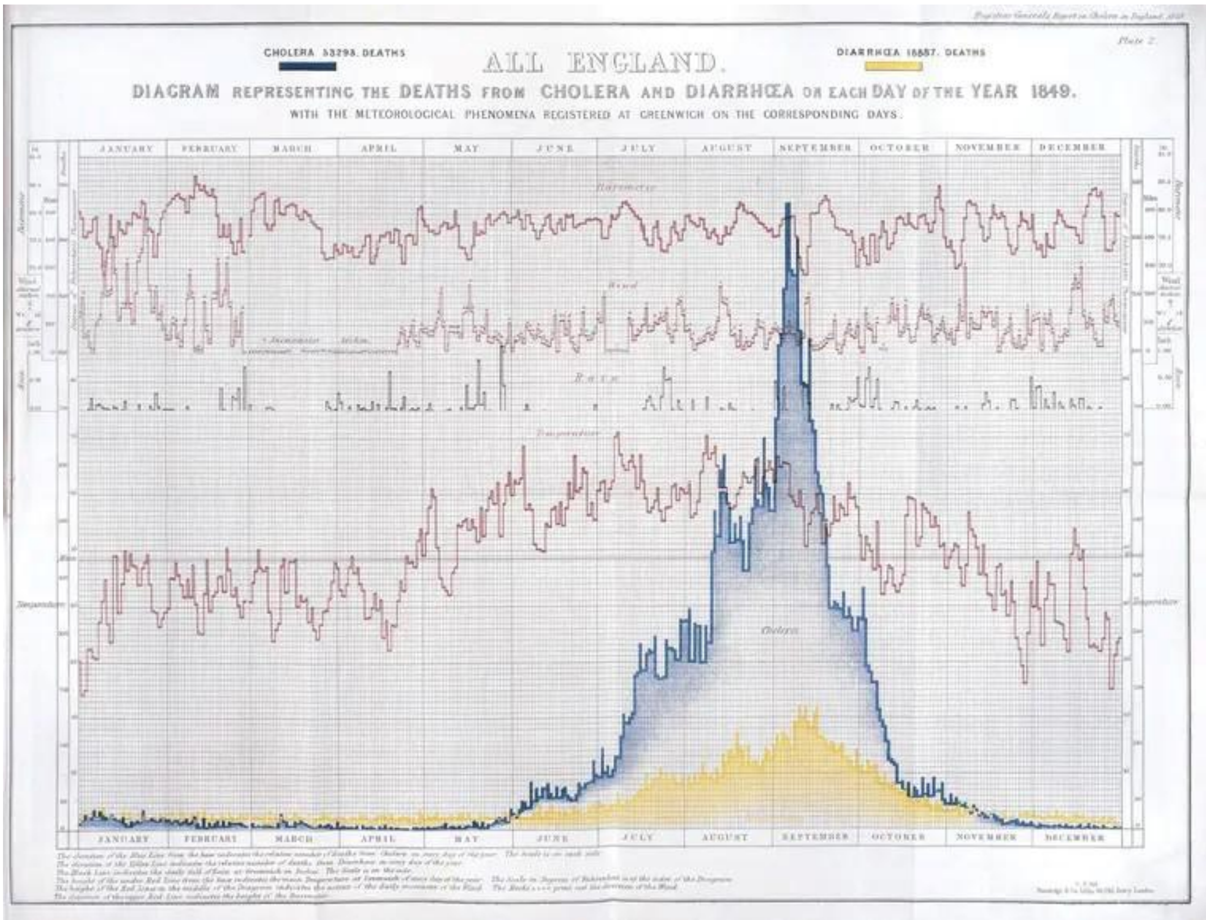

Figure 6: Diagram representing the deaths from Cholera and Diarrhoea on each day of the year 1849: from the Report on the mortality of cholera in England, 1848-49. Compiled by William Farr for the Registrar General. In 1852

James Cheshire, Professor of Geographic Information and Cartography at University College London, reminds us in his Blog on How the Victorians Mapped London's Cholera, of John Snow, who used maps to demonstrate that the clusters of deaths from cholera in London's Soho during London's 1854 outbreak were caused by contaminated water. ..

John Snow was part of an arms race to get the best data communicated by the most compelling maps and charts... William Farr was also a master data visualiser. The Wellcome Collection's image catalogue contains many high resolution examples of maps and charts from around the 1850s, which are free to use under a CC-BY 4.0 license.

Figure 7: Plan Showing the Ascertained Deaths from Cholera (John Snow) in the parishes of St James Westminster and St Anne Soho, mapping their relationship to the contaminated water pump in Broad Street

Credit: https://jcheshire.com/visualisation/mapping-and-visualising-cholera-data/

Data Visualisation using Computers

My interest in the subject of data visualisation was further expanded in the early 2000s, when I was prompted through observations in the deficiencies in clinical cancer data in my role as Editor in Chief of The European Journal of Surgical Oncology, the EJSO. I reflected at length upon the challenges in our understanding of the whole of life progression of various types of cancer and the measurement of the impact of various treatments during various phases of the disease and along the timeline of disease progression.

These thought processes led me from paper based information to research the evolving digital science of dynamic data visualisation through computer systems, and hence of the computer based analysis of clinical timelines and data. I will return to this matter in greater detail in the essays on the development of the Southampton Breast Cancer Data System.

The outputs of the Human Computer Interaction Laboratory at the University of Maryland during the 1990s were foremost in the field of computer based data visualisation. This work was conducted under the direction of Professor Ben Shneiderman, whose work and guiding principles I will describe later in this essay.

The digitisation of information has many advantages. It allows for data storage in huge quantities; for data transmission at the speed of light to any point across the globe; and for data manipulation of data and images in many novel ways.

The invention of the Computer Mouse and the Graphical User Interface (GUI) at the Xerox Paulo Alto Research Centre in the early 1970s and their implementation on the Xero Alto computer workstation hugely simplified and democratised the interaction of the user with complex data on commercial and personal computers.

The GUI it is now universal on hand held devices and smart phones, all be it that touch sensitive screen control is now as commonplace as the mouse. Dynamic iconography has eliminated the need for typed instructions in many applications.

The GUI introduced a series of operating functions with which are central to the ease of modern computer use, including:

- drop down menus
- Multiple windows, frames and tabs on a single screen
- point and click functionality of the mouse, electronic pen or finger
- graphical tools such as symbols and icons to represent documents, folders and events
- Sliders and windows which can be opened, closed, expanded and contracted at will
- highlighting of interactive function by shape, colour, fonts, underlining, flashing, vibration
- Zooming: the ability to select an area of interest and expand it.
- Lensing: the ability to move over a data set, magnifying areas of interest within the lens field

In parallel, the development of GUI enhanced word processing programs such as MS-Word and similar formats, and the introduction of the Portable Document Format (PDF), simplified the technology of document generation, management and transmission.

The Challenges of Information and Activity Overload in Digital Systems

Despite the extraordinary advances in the technical development of hardware and software tools to simplify our interaction with computers, all is not rosy in our digital Garden of Eden.

Computers have transformed our access to information in many ways, but many aspects of their design impose substantial time and energy penalties upon us in the use of their basic functions.

The Anon columnist Bartelby captured this problem in the above image and in an Essay in The Economist Journal in an article which was published on 16th June 2022 under the title

"**Work: The Wasted Years:**

Logging in, deleting emails, mistyping things. It all adds up"

With tongue firmly in cheek, he/she observed that:

"Few things are more depressing than estimates of how much time people spend on a specific activity over the course of their lives. ...

A new study, by academics from the Maryland and Delaware Enterprise University Partnership (MADEUP), applies this approach to the workplace. The researchers identify the number of minutes that people waste on pointless activities each working day. The authors then extrapolate these figures to come up with a "Weighted Total Futility" (WTF) lifetime estimate of time that could have been better spent. The results are literally unbelievable.

Correcting typos takes up an average of 20 minutes in every white-collar worker's day, the equivalent of 180 days, or half a year, over a 45-year career. Some words are mistyped so frequently that on their own they can waste days of the average employee's existence. "Thnaks" is the worst offender in the English-speaking world, followed by "teh", "yuo", "remeber" and "Bets wishes".

The gestation period of a goat at around 145 days is also how long the average worker spends logging into things during his or her working life. Months are wasted trying to remember passwords, entering them wrongly or updating them. Just as much time is spent waiting for something to happen, a great economy-wide period of vacant staring at a screen.

If getting into things wastes lots of time, so does closing them down. Eliminating help windows and tool-tip boxes takes up days over a career. Rejecting repeated requests to schedule updates to your operating system is another chunk of existence that you will never get back. Zapping pop-up ads and trying to pause auto-playing video absorbs time that could have been spent learning to knit or visiting Machu Picchu. A bundle of "tidying up" activities absorbs over four months of the average worker's life. Deleting emails takes up *about six weeks of your life. Clicking on Slack channels to read through messages that are not meant for you, or clearing notifications on your phone screen for articles that you will never look at: tasks like these each eat up several days.*

Various types of formatting tasks constitute another huge time-suck. Think of those attempts to change the margins on Word or Google documents, or the hours spent trying to work out where exactly you need to put the missing bracket in that broken spreadsheet formula. Shakespeare wrote "King Lear" in the time an average office worker spends changing font sizes during their career.

Redoing work that you have failed to save is in a category all of its own, because of the psychological trauma involved. Batteries still run out at crucial moments, internet connections still fail. Failing to save a series of deeply insightful comments in a Google doc them and then closing everything down causes a special kind of despair. So does creating an org chart with hundreds of arrows and text boxes, and realising you missed someone out.

These are only some of the many ways in which time is routinely wasted. Co-ordinating diaries for meetings that will later be cancelled: another month. Waiting for people to repeat themselves because they were on mute by mistake: a fortnight. Spending hours crafting an email and then leaving it in the drafts folder: two days. Desperately opening and shutting various flaps on a recalcitrant printer: a day.

The MADEUP study highlights that technology lies at the heart of this squandered time. Technology can also help. Services that sync up diaries and autocorrect options already do; passwords will doubtless end up being replaced by facial recognition and fingerprint logins. Whether the time thereby saved would be put to more productive use, like reading this column, is a reasonable question. But years of workers' lives are wasted on utterly pointless activities. All improvements warrant heartfelt THNAKS."

These factors are particularly familiar to users of the electronic patient record, as clearly articulated by Dr Athul Gawande in his essay on "Why Doctors Hate Their Computers" in the issue of the New Yorker Magazine of November 12th 2018.

Logical Information Assembly

Information has exploded in volume and in the speed of generation and transmission, so individual decision makers are now faced by the limitations of time and mental energy to collate, analyse and interpret information, and to sort the wheat from the chaff.

Computer users often need to re-assemble multiple documents, events, reports and images to help the user to understand and solve a practical problem, or to make a decision based upon the best available data. The form and patterns of presentation of that data are critical.

Unfortunately, the development of screen methodologies to re-assemble information on computer screens has lagged behind many other digital technologies. Users often need to reassemble document and event collections from many diverse sources, from separate record filing systems, and from lists of documents in multiple menus and screen windows. This need for data assembly in any given user context creates delay and decision risk, because the user may need to cut corners against practical time constraints and deadlines, and against limits to personal mental energy and concentration.

The Newsagents Analogy of Logical Information Assembly

To better understand the problem that Athul Gawande has so clearly articulated , imagine by analogy, that you are seeking to purchase a newspaper from a bookstore. The very helpful sales assistant directs you to all of the filing cabinets and shelves where the pages of the newspaper are individually stored.

He or she also offers the additional helpful information that some of the pages may well have been mislabelled or mis-filed, but you should find them if you search long and hard enough.

You will be able of course to purchase and read the newspaper as soon as you have completed the reassembly task in the store, if the business is still trading when you eventually reach the checkout till, and if you can remember why you wanted to read the newspaper in the first place, and if tomorrow’s edition has not already arrived.

The Reconstruction of Operational Information on Computer Screens

The need to reconstruct sequential events from "traditional" screen list mode display formats, whether in Microsoft, Apple or other standard systems, therefore wastes a huge amount of user time and mental effort to complete a core task. The chances are that the user who is working under time constraints will only have the time to dip into the most immediate accessible and relevant documents, at the risk of missing critical information under the constraints of time and mental effort.

If we are to solve this problem of data reassembly into a coherent and usable format, we must find new ways to use computer systems to simplify the assembly and presentation of the record in real time and in way that is optimised of the human visual brain.

The Representation of Time within the Electronic Patient Record

Everything in the natural world changes continuously with time. Every life, from conception to our death, creates a cometary trail of disjointed information about out health, through phases of good health, disease and injury. All citizens pass through a wide range of health care facilities and through the hands of many health professionals in many circumstances.

The personal health record of every life is therefore a complex admixture of many types of information on multiple storage systems which can stretch out over 100 years. They may be kept or lost in many locations and formats in one or more countries and health systems, and most records of more than 25 years old will still be held on paper. Individuals may generate hundreds or even thousands of separate documents, reports and events about their health over their lifetime.

The life course of an individual may also follow a complex course, with a complex mix of acute conditions and chronic diseases in a variety of body tissues, organs and systems. Although all body systems and organs are all related, it is often convenient and practical to consider them separately for the purposes of medical reasoning, patho-physiological study and record keeping. These systems include the neurological system; the respiratory system; the cardiovascular system; the gastrointestinal system, and/or the musculoskeletal system

We can also categorise health information by the sequence of investigations and tests that the patient undergoes in health and disease; or by age related factors, for example relating to childhood (paediatrics) and late life (elderly care), and even beyond death (the death certificate and the post mortem report).

In this context, it is helpful to consider an ideal framework for the practical visualisation and economical navigation of the complex health information space which describes our individual lives, so as to define a working model of the Ideal EPR.

Imagineering the Visualisation of the Ideal Personal Health Record

An ideal Personal Health Record **system** would collect, collate, organise and display all information about our health from conception (including our foetal medicine scans and mother's obstetric records) to death and beyond (including our post mortem reports and coroner's reports); in a standard system and in a standard format, along with our birth, death and unique identity records.

It would be systematically filed, and immediately accessible to the individual to whom it related and to any permitted health professional, subject to reasonable controls to confidentiality; in any location on a need to know basis in any health system; from one country to another; and in any language for globally mobile citizens and travellers.

It would be under the ownership of the individual citizen himself or herself, who would also be taught and trained in the responsibilities and principles of its use and accurate maintenance throughout life.

It would be displayable in a unifying format on a wide range of different systems, including desktop, laptop and hand held devices.

This ideal has never yet been achieved, but it is now technically possible using current generation digital systems for future generations of citizens. The constraints to full implementation are primarily of a political and educational nature. Nevertheless, it sets an achievable goal for governments and citizens, with the worthy objective of better individual and collective healthcare for all.

The Representation of the Ideal EPR on a Unitary Graphical User Interface

The development of an effective visualisation, navigation and interrogation system for the ideal EPR mandates a series of achievable digital design principles, which are as follows:

- All contributing records, reports documents, images and multimedia content should have standard Metadata. Metadata are the identity tags of electronic records. Dates on hand-written notes, letters and documents as a basic writing discipline. Similarly, metadata should also be applied to all electronic documents and events, to include (as a minimum) a date/time stamp of origination, a description of content according to a standard subject taxonomy or classification, and a place of origin.

- All contributing records, wherever and however held, should be accessible to the relevant central “brain”, or information assembler, of each personal EPR.

- The information should be available in real time and at any point of need, by any authorised individual, and with all appropriate security and access controls

- The content of the record should be under the ultimate ownership in law with the individual citizen whose health it describes, subject to mental capacity and reasonable practical considerations.

- the system should present the information to the owner or user in the most efficient, understandable and actionable way to optimise clinical decision making.

- The system should therefore present the data its correct and sequential timeline and spatial order, and in a way that clearly indicates the relationships between conditions, tests and significant clinical events, that would not otherwise be apparent in a linear filing system. Intelligent design offers a simple and elegant solution to the particular challenge of efficient information retrieval, organisation and presentation of data in the EPR.

The Timeline of the Human Life

The Timeline is a well established graphical format for the visual representation of a human life. The life and career of President Benjamin Franklin is often used as an exemplar.

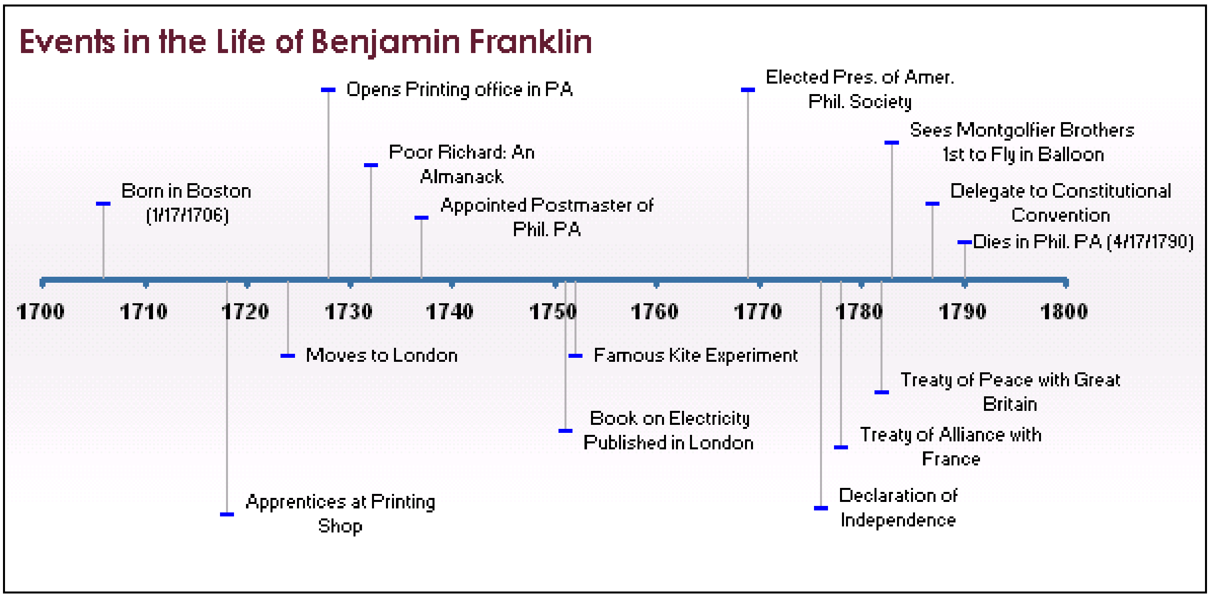

Figure 8: is a simple timeline of key events in the life of President Franklin

Ref https://graphaday.blogspot.com/2009/11/timeline-chart-for-benjamin-franklin.html

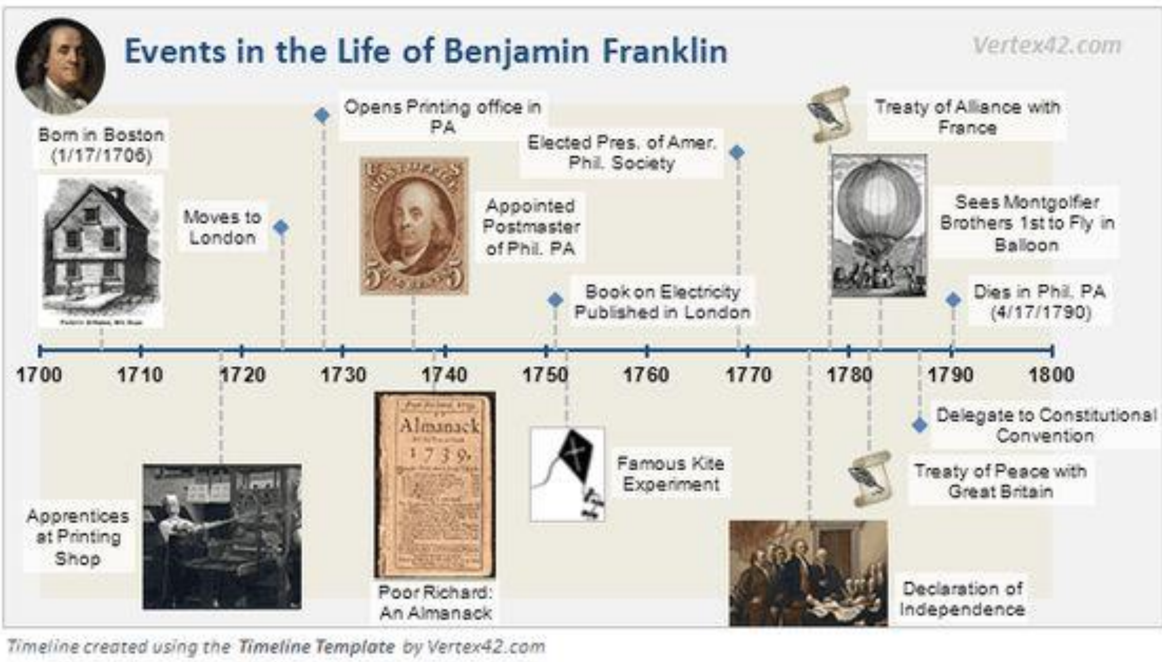

Figure 9: President Franklin’s timeline with added “multimedia”: images and text

Figure 8 illustrates both a simple representation of the timeline of his life, but also the problem of labelling of those key events when we start to populate it with descriptions of the events, whether on the printed page or the computer screen. The written descriptions of the time points are in effect the metadata of those time points, and for ease of reading in a reasonably sized font they take up a lot of screen estate.

Figure 9 shows how the problem of screen clutter is further compounded when we add images or multimedia material.

Fortunately, there are simple programming tricks to overcome this constraint, to which we will return later. We can hide the actionable information (the images in Figure 9) behind an actionable dynamic icon at the relevant time point, so that clicking on the icon opens up the image. We can also place the metadata in a dynamic balloon at the time point, such that hovering over the time point displays the balloon with its relevant metadata.

Experimental Formats for The Representation of Time on Computer Graphical Interfaces:
The commonest and simplest time-structured computer format is the list order. Computer software designers commonly display documents and events in list mode, in date order, which may be ascending or descending by preference. The list mode is useful for the correct sequencing of documents or events, but it gives no simple visual indication of the temporal separation of sequential events, which may be 20 minutes or 20 years between two related events.

Through the 1990s and 2000s, a range of experimental approaches to timeline visualisations in the EPR were tested and published. They included:

The Flat Calendar

The Sequential List

The Multilayered Timeline

The Layered Timeline

Rich graphical formats, including two and three dimensional interactive graphics.

Synchronised and parallel timelines

In 2004, Dina Goren-Bar and colleagues in Israel described KNAVE-II in the Proceedings of the working conference on Advanced visual interfaces, AVI 2004. This was a presentational and interpretational model for time-structured haematology data, and treatments. (Figure 10). The very cluttered nature of the interface is evident, and the model does not appear to have evolved into a practical system.

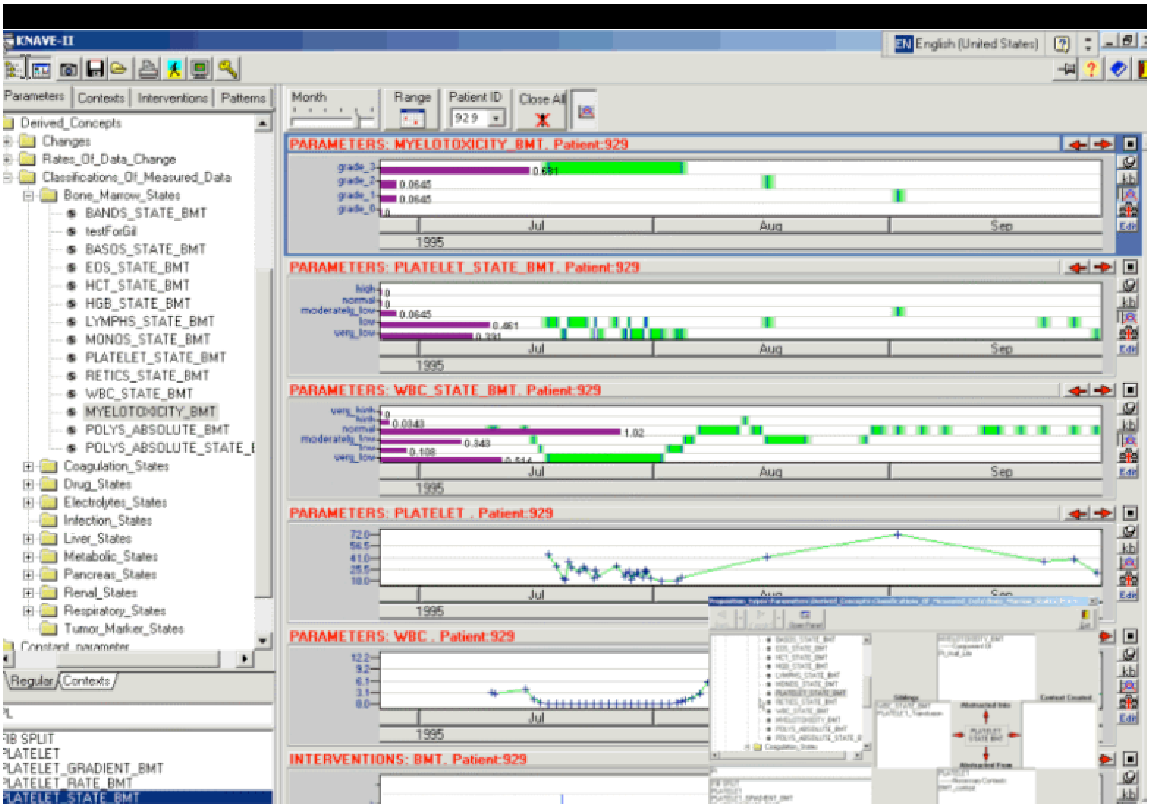

Figure 10: A Screen shot from KNAVE-II

In 2007, Gary Marchionini, Barbara K. Rimer, and Barbara Wildemuth reviewed the formats for EHRs for the US National Cancer Institute. They created mock-ups of the designs, using pregnancy as the model. Their Flat Calendar Design displayed the data for the entire month, with detail on the right hand panel. However, most of the calendar screen “real estate” is inefficiently used (Figure 11).

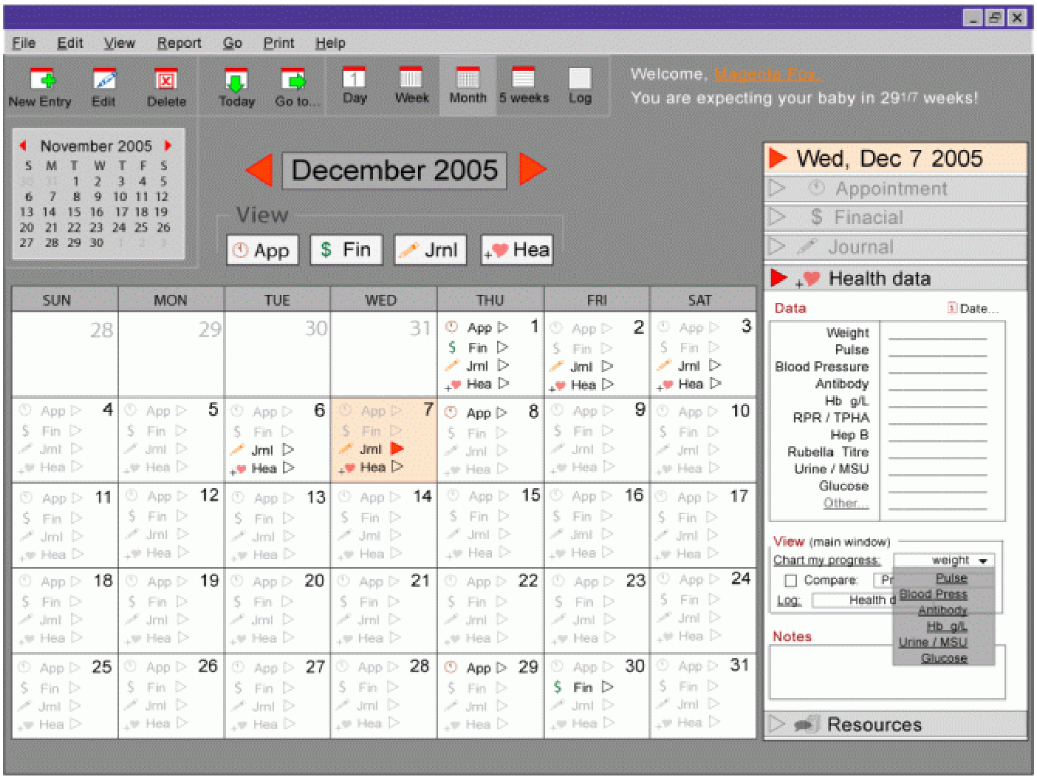

Figure 11. **The Flat Calendar Design** (Marchionini et al 2007)(see text)

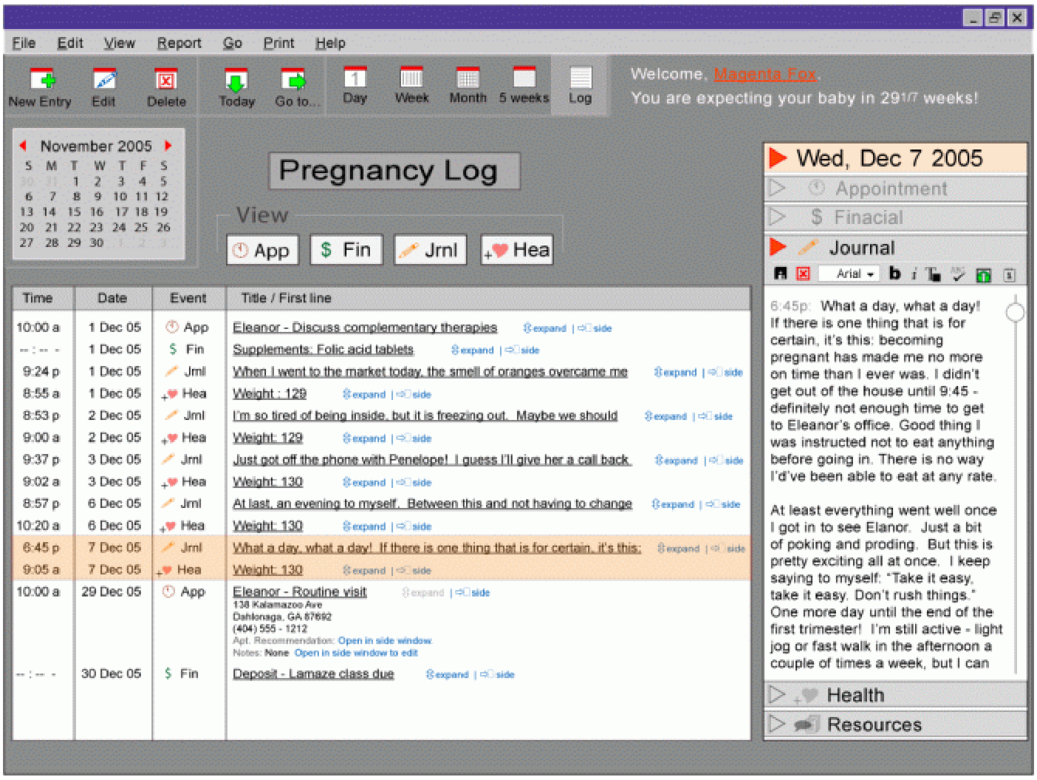

Figure 12. The Sequential (Day List) Design (Marchionini et al 2007)

The Sequential (Day List) Design sets out the events in a temporal list which can be traversed, much as in a conventional diary.

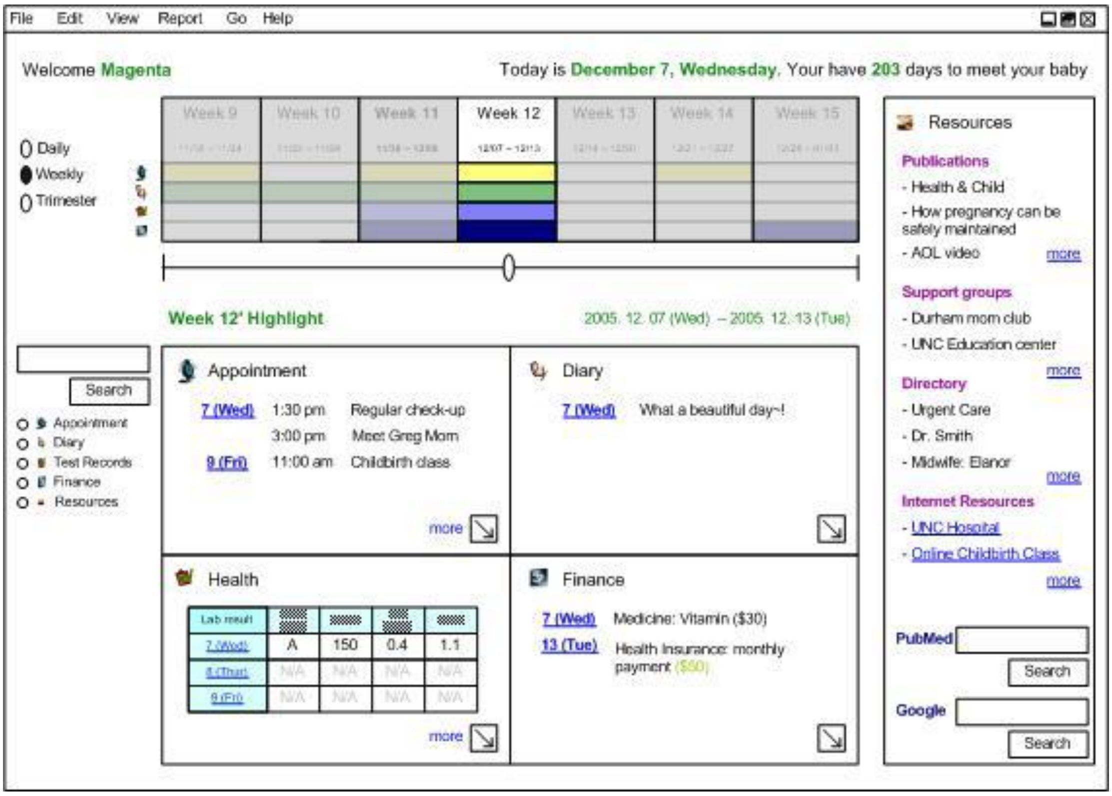

Figure 13. The Multilayered Timeline. (Marchionini et al 2007)

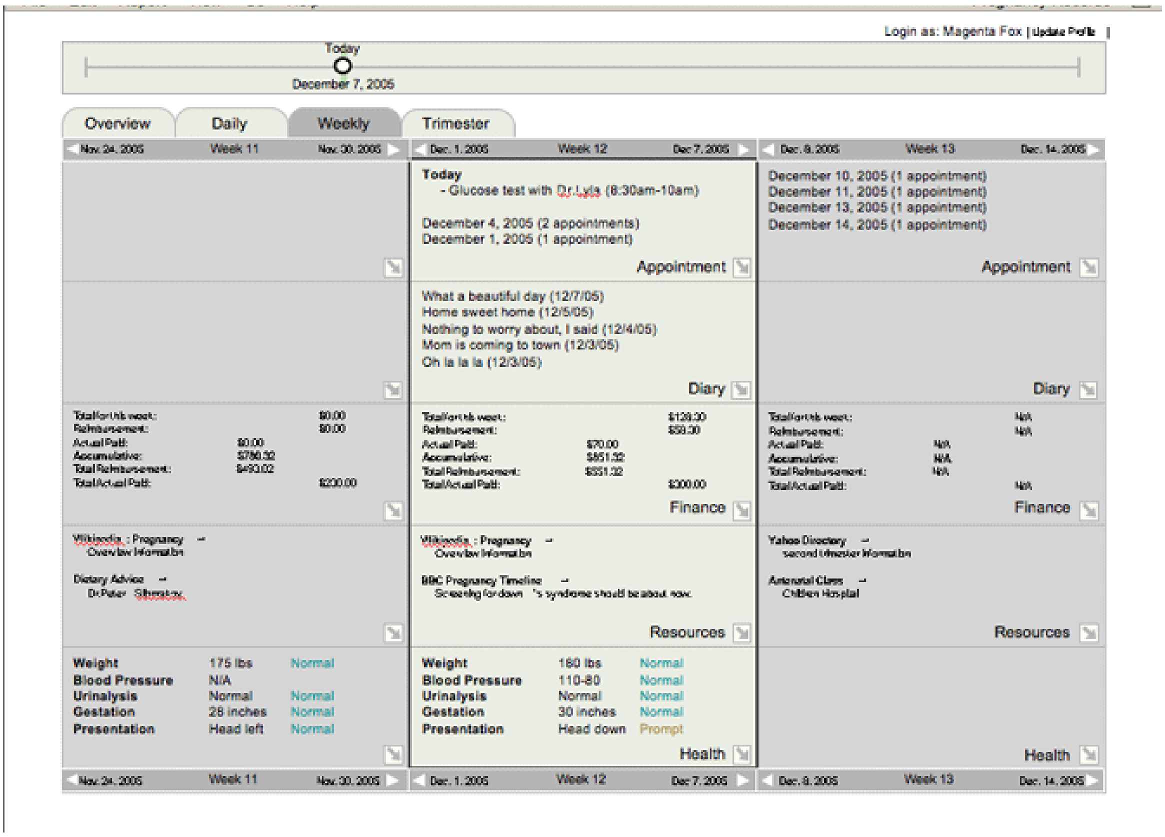

Figure 14. The Layered Timeline (Marchionini et al 2007)

In their **Multilayered Timeline model** (Figure 15), the timeline can be adjusted with a button. The remaining screen is allocated to five information facets. This design provides detailed information for a specific day at the cost of the familiarity of a calendar array. In their **Layered Timeline model** (Figure 16), they use tabs and a slider bar to expand the timeline. This user interface provides flexibility to move within and across time, but it is the least familiar navigation style.

Each of the models proposed by these authors has significant practical limitations in terms of the user experience, in that they do not present information to the user in a practical or intuitive format such that the EPR can be easily read.

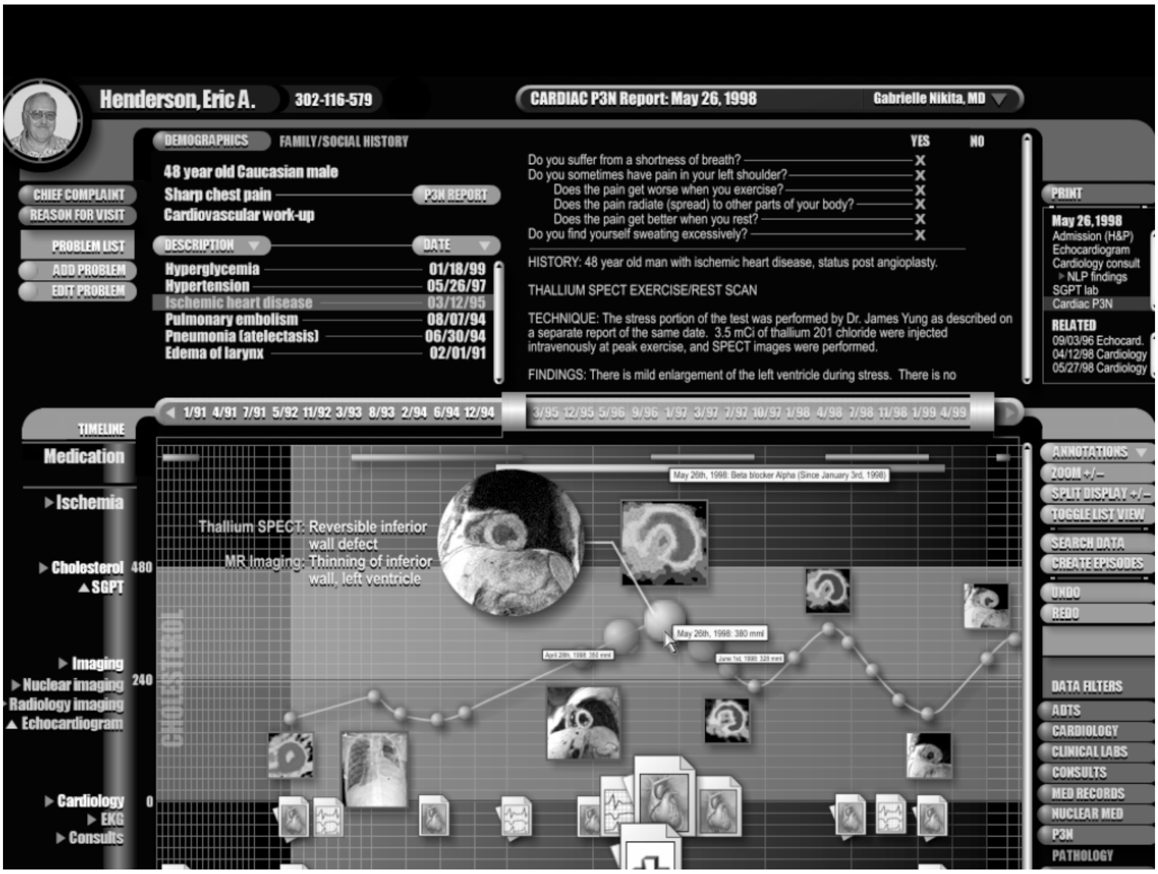

Figure 15: The TimeLine model (Bui et al 2007) A screenshot of the graphically rich, time structured interface. The model used here is a cardiac history.

The TimeLine system (Figure 15) was proposed by Alex Bui and colleagues in 2007. They described it as a problem-centric temporal visualization of patient records, in which the medical records were reorganised around medical disease entities and conditions. The contents of the EHR were integrated, reorganised, and displayed within the user interface along a timeline. The different elements of the EHR are grouped along the y-axis: imaging, reports, lab tests, etc are collapsible categories. The TimeLine system used an eXtensible Markup Language (XML) to manage data from distributed, heterogeneous medical databases. The automatic reconstruction of the TimeLine display from existing clinical repositories followed three steps, vis:

1) data access from distributed, heterogeneous medical databases;
2) data mapping and reorganisation into hierarchical, problem-centric views; and
3) data visualisation.

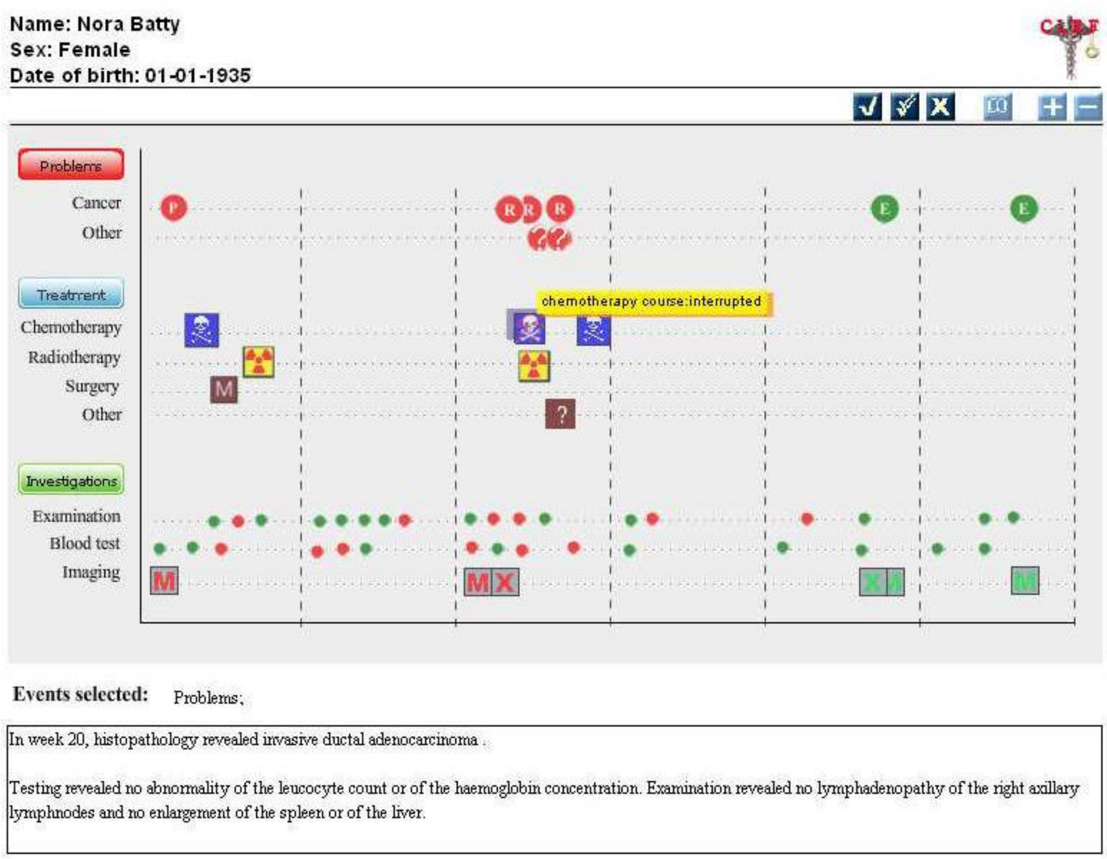

Figure 16. Low resolution image (Hallett C 2006) of the CLEF Visual Navigator

The Clinical e-Science Framework (CLEF) Visual Navigator was described by Catalina Hallett and colleagues in 2006. It depicted a high level overview of a patient's medical history by plotting events along three parallel timelines, corresponding to diagnoses, treatments and investigations, with a particular emphasis on breast cancer.

CLEF was described as a repository of detailed clinical histories for use in in-silico experiments. Their model generated a range of textual and graphical summaries of a clinical history, using a *chronicle* as the representation of the EP/HR for cancer and other patients.

In 2010, Francisco Roque, Laura Slaughter and Alexandr Tkatsenko reviewed a range of time structured information visualisation models for viewing EP/HRs, which included KNAVE II, CLEF Visual Navigator, TimeLine and Lifelines. They concluded that information visualisation systems for EHRs were frustrated by missing and inconsistent data.

The key principles in digital data visualisation

In the mid 1990s, the Human Computer Interaction Laboratory (HCIL) of the Department of Computer Science, of the Institute for Systems Research at the University of Maryland produced a rich output of novel systems for the dynamic (continuously adaptive) graphical representation of complex data sets. Critically, the work of the Laboratory was founded in a crucial set of rules for the visualisation of complex data sets, which were formulated and clearly articulated Professor Ben Shneiderman.

In a seminal paper on information visualisation, Professor Shneiderman set out his mantra for the design of advanced graphical user interfaces (Shneiderman 1996).
See: The eyes have it: A task by data type taxonomy for information visualizations.
Professor Shneiderman identified a core set of principles for the visualisation of complex data sets. He described seven key tasks for the designers of such systems. These are:

- The ability to secure an **Overview** of the entire data set, which allows the user to always orientate himself or herself as he or she navigates the data
- The ability to **Zoom In** on features of interest within the data set
The ability to **Filter Out** information (clutter) which distracts from the user's immediate task
The ability of obtain **Details on Demand** of any element of interest in the data set.
The ability to **Relate** features within the data set to each other
The ability to keep a **History** of interactions and exploration of the system
The ability to **Extract** sub-collections and query parameters

Professor Shneiderman and his team also characterised seven digital data types, vis:

- One-dimensional information, for example alphanumeric text
- Two-dimensional data, for example the space occupied by an XY graph, and
- Three-dimensional data, for example the space defined by a solid object or an XYZ graph;
- Temporal (time structured) data;
- Multi-dimensional data, of various dimensions of complexity;
- Tree structures, in which items are linked to a parent item.
- Network data sets, in which items may be linked to an arbitrary number of other items.

This work led to Shneiderman's "Eight Golden Rules of Interface Design", which have stood the test of time. https://www.cs.umd.edu/users/ben/goldenrules.html . These are:

1. Strive for consistency. Consistent actions should be required in similar situations

2. Seek universal usability. Recognise the needs of diverse users and design for plasticity

3. Offer informative feedback: For every user action, there should be feedback from within the interface, for example an actionable sound (eg a click)

4**. Design dialogues to yield closure**: Sequences of actions should be organised into groups with a beginning, middle, and end

5. Anticipate and Prevent errors: design so that users cannot make serious errors.

6. Permit easy reversal of actions: This feature relieves anxiety, since users know that errors can be undone, and it encourages the exploration of unfamiliar options.

7. Keep users in control: users must sense that they are in charge of the interface and that the interface responds to their actions

8. Reduce short-term memory load: The rule of thumb is that people can remember "seven or so chunks" of information). This requires designers to avoid interfaces in which users must recall and carry over information from one display to another display.

The HCIL Maryland FilmFinder Model

My exploration of the outputs of Maryland HCIL team led first to Filmfinder (Figure 17). This was a fun project to condense all of the Hollywood film outputs from 1960 to 1995 onto a single computer screen in a dynamic fashion, such that films were colour coded by category and they could be further searched by a number of variables.

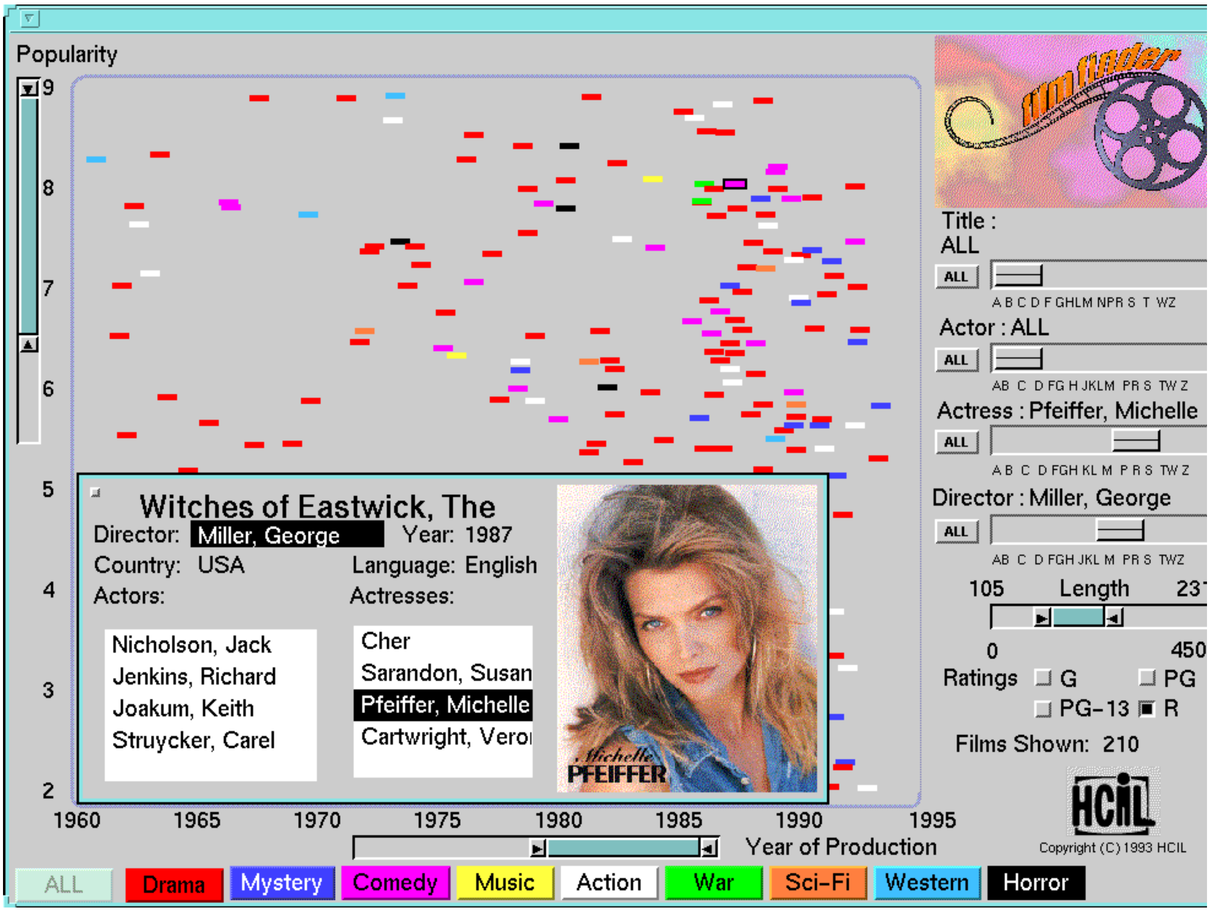

Figure 17. Screenshot of an output from The HCIL University of Maryland Filmfinder project

This elegant design pointed up a number of key elements which educated our search in Southampton for the design conundrum of the universal EPR interface, in that:

1. The two dimensional X-Y graphic could be used to display a large quantity of information in a dynamic and interactive fashion

2. The single dynamic icon could be used as a window (or a shutter) to a large quantity of information, in this case the detail on the individual film

3. Related data sets could be sub-classified in this data set by visual clues, in this case the colour and geographic distribution of the icons on the computer screen.

I realised that dynamic Iconography could be used as a powerful technique for the efficient organisation and presentation of documents, events, images or multimedia files on the computer screen as a key design element of the EPR.

Such functionality is well understood by all users of modern computers and hand held devices. Graphical imagery on icon can give further indication as to its content and function, such as the recycle bin icon on a computer screen. Dynamic activity of the icon, such as flashing or shaking, can be used to draw attention to an icon whose content which requires action. Icons can be further organised in patterns, vertically, horizontally and linearly or otherwise. Timestamp metadata allows icons to be serially displayed linearly on one or more timelines

The Maryland HCIL LifeLines Interface and the origins of UHS Lifelines

I was now very close to discovering the solution to the search for a visualisation framework for the EPR on which it was possible to presents the entire clinical record on a single screen in a logically structured and visually effective format; which would save the health professional a lot of time to assemble from the distributed records; and which would substantially reduce the risk of missing key information and themes in the patient's clinical record.

This solution was compellingly described in a format which displayed data in a two dimensional graphical concept of digital data visualisation in theoretical form by the team at the University of Maryland HCIL.

HCIL's "LifeLines" was originally developed between 1994 and 1996 as a research project for the Maryland Department of Juvenile Services for the graphical presentation study of young persons' offending records. In pulling together disparate records, it offered a better understanding of the development of criminality and to identify patterns of individual behaviour (Figure 18). (http://www.cs.umd.edu/hcil/youth-services/). This system adopted features of Filmfinder. Importantly, it also introduced Content Categorisation timelines , which in this case were described as Cases, Placements, Assignments and Reviews.

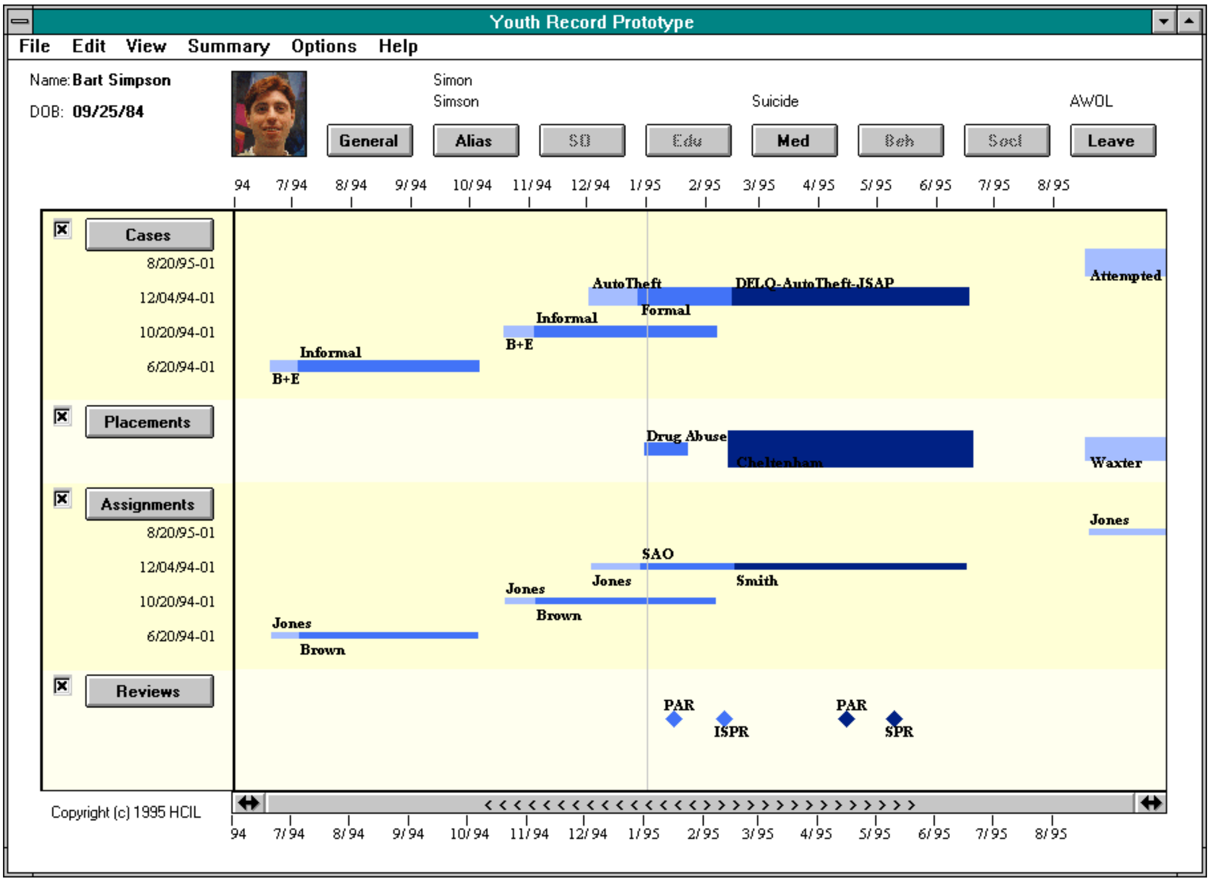

Figure 18: A life of crime: The evolving criminal career of Bart Simpson

http://www.cs.umd.edu/hcil/trs/97-19/97-19.html: Courtesy of HCIL, Univ. Maryland

The Juvenile Services Study evolved into a wider study into the information architecture required to support the visualisation of personal medical histories, see: https://hcil.umd.edu/publications/

The following screenshots shows alternative formats of data presentation, and the effect of calling up letters and handwritten clinical records. Figure 19 illustrates an early iteration of this theme, in which timelines are used with a subject taxonomy to provide structure to the medical record.

However, the use of continuous bars to describe continuing ailments and events is an inefficient use of screen space and it is an impractical display device for the description of chronic conditions rather than unitary events and documents. The co-display of pages of alpha-numeric text adds to the sense of clutter.

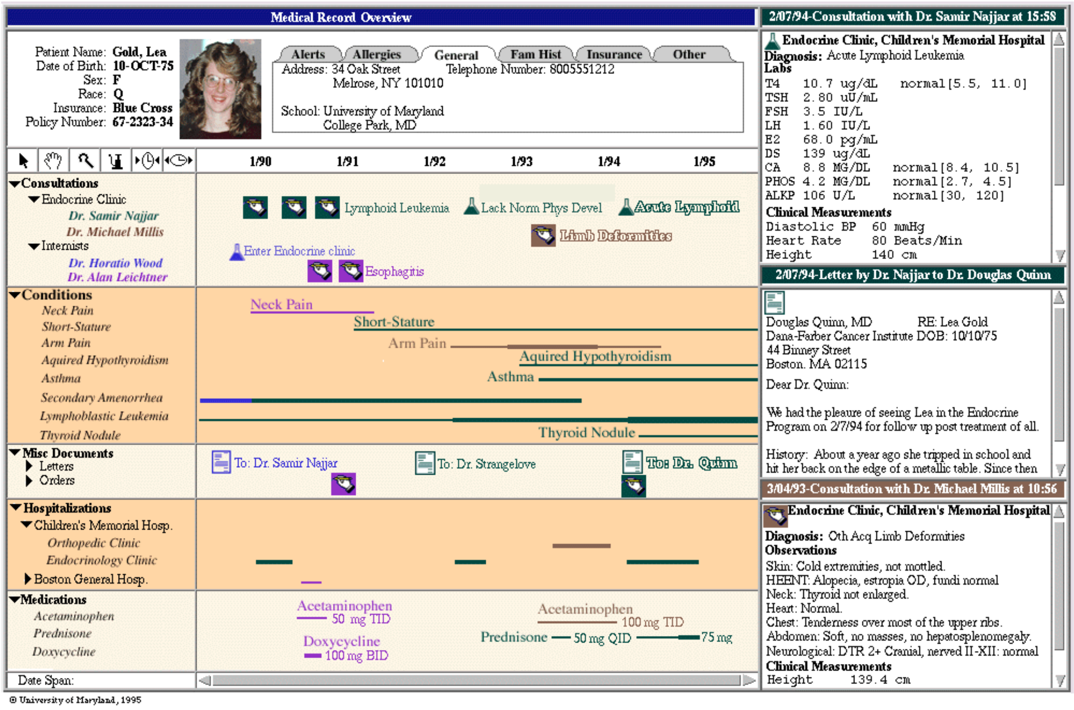

Figure 19. An early concept rendition of the EPR by the HCIL Maryland Team

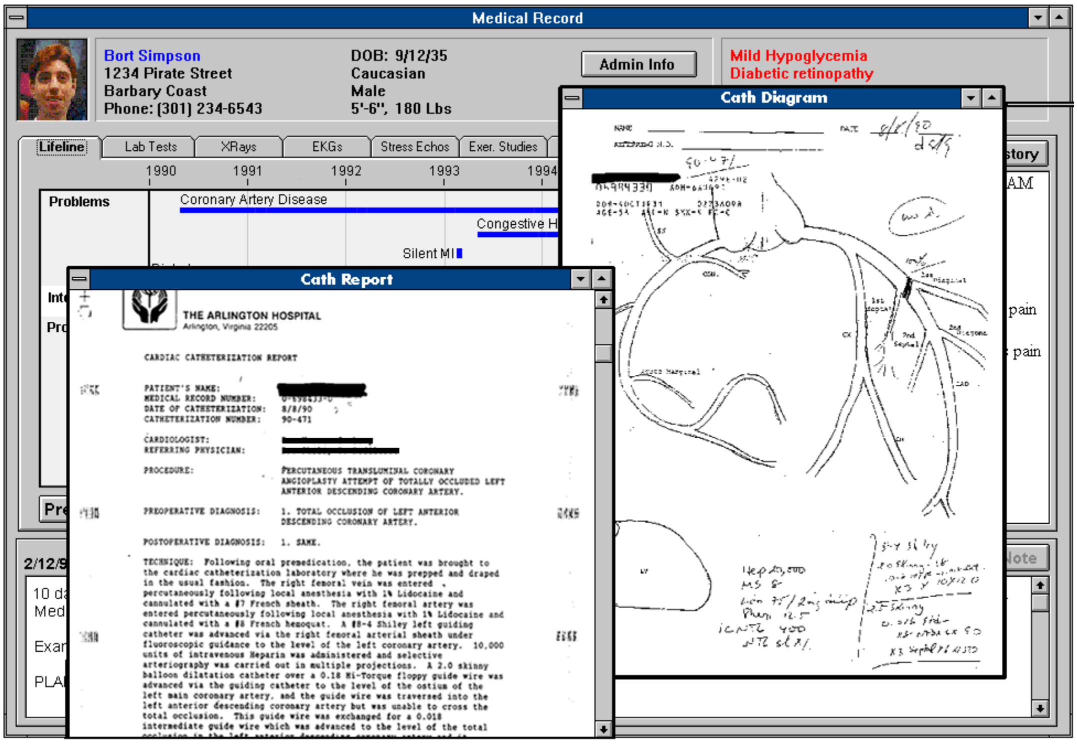

Figure 20: This image illustrates an evolution of the display model shown in the previous figure. In this version, the continuous display bars are retained, but individual documents, reports and images are now displayed independently on command in frames. This reduces the clutter on the primary interface.

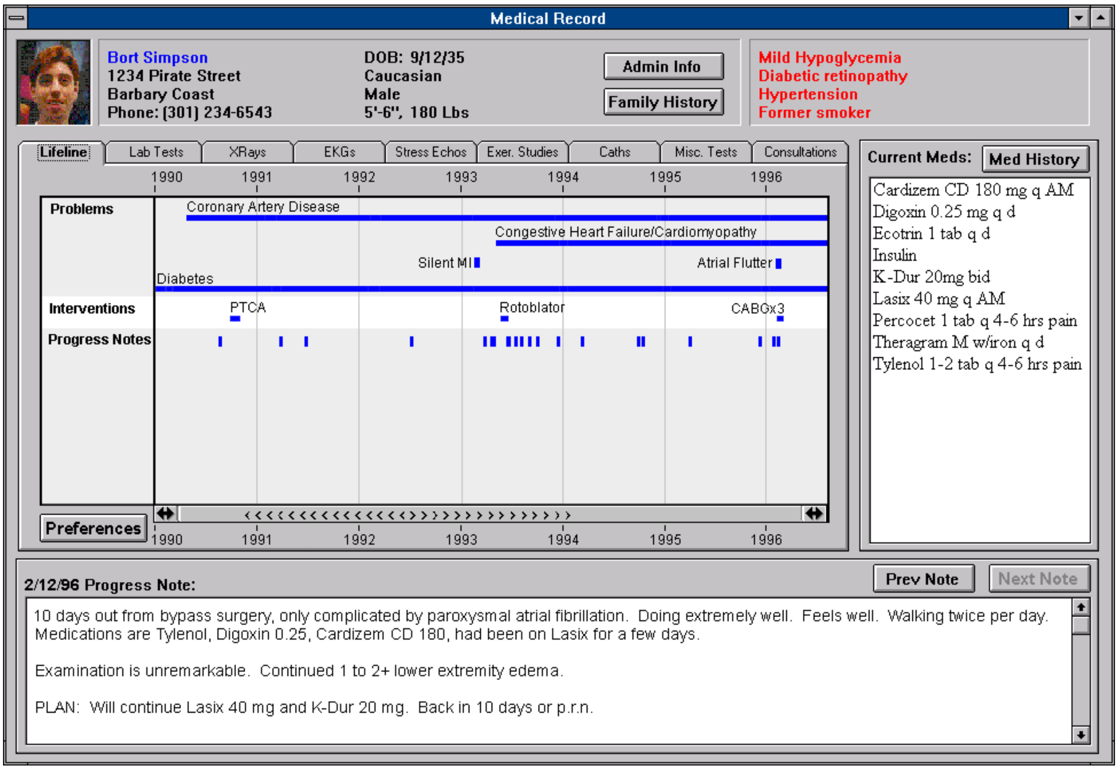

Figure 21: A Further exploration of LifeLines for Visualizing Patient Records Copyright University of Maryland, 1996-97: Source w3.cs.umd.edu/hcil/lifelines/medprints.shtml

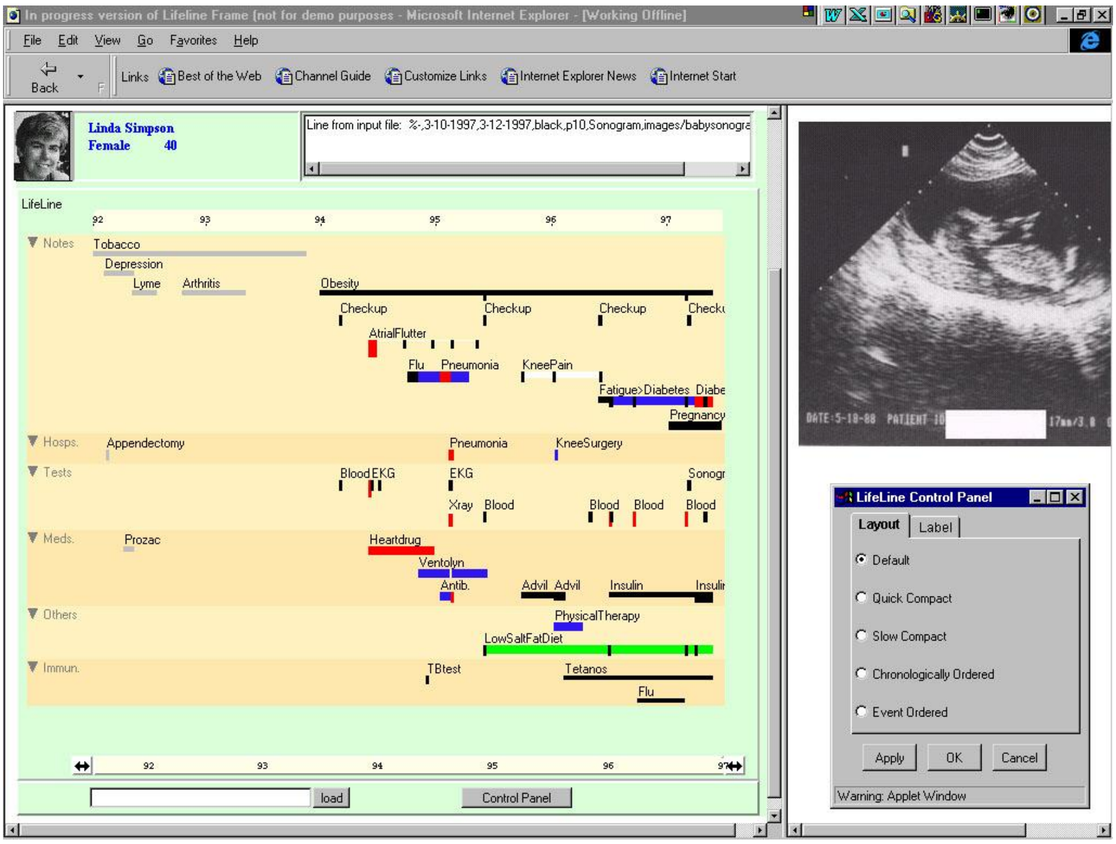

Figure 22: HCIL Lifelines as a concept interface for the EPR or Linda Simpson

The “Apex” Model of the HCIL Maryland Lifeline EPR interface

Figure 22 illustrates the “mature” format of the HCIL Maryland Lifeline EPR primary care interface. This is much cleaner, and more records are displayed as small actionable icons on the timelines. However, the bar format of data representation is retained, as is text insertion to describe the purpose of each icon.

The interfaces represent test data and to the best of our knowledge the various models never progressed to development in a live clinical system. More information on these visual models can be found on the web pages of “LifeLines for Visualizing Patient Records Copyright University of Maryland, 1996-97: Source: cs.umd.edu/hcil/lifelines/medprints”

The Key Features of the HCIL Lifelines EPR concept interface

In this model, the user is given a series of options to configure the display, which include overviews, chronological ordering, event ordering, and movement along the time line. It is easy to see how much information is immediately conveyed about this individual in respect of the frequency of contacts with the health services; the time, range and nature of concurrent and serial health conditions, and even such visually rich information as a photograph of the patient.

The Lifelines Model therefore contains a number of visualisation techniques and design concepts which eluded mainstream EPR designers for a further 15 years or so, such that:

- The entire data set is rendered on a single screen
- The X axis is a user definable master timeline which coordinates dates and events,
-The Y axis is used to stack subject timelines in parallel, according to a predefined taxonomy
- Individual documents and events are defined by a range of colours and by actionable icons

- clicking on the icons opens up the source documents, records and graphics, thus allowing the user to enter the underlying documents, which are in effect a third dimension of data.
- The placement of dynamic (interactive) icons on the timelines, which allows the direct opening of specific documents and images.
- the temporal relationships between documents and events can be readily visualised.

Essay Summary

Through a curiosity driven interest in the art and science of data visualisation, by 2009 I had identified the University of Maryland HCIL "LifeLines" project as a suitable precursor candidate for an EPR interface solution in Southampton,.

From 2009 onwards, Alan Hales and I were able to build upon this concept work to create a practical working system in the University Hospital Southampton Clinical Data Estate, with the help and support of colleagues in the Hospital IT Department.

In the next essay, I describe the early practical evolution of the concept into a working EPR interface. I have been very pleased to acknowledge the origins of the Southampton system by calling it UHS Lifelines, and I have been in regular communication with Professor Shneiderman and with his principal collaborator, Professor Catherine Plaisance, over the years.

Selected References and Further Reading

Anon. Evidence Base for Personal Health Record Usability: Final Report to the National Cancer Institute: University of North Carolina at Chapel Hill. February 10, 2007

Bui, A., Aberle, D.R., Kangarloo, H., 2007. TimeLine: Visualizing Integrated Patient Records. IEEE Transactions on Information Technology in Biomedicine 11(4): 462-473.

Chittaro L and Combi C. Representation of Temporal Intervals and relations.

Electronic Health Record Overview. The National Centre for Research Resources, National Institutes of Health April 2006

DOI: 10.1111/j.1471-8286.2005.01135.x Published by the MITRE Corporation for the NIH

Fails, J., Karlson, A., Shahamat, L., Shneiderman, B. A Visual Interface for Multivariate Temporal Data: Finding Patterns of Events over Time in *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology* (VAST 2006) 167-174

Gawande A. Why doctors hate their computers. New Yorker Magazine, November 12th 2018

Glasspool, D. W., Fox, J., Castillo, F. C. and Monaghan, V. E. L. (2003). Interactive Decision Support for Medical Planning. In M. Dojat, E. Keravnou and P. Barahona (Eds.): *Artificial Intelligence in Medicine*. Proceedings of the 9th Conference on Artificial Intelligence in Medicine in Europe, AIME 2003. Lecture Notes in Artificial Intelligence (LNAI 2780), Berlin: Springer-Verlag. pp. 335-339.

Goren-Bar, D., Shahar, Y., Galperin-Aizenberg, M., Boaz, D., Tahan, G., 2004. KNAVE II: the definition and implementation of an intelligent tool for visualization and exploration of time-oriented clinical data, AVI '04: Proceedings of the working conference on Advanced visual interfaces, pp. ACM--174.

Hallett C, Power R, Scott D. Summarisation and Visualisation of e-Health Data Repositories Computing Research Centre, The Open University. UK E-Science All-Hands Meeting, Nottingham, UK. (2006)
Source: informatics.sussex.ac.uk/users/drs22/projects/CLEF/publications/AHM06.pdf

Hallett, C., 2008. Multi-modal presentation of medical histories, IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces, pp. ACM--89.

Lindwarm D., Rose, A., Plaisant, C., Norman, K. Viewing personal history records: A comparison of tabular format and graphical presentation using LifeLines (May 1997) CS-TR-3795, UMIACS-TR-97-45. Revised version appeared in Behaviour and Information Technology, 17, 5, 1998, 249-262

Marchionini G, Rimer BK, and Wildemuth B. Evidence Base for Personal Health Record Usability: Final Report to the National Cancer Institute: University of North Carolina at Chapel Hill. February 10, 2007 Citeseer: ils.unc.edu/phr/files/final report

National Institutes of Health National Centre for Research Resources ncrr.nih.gov/publications/informatics/EHR.pdf accessed 27.08.2010

Plaisant, C., Lam, S., Shneiderman, B., Smith, M., Roseman, D., Marchand, G., Gillam, M., Feied, C., Handler, J., Rappaport, H. (May 2008). Searching Electronic Health Records for Temporal Patterns in Patient Histories: A Case Study with Microsoft Amalga. AMIA 2008 Fall symposium (cs.umd.edu/hcil/patternFinderInAmalga/)

Plaisant, C., Milash, B., Rose, A., Widoff, S., Shneiderman, B. (Sept. 1995). Life Lines: Visualizing personal histories. ACM CHI '96 Conference Proc. (Vancouver, BC, Canada, April 13-18, 1996) 221-227

Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., Shneiderman, B. LifeLines: Using Visualization to Enhance Navigation and Analysis of Patient Records. 1998 American Medical

Informatic Association Annual Fall Symposium (Orlando, Nov. 9-11, 1998) AMIA, Bethesda MD, pp. 76-80

Plaisant, C., Rose, A. (March 1996). CS-TR-3620, CAR-TR-819. 1996 American Medical Informatic Association Annual Fall Symposium (Washington, DC, Oct. 26-30, 1996), pp 884, AMIA, Bethesda MD.

Plaisant, C., Shneiderman, B., Mushlin, R. An Information Architecture to Support the Visualization of Personal Histories CS-TR-3855, UMIACS-TR-97-87. Information Management and Processing, 34, 5, pp. 581-597, 1998.

Rew DA. The Clinical Informatics Revolution. An issue in Professional Practice for the Association of Surgeons of GB and Ireland. April 2015
http://www.publications.asgbi.org/ciiipp_04_15/clinical_informatics_iipp.html

Rogoway T. Boeing Is Being Paid $84 Million Just For Manuals For New Air Force One Jet
The War Zone website, April 15th 2020
https://www.thedrive.com/the-war-zone/33034/just-manuals-from-boeing-for-new-air-force-one-jets-cost-a-whopping-84-million

Roque FS, Slaughter L, Tkatsenko A. A Comparison of Several Key Information Visualization Systems for Secondary Use of Electronic Health Record Content: Proceedings of the NAACL HLT 2010 Second Louhi Workshop on Text and Data Mining of Health Documents, pages 76–83, Los Angeles, California, June 2010. c 2010 Association for Computational Linguistics
Source: aclweb.org/anthology/W/W10/W10-1112.pdf

Sharar Y & Musen MA: Knowledge based temporal abstraction in clinical domains
SMI-95-561 CiteSeer

Shahar Y, Goren-Bar D, Boaz D, Tahan G. Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions. Artif Intell Med. 2005 Dec 9

Shneiderman, B., The eyes have it: A task by data type taxonomy of information visualisations, *Proc. IEEE Symposium on Visual Languages '96*, IEEE, Los Alamitos, CA (September 1996), 336-343.

Shneiderman B: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. , Department of Computer Science, Human-Computer Interaction Laboratory, and Institute for Systems Research, University of Maryland July 1996
Source: citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.25.7925

Wang TD, Plaisant C, Shneiderman B: *Lifelines2*: Discovering Temporal Categorical Patterns Across Multiple Records http://www.cs.umd.edu/hcil/lifelines2 accessed 17.10.2009

Wang TD, Plaisant C, Shneiderman B, Spring N, Roseman D, Marchand G, Mukherjee V, and Smith M. Temporal Summaries: Supporting Temporal Categorical Searching, Aggregation and Comparison. HCIL technical report(HCIL 2009-09). March, 2009.

Wang TD, Plaisant C, Quinn A, Stanchak R, Shneiderman B, and Murphy S. Aligning Temporal Data by Sentinel Events: Discovering Patterns in Electronic Health Records. SIGCHI Conference on Human Factors in Computing Systems (CHI 2008).

Wongsuphasawat, K. and Shneiderman, B. Finding Comparable Temporal Categorical Records: A Similarity Measure with an Interactive Visualization: Proceedings of IEEE Symposium on Visual Analytics Science and Technology (IEEE VAST), 2009.