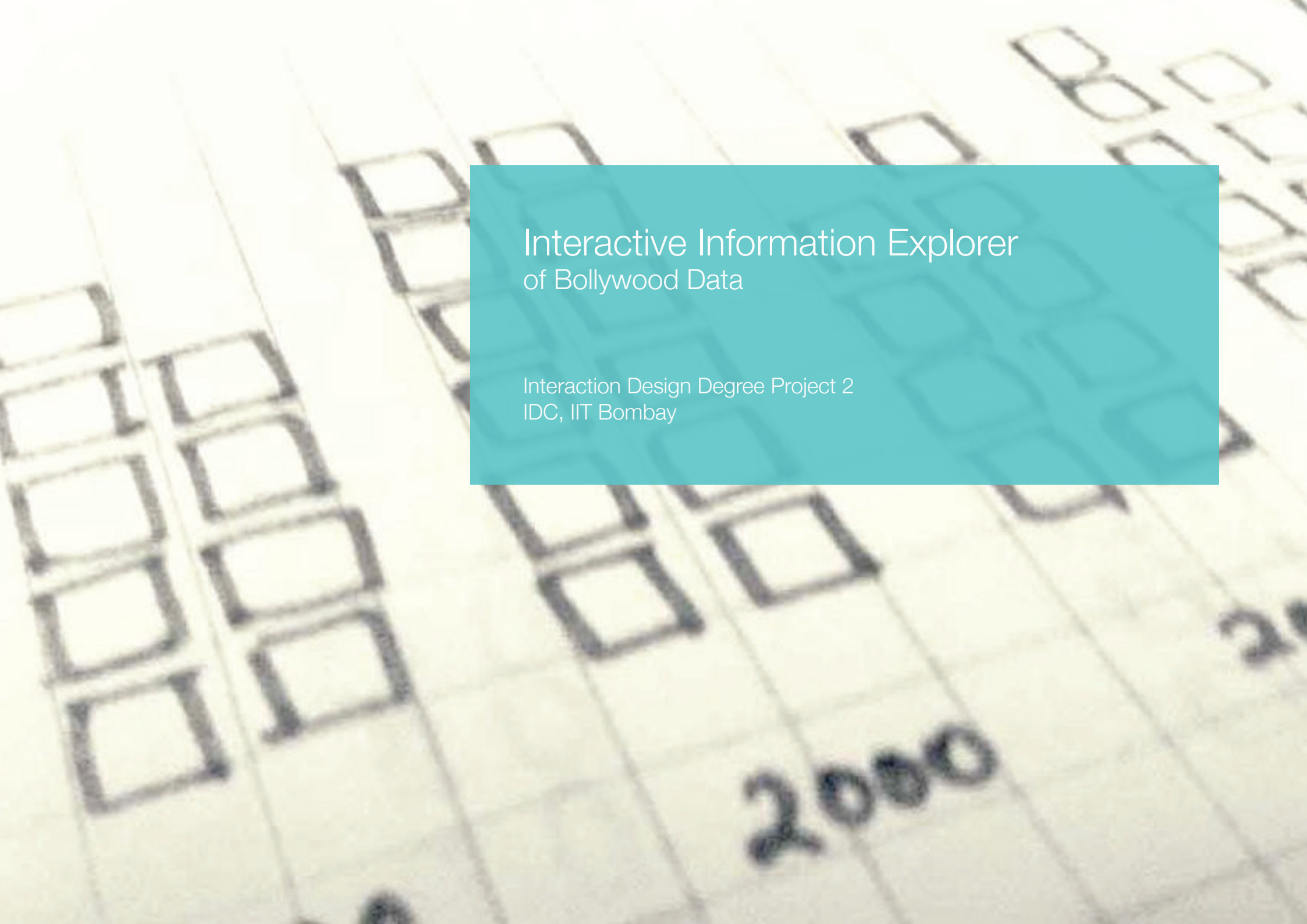
Interaction Design Degree Project 2 on

Interactive Information Explorer
of Bollywood Data

Pritam Pebam
126 330 003
M.Des, Interaction Design

Under the guidance of
Prof. Ravi Poovaiah

Industrial Design Centre
Indian Institute of Technology Bombay



Interactive Information Explorer
of Bollywood Data

Interaction Design Degree Project 2
IDC, IIT Bombay

Declaration

The research work embodied in the written submission titled "**Interactive Information Explorer for Bollywood Data**" has been carried out as Project 2 by the undersigned as part of the post graduate program in the Industrial Design Centre, IIT Bombay, India under the supervision of **Prof. Ravi Poovaiah**.

The undersigned hereby declares that this is an original work and has not been plagiarized in part or full from any source. Appropriate reference information or links have been provided whereever due. Furthermore, this work has not been submitted for any degree in this or any other University.

I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action if need arises.

Pritam Pebam
M.Des
Interaction Design
IDC, IIT Bombay

Approval Sheet

This interaction design project entitled "**Interactive Information Explorer for Bollywood Data**" by Pritam Pebam, 126 330 003, is approved in partial fulfilment of the requirements for Master of Design Degree in Interaction Design.

Project Guide: 

Chair Person: 

Internal Examiner: 

External Examiner: 

Date: 10/1/2014

Place: MUMBAI

Acknowledgements

I express my sincere gratitude to my guide, **Prof. Ravi Poovaiah**, for his valuable feedbacks and the resources shared over time that had helped me tremendously during the entire course of my project.

I would also like to express my gratitude to **Prof. Mandar Rane** for providing me the springboard to head on and start ideating. Furthermore, I would like to thank **Prof. Anirudha Joshi** and **Prof. Girish Dalvi** for their valuable ideas and insights during the various stage presentations of the project.

I would also like to thank my friends in IDC who has helped me with leads and suggestions to projects in similar problem areas during my secondary research.

I would also like to thank IDC IIT Bombay for providing me with all the facilities and necessary materials and an environment that encourage me to work on such a challenging project.

Pritam Pebam
Mumbai
18th November 2013

Table of Contents

Abstract

In the age of big data, accessing information, collating them and of all making sense out of them has become challenging.

Interactive Information Explorer (IIE) attempts to bridge the gap between human cognitive limitation and deciphering/accessing big data. It provides a structural model to access data, not limiting to the current trends of visualization and summarisation that prevails in the information design-sphere.

IIE provides a minimal depth UI framework through which one can access data which invokes users to explore a particular data set. During the process of ‘exploring’, users picked up subtle insights revealing patterns and establish knowledge about the data set which otherwise would’ve been hard to notice or understood.

The framework is designed to handle different varieties of data-set in an attempt to universally establish a generic structure to access information and data.

In short, IIE may be roughly compared as a more insightful version of the Windows Explorer or Mac Finder.

Introduction

Every day, billions of gigabytes of digital data are generated over the internet. Among which only a much smaller section is available in the form of structured data. This data however is too vast and big enough for the human cognitive mind to juxtapose, find patterns or relationships that helps make sense out of it.

Over time, big companies like Google, Yahoo, Microsoft etc. have tried to structure their own data and it has become one of the most valuable piece of property each company owns.

Beyond this, there comes the challenge of “How do we access this data?” Millions & Billions of bytes of data are sitting on a remote server but how are we making use of it? Can we organize data into a generic framework to help understand them?

In an attempt to give an answer, IIE was conceptualized.

IIE starts with the data-set of bollywood movie information, organises it into a structured information architecture and provides a UI framework to access the data. The data itself is presented in a collated and visualized form so as to yield insights that otherwise would have been hard to relate or be seen.

The IIE framework is attempting to not be restrained only with just Bollywood data but extend to or scale over multiple data-sets including, but not limited to, music albums, books, medical data, sports data or so on.

Microsoft Pivot's[1] Interaction Model has been one of the main influence that shaped how the internet can be accessed; a novel idea that transformed, or at least attempts to, how we organized & explore the data available online.

The fundamental question that IIE tries to answer is that "data speaks, how can we access and make sense out of it?"

"IIE is an interaction model which represents collated data in such a manner that reduces human cognitive load"

[1] Microsot Pivot TED talk by Gary Flake - http://new.ted.com/talks/gary_flake_is_pivot_a_turning_point_for_web_exploration

Project Brief

The main focus area of the project is to simplify the way data is consumed by the users. The focus user groups varies exhaustively. Not restricted to specific age groups, IIE is targeted to major day-to-day users who just needs (or wants) a holistic perspective of the whole data. In this context, bollywood movie buffs regardless of age, gender or location is targeted.

As such, IIE do not aim to handle direct manipulative inputs but provides only a generic overview of the data that can be boiled down as per the user's interest. The reason being that most users care less about hard lines that segregates different items. For example, two movies one with 4.2 rating and the other with 4.7 rating hardly makes any difference to them. From basic user study results, majority of the users are not indulged in getting minutely detailed & accurate results. Rather users have a tendency of looking into a generic overview that gives a general idea of the data-set. This drives the primary reason for restricting the level of customization available in the interaction model as the system tries to emphatically reflect the users' behaviour. As such majority of the users do not have a pre-specified objective that they plan to achieve using IIE. The use case scenario of IIE rationales its own use as mostly ***'exploratory'*** & '***discovery serendipity***'.

The nature of the problem area also has demanded to understand previous human cognitive stress as concluded from previous infographics or visualisations of similar kind.
Following which, rather than performing an exhaustive user study, an approach in which the system is created first and then evaluated from user feedback is taken up.

The primary area to be focussed while dealing with the project is to understand and create on two main areas:

Structuring of data:

Each data-set needs to be translated and transform into a shape which truly represents its internal information architecture and represent the data in a wholly accessible and understandable format. The main challenge lies in identifying a pattern among various data-sets to establish a framework that fits all these data-set.

Presentation of data:

When data is represented contextually, there are greater chances of it being understood. Similarly, various methods of infographics and visualization, summarization etc. has helped in getting into the gist of the data in the shortest of time.

Data Collection

The first step to “organizing” & “visualizing” data is to ‘acquire’[2] the data.

Obtaining data for such projects have always been a major challenge because of the scope and nature of data availability. Wikipedia has an exhaustive collection of data of which hardly only a few portion is available in the “structured” form. This data is available through APIs from www.dbpedia.com.

Likewise, the National film archives have a website[3] through which you can obtain data however it is not readily available as an API call. Another option is to collect the data manually. However given the fact that the amount of data is unsurmountable, manual collection is almost impossible.

There are couple of website which have their own database out of which Google own users curated Freebase database seems to be quite extensive and robust and readily available through API calls. The data is also available in a huge 15gb dump in case more API calls are expected.

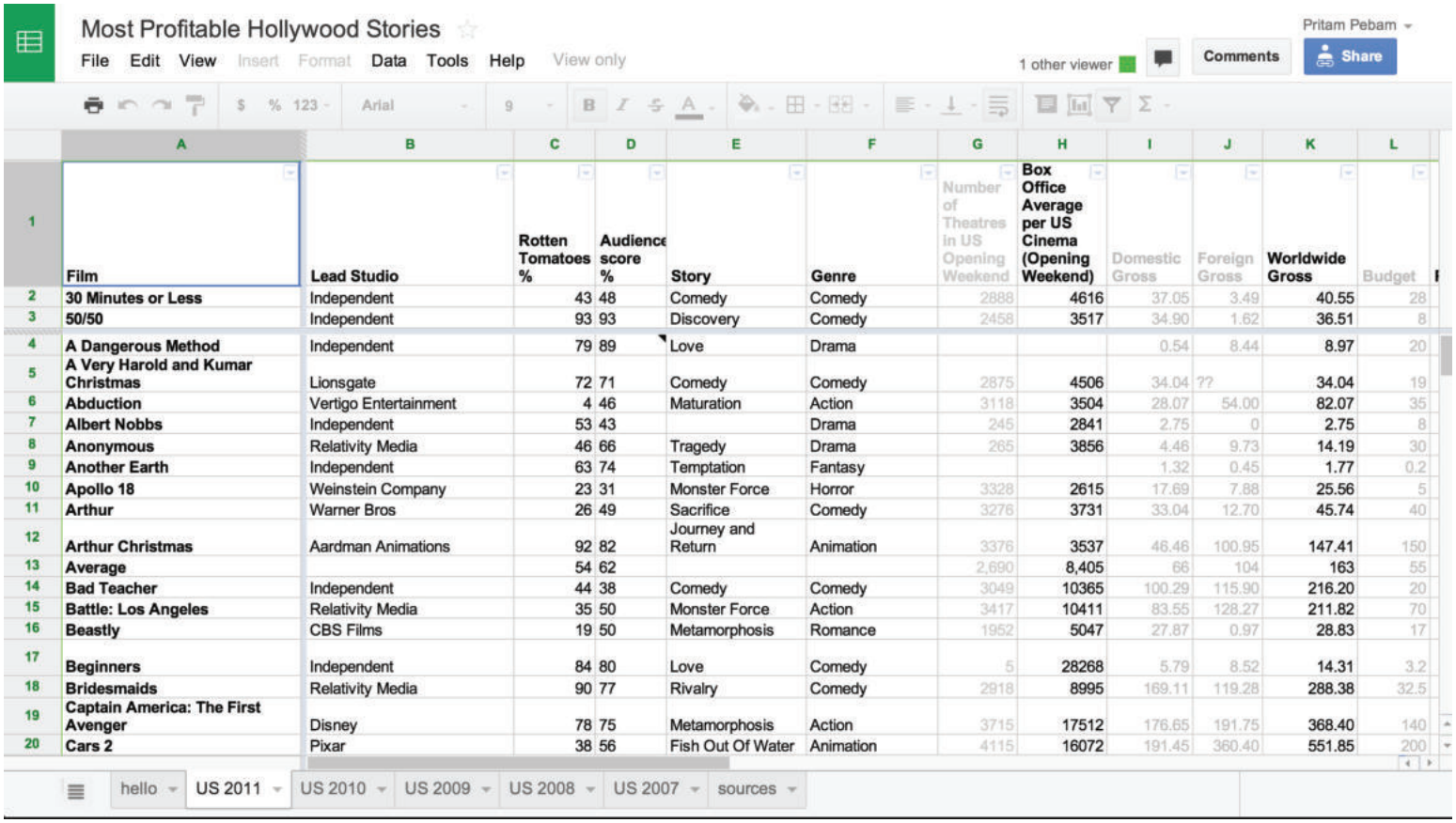
The nature of how the data is curated however yields some garbage data which makes no sense. This limitation however is outside the scope of this project. As a result, some visualized graphical elements

2 Visualizing Data by Ben Fry

3 National Film Archives website http://www.nfaipune.gov.in

may display random irrelevant data as it is fetch directly from the Freebase APIs.

To help understand how existing data are organized, the structure of the data as collected and published by '*Information Is Beautiful*'[4] was studied.



Most Profitable Hollywood Stories

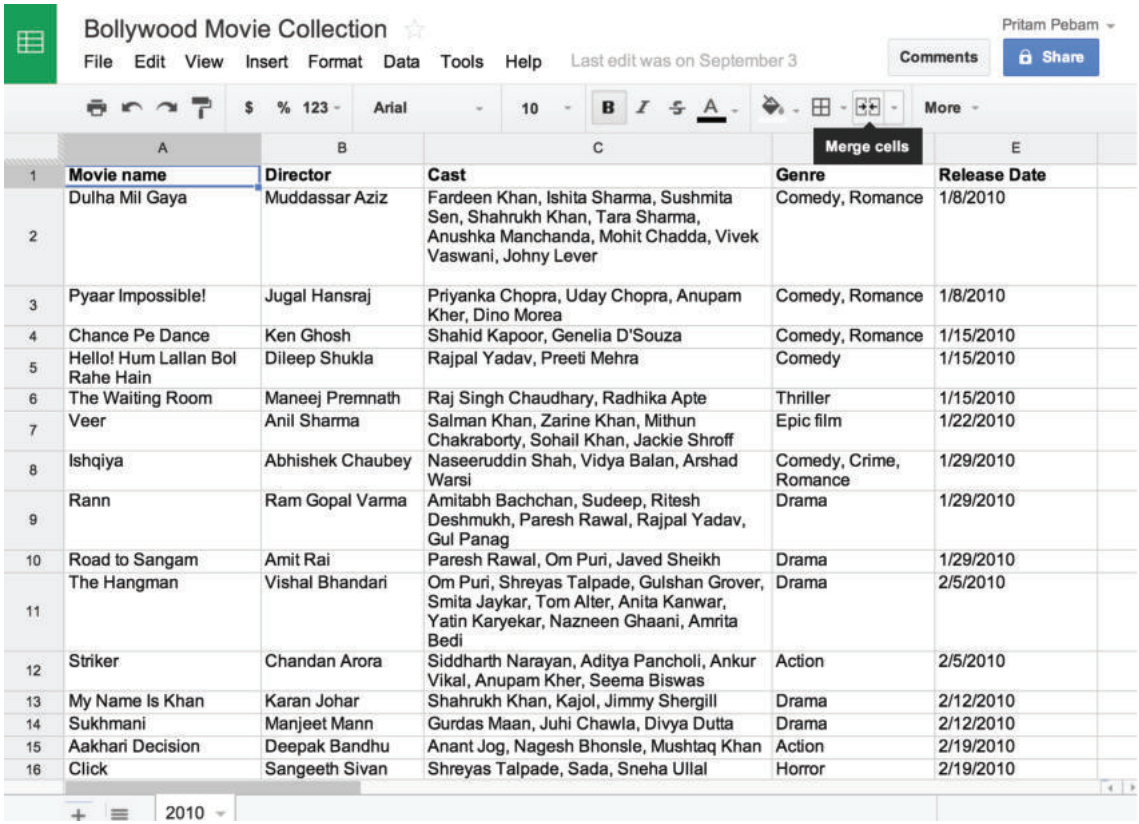
File Edit View Insert Format Data Tools Help View only — 1 other viewer — Comments — Share — Pritam Pebam

	Film	Lead Studio	Rotten Tomatoes %	Audience score %	Story	Genre	Number of Theatres in US Opening Weekend	Box Office Average per US Cinema (Opening Weekend)	Domestic Gross	Foreign Gross	Worldwide Gross	Budget
2	30 Minutes or Less	Independent	43	48	Comedy	Comedy	2888	4616	37.05	3.49	40.55	28
3	50/50	Independent	93	93	Discovery	Comedy	2458	3517	34.90	1.62	36.51	8
4	A Dangerous Method	Independent	79	89	Love	Drama			0.54	8.44	8.97	20
5	A Very Harold and Kumar Christmas	Lionsgate	72	71	Comedy	Comedy	2875	4506	34.04	??	34.04	19
6	Abduction	Vertigo Entertainment	4	46	Maturation	Action	3118	3504	28.07	54.00	82.07	35
7	Albert Nobbs	Independent	53	43		Drama	245	2841	2.75	0	2.75	8
8	Anonymous	Relativity Media	46	66	Tragedy	Drama	265	3856	4.46	9.73	14.19	30
9	Another Earth	Independent	63	74	Temptation	Fantasy			1.32	0.45	1.77	0.2
10	Apollo 18	Weinstein Company	23	31	Monster Force	Horror	3328	2615	17.69	7.88	25.56	5
11	Arthur	Warner Bros	26	49	Sacrifice	Comedy	3276	3731	33.04	12.70	45.74	40
12	Arthur Christmas	Aardman Animations	92	82	Journey and Return	Animation	3376	3537	46.46	100.95	147.41	150
13	Average		54	62			2,690	8,405	66	104	163	55
14	Bad Teacher	Independent	44	38	Comedy	Comedy	3049	10365	100.29	115.90	216.20	20
15	Battle: Los Angeles	Relativity Media	35	50	Monster Force	Action	3417	10411	83.55	128.27	211.82	70
16	Beastly	CBS Films	19	50	Metamorphosis	Romance	1952	5047	27.87	0.97	28.83	17
17	Beginners	Independent	84	80	Love	Comedy	5	28268	5.79	8.52	14.31	3.2
18	Bridesmaids	Relativity Media	90	77	Rivalry	Comedy	2918	8995	169.11	119.28	288.38	32.5
19	Captain America: The First Avenger	Disney	78	75	Metamorphosis	Action	3715	17512	176.65	191.75	368.40	140
20	Cars 2	Pixar	38	56	Fish Out Of Water	Animation	4115	16072	191.45	360.40	551.85	200

hello | US 2011 | US 2010 | US 2009 | US 2008 | US 2007 | sources

4 Information is beautiful Hollywood data - http://goo.gl/2XD8hO

Major parameters that held responsible for the system to get organised consist of budget, opening weekends, grosses, story type, genre, ratings and score etc.



	A	B	C	D	E
1	**Movie name**	**Director**	**Cast**	**Genre**	**Release Date**
2	Dulha Mil Gaya	Muddassar Aziz	Fardeen Khan, Ishita Sharma, Sushmita Sen, Shahrukh Khan, Tara Sharma, Anushka Manchanda, Mohit Chadda, Vivek Vaswani, Johny Lever	Comedy, Romance	1/8/2010
3	Pyaar Impossible!	Jugal Hansraj	Priyanka Chopra, Uday Chopra, Anupam Kher, Dino Morea	Comedy, Romance	1/8/2010
4	Chance Pe Dance	Ken Ghosh	Shahid Kapoor, Genelia D'Souza	Comedy, Romance	1/15/2010
5	Hello! Hum Lallan Bol Rahe Hain	Dileep Shukla	Rajpal Yadav, Preeti Mehra	Comedy	1/15/2010
6	The Waiting Room	Maneej Premnath	Raj Singh Chaudhary, Radhika Apte	Thriller	1/15/2010
7	Veer	Anil Sharma	Salman Khan, Zarine Khan, Mithun Chakraborty, Sohail Khan, Jackie Shroff	Epic film	1/22/2010
8	Ishqiya	Abhishek Chaubey	Naseeruddin Shah, Vidya Balan, Arshad Warsi	Comedy, Crime, Romance	1/29/2010
9	Rann	Ram Gopal Varma	Amitabh Bachchan, Sudeep, Ritesh Deshmukh, Paresh Rawal, Rajpal Yadav, Gul Panag	Drama	1/29/2010
10	Road to Sangam	Amit Rai	Paresh Rawal, Om Puri, Javed Sheikh	Drama	1/29/2010
11	The Hangman	Vishal Bhandari	Om Puri, Shreyas Talpade, Gulshan Grover, Smita Jaykar, Tom Alter, Anita Kanwar, Yatin Karyekar, Nazneen Ghaani, Amrita Bedi	Drama	2/5/2010
12	Striker	Chandan Arora	Siddharth Narayan, Aditya Pancholi, Ankur Vikal, Anupam Kher, Seema Biswas	Action	2/5/2010
13	My Name Is Khan	Karan Johar	Shahrukh Khan, Kajol, Jimmy Shergill	Drama	2/12/2010
14	Sukhmani	Manjeet Mann	Gurdas Maan, Juhi Chawla, Divya Dutta	Drama	2/12/2010
15	Aakhari Decision	Deepak Bandhu	Anant Jog, Nagesh Bhonsle, Mushtaq Khan	Action	2/19/2010
16	Click	Sangeeth Sivan	Shreyas Talpade, Sada, Sneha Ullal	Horror	2/19/2010

Inferring to it, as a start, a small set of Bollywood data is collected whose parameters consists of the name, director, casts, genre and release date. As manual collection of the data can not be the final solution, hence steps for using dynamic API calls from freebase database (or whichever data providers) is considered.

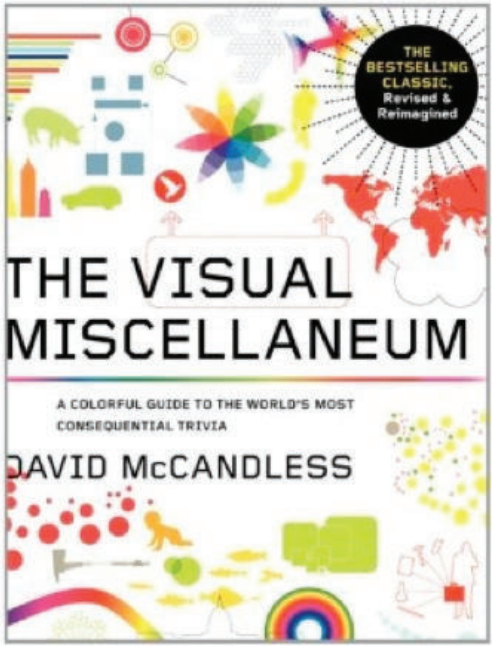
IIE handles movie data from time-period ranging from 1980 to 2013 that has been collected from Wikipedia. The data contains the elements above.

What is IIE?

Several infographics or data visualisations of different kinds and format are readily available in various forms. The main question “Why IIE?” lies in the fundamental need of accessing data in a fathomable approach. This can be achieved by coupling an interactive infographics/data visualisation system along with a solid Information Architecture.

One of the main challenge of visualizing data is its unique nature - that different data-sets have different structure and making a visualization of a particular data-set is considered an art in itself. The same method usually is not applicable on another data-set and hence certain information architectural structure needs to be build as the foundation so that data, in whichever form, be able to adapt to the interaction model.

IIE creates a generalised framework that can help represent different data-set and fit in one UI model. This model is scalable and thus help cover different structural data-sets into the framework.



5 The Visual Mescellaneum by David McCandless

6 Feltron Report

7 Interactive Analysis of Big Data

8 Stat! – Interactive Analytics Environment

Research

Secondary Research

As part of the secondary research, numerous papers on Big-Data visualizations are studied.

David McCandles's '*Information is beautiful*', the *Feltron Report* as well as sites like *Visual.ly* yields several visualization inspirations and insights on how to organize and represent big data.

As a result of the secondary research, the 2 main Primary Research area are narrowed down as:

Understand the building block
To create an interaction model from scratch, a thorough understanding of it's building block needs to be understood essentially. As an example, as the name and the definite of IIE suggestion, one needs to identify what is 'interactive', what is 'Information' and what is 'explorer'.

Understanding Microsoft Pivot
Microsoft Pivot is a Interaction Model which retrieves data, organizes them and deliver a rich visual and graphical representation of it.
It makes it easier for people to get insights over large chunk of data.

Analyse Existing Visualizations
An analysis of existing products which solves or try to solve similar project needs to be considered before starting with the concepts.

The results of these areas is discussed ind detail under Primary research header.

Research:

Primary Research

As part of the primary research, several studies over existing visualizations & interaction models are conducted.

Apart from an understanding of data visualization methods and interaction models (discussed in the concepts section), one of the seemingly mundane but important finding from the primary research is a dangerous pitfall while begining designing with data.

"Average infographics focuses more on the overall aesthetic of the piece, the 'cool' factor, than on the data itself"[9]

9 A Common Misconception About Designing With Data - http://www.smashingmagazine.com/2013/07/29/common-misconception-designing-data/

As inferred from Secondary research results, the following focus areas are discussed.

Understanding the building Blocks

From the two points, there arose two basic research areas to focus on:

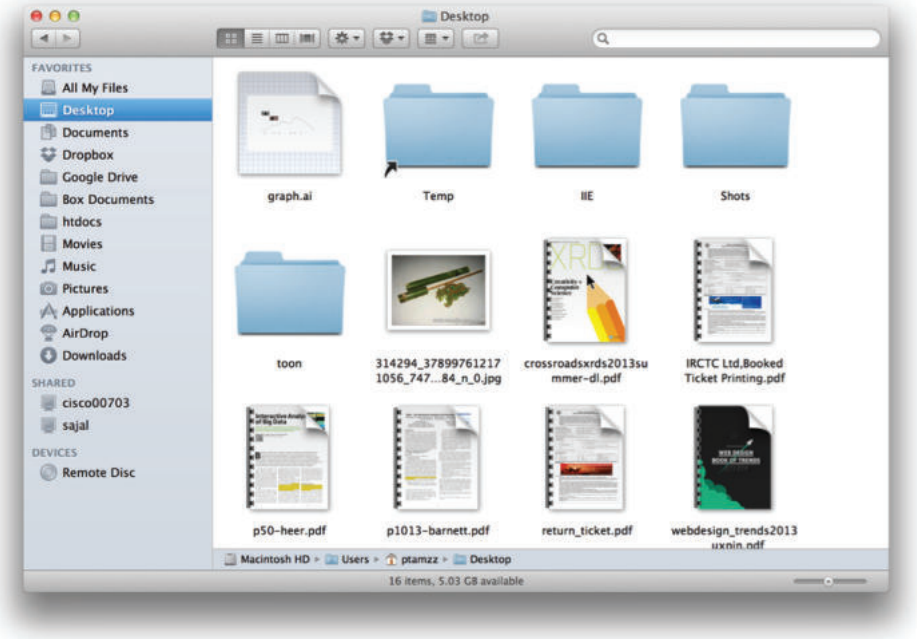
1) **Structure**:
   How are we structuring and accessing data curently?

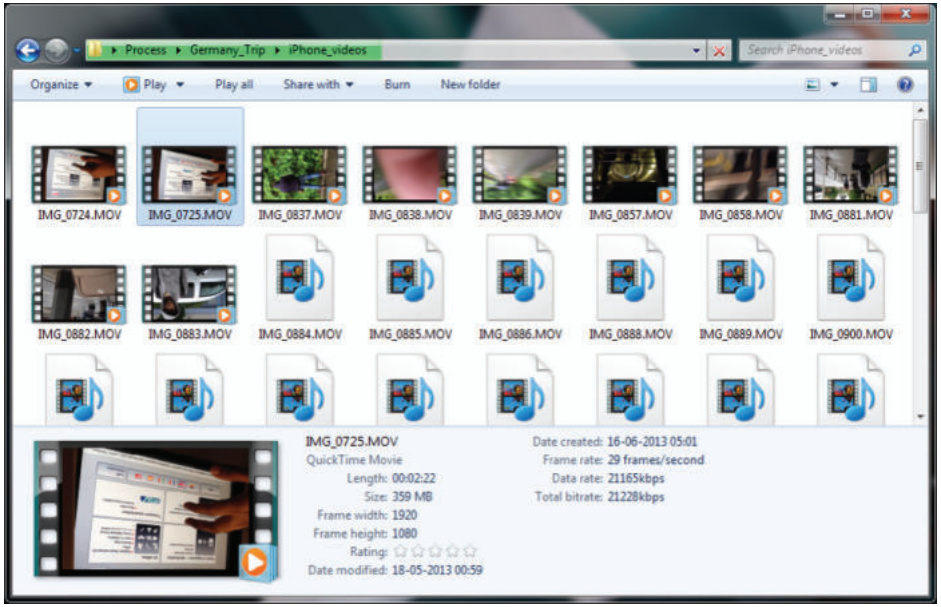
2) **Presentation**:
   What is the best way to represent data to reduce the cognitive load?

For a start, going back to the basics, to understand and answer the two questions above, IIE needs to be broken down to what it stands for and study its building blocks.

Considering the term ‘explorer’ in ‘Interactive Information Explorer’, one needs to understand what an explorer is and what purpose does it serve.

As such, we have a very familiar file explorer that every computer user is aware of - the Windows Explorer or the Mac Finder.

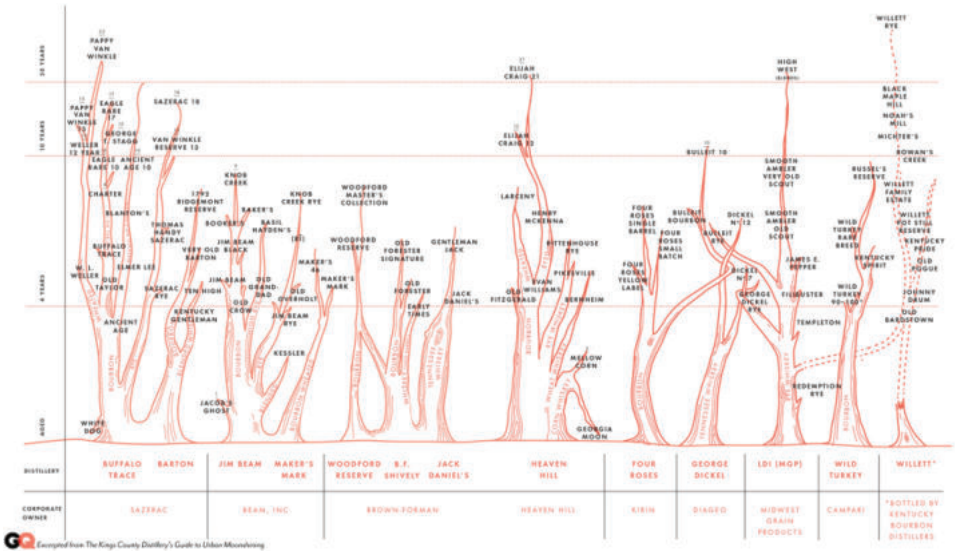




"Windows Explorer is a file manager application and also a navigation tool.... It provides a graphical user interface for accessing the file systems"[10]

[10] File Explorer - http://en.wikipedia.org/wiki/File_Explorer

11 http://visual.ly/what-each-country-leads-world



12 http://visual.ly/bourbon-family-tree

"Information Visualization is the study of (interactive) visual representations of abstract data to reinforce human cognition"[11]

11 Information Visualization - http://en.wikipedia.org/wiki/Information_visualization

As concluded, the two possible ways to achieve what IIE aims to serve its purpose can be implemented as

File Explorer as Structural Organization

The existing Windows or Mac explorer organizes data into a linear architecture and stack data in layers. These data is accessible through a straight linear interaction. IIE can solve the main issue using this exploratory model.

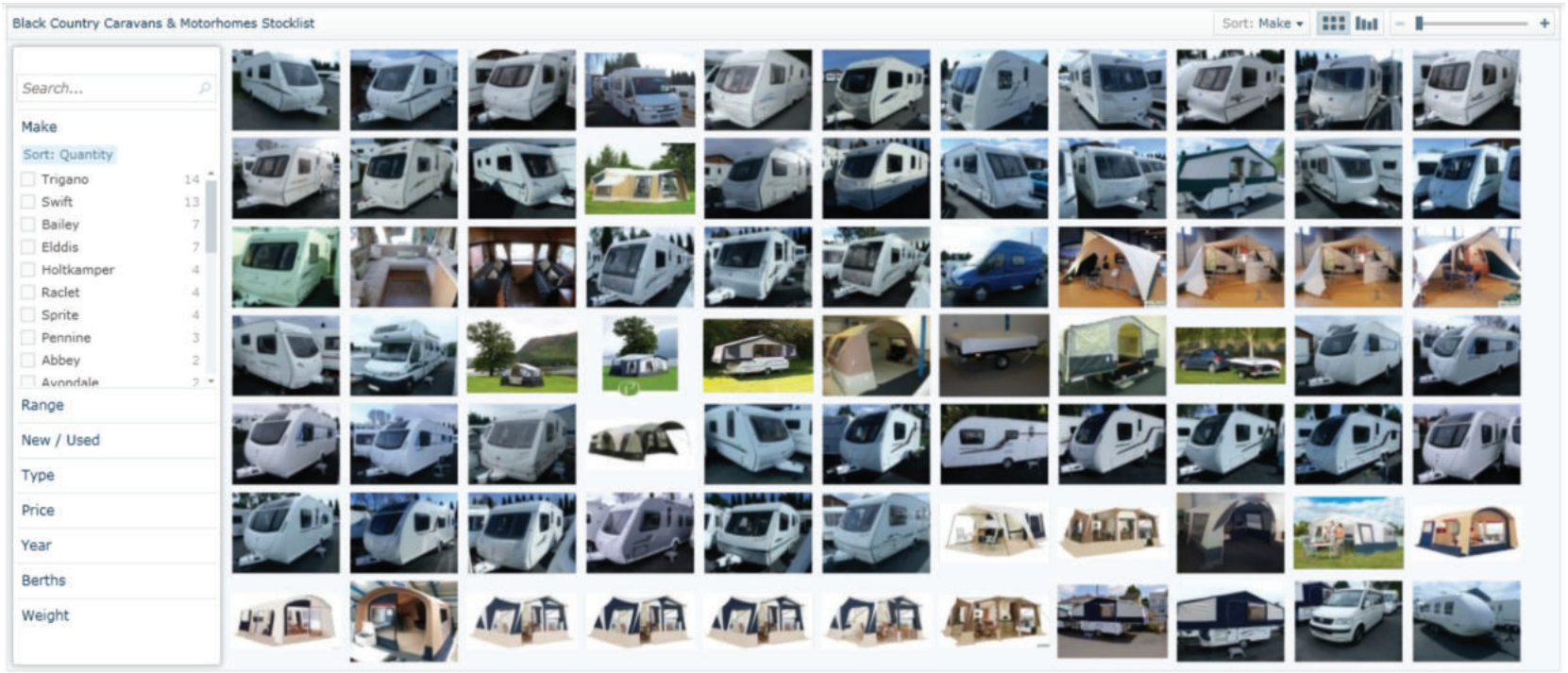
Visualization as Presentational format

As a method to represent data, visualizations become and essential area to be studied so as to conceptualized a solid framework for IIE.

Understanding Microsoft Pivot

Going forward, 2 products have been identified for a thorough understanding.

PivotViewer is a interaction model that helps interact with massive amounts of data on the web. It helps understand and skim down thousands of related items to specific data set through it’s interactive user interface.



Pivot Implementation [12]

The way it is visualized help users see patterns and trends over the huge data. The systems links up the image meta-data as attributes of the image through which it is categorised and organised into visualized filtered results.

Pivot is built on *Seadragon technology* and cleverly takes advantage of extreme zoom in & zoom out interaction model to browse and view the data. It provides and alternative to traditional standard web browsing through browsers.

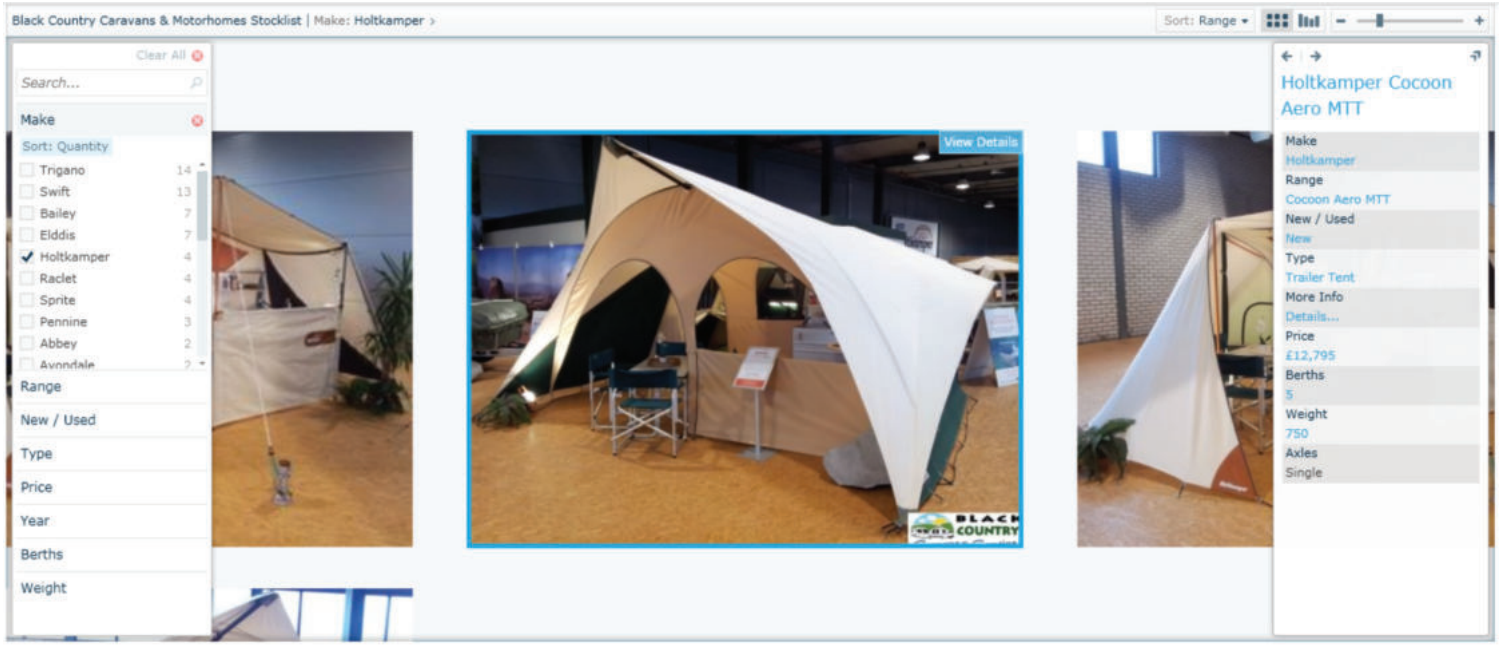
---

[12] Black Country Caravans - http://goo.gl/YQjEzw

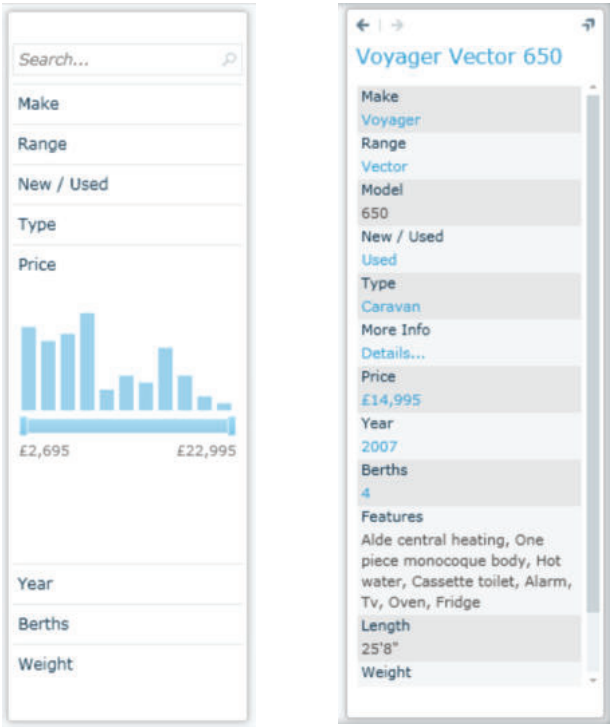
PivotViewer Interface:

Pivot takes data in the form of CXML which can be generated through its Excel tool. One can organize a data on a spreadsheet, assign meta data information over the rows along with an image. These meta data is then picked up and used to understand and organize the images into categories.

The User Interface consist of 3 main architectural elements: Zoomable Display Area, Filter Tool & Meta Data display.



Black Country Caravans - http://goo.gl/YQjEzw



PivotViewer UI Modules
(from http://goo.gl/YQjEzw)

Zoomable Display Area

The Zoomable data display area carries the visuals of the entire display. One can interact with the elements using the keyboard and the mouse. Data are constantly updated over the view when a filter parameter changes and fits in within the specified zoom area itself.

Filter tool

The filter tools carries a list of filtering parameters which are searchable and organised into chunked groups. As per the filters, the data displays gets updated over the display area subsequently.

Meta Data display

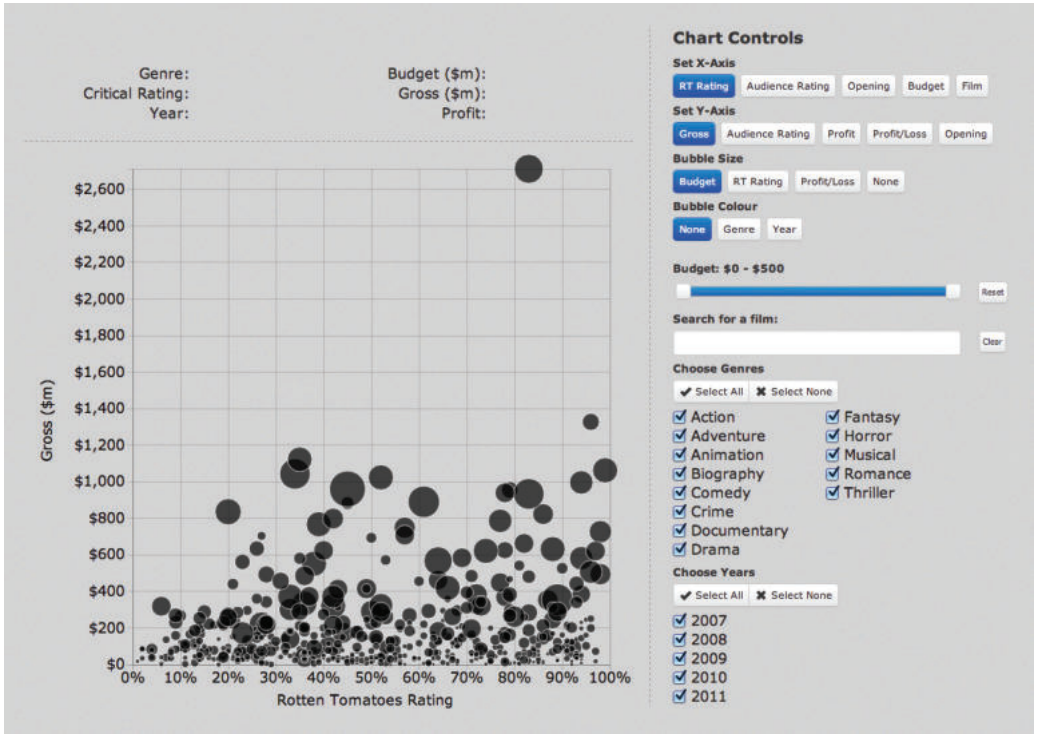
The Meta Data pane hides or displays as per the context. When a user clicks on a particular image, the data pertaining to that image shows up on the display pane.

Analysing Existing Visualizations

Apart from a generic UI model like Pivot, the second project/product deals with specific visualisations made on a much closer data area.

The Hollywood Data Explorer[13] (as shown in the diagram) from Information is beautiful has made an interesting visualization. It has several filters on the right which controls the data that is being visualized. Using only distribution charts and colors, it help understand the data in hand. For instance *outliers* like the *movie with the highest rating and maximum profit (Avatar)* jumps out and can easily be noticed which otherwise would have been tough.

Hollywood Data Explorer



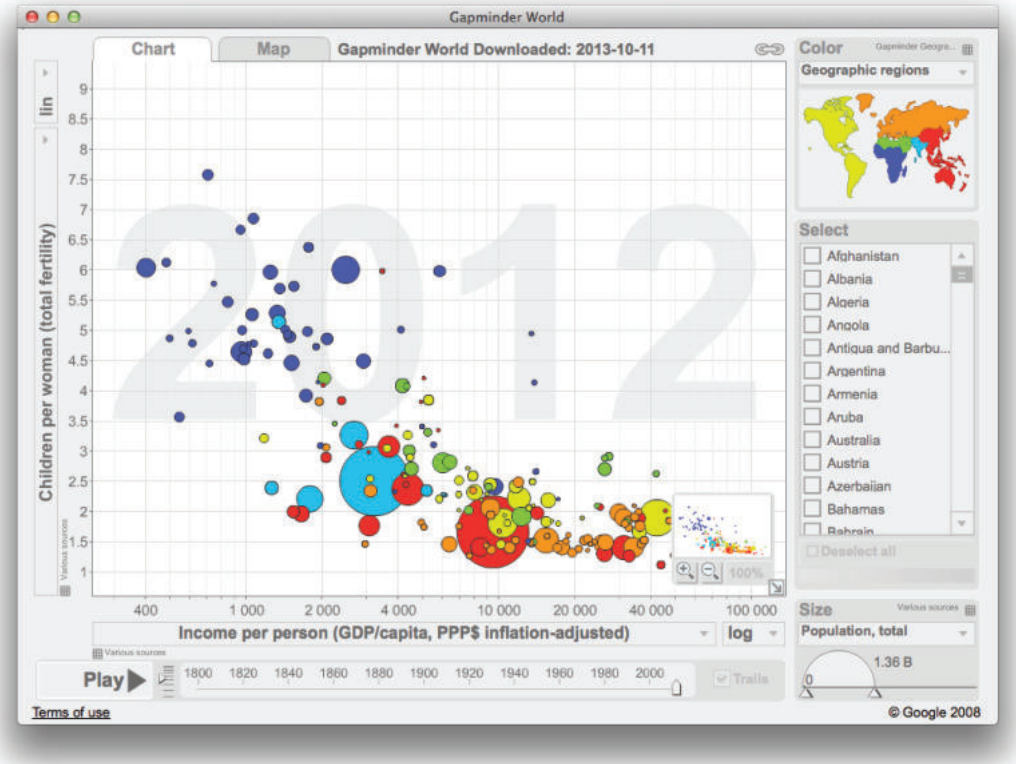
The Hollywood data explorer however merely visualize the data within the parameters assigned and does not show relationships between the entities. Categorizations are done using colors and not through Gestalt's grouping. Furthermore, the direct connection between the movie and the visual element representing the movie in the graphical expression is less prominent as merely circular shapes indicates each element. However, the size of these graphical units represents certain parameters which accordingly acts as the 3rd parameter in the x-y coordinate display.

13 Hollywood Data Explorer – http://indexity.net/vis/hw/

Trendalyzer:[14] Trendalyzer is another visualization tool from Gapminder Foundation in Sweden. This tool uses interactive bubble visualization charts to visualize a 5 variable data-set (x axis, y axis, bubble color as country, time and bubble size). Gapminder uses statistical data from the UN, National Statistical agencies and various NGOs[15].

Similar to the Hollywood Data Explorer, the Gapminder interaction is mainly govern by filters with an additional timeline animation. The bubble elements in the chart represents each country or region depending on the filter. The radius of the bubble indicates the value of the filtered data while its color represents the region in which the country belong. The x and y axis can be adjusted or changed according to the type if information a user wants to view through a series of fly-out menus. The graph can also be switched between a linear or a logarithmic one depending on the user needs.
The 'play' button allows users to view the progress or changes in the data over a period of roughly 200 years (1800 to 2000+). As the time progress, the location, size etcetera of the bubble in the chart shifts as per the data-set values providing a animated movement of the elements overtime.

Trendalyzer from Gapmind Foundation



[14] Gapminder World: http://goo.gl/UQVTe9

[15] The best stats you've ever seen: http://new.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen

Apart from the x-y coordinated chart, there is also a provision of a world-map view. Users can simply change the tab to display the world-map view and make sense out of the data bubble overlaid on top of its respective country in the map.
Similar to Microsoft PivotViewer, Gapminder's UI can be modularised into

Zoomable Display Area

This area has a tabulated interface where user can switch between a bubble-chart view or a map view. The Canonical adjust[16] operation is used as a zooming method where a zoom area is selected and focused which forces data outside the region to be pushed out of the screen while the whole scale on the axes re-adjust.

Filter tool

Gapminder has a huge data-set and as a result, appropriate filters need to be applied to view certain information out of them. Filters are provided through extensive flyout menu lists, a scrollable selectable checkbox and a draggable range selector.

Timeline Seeker Bar

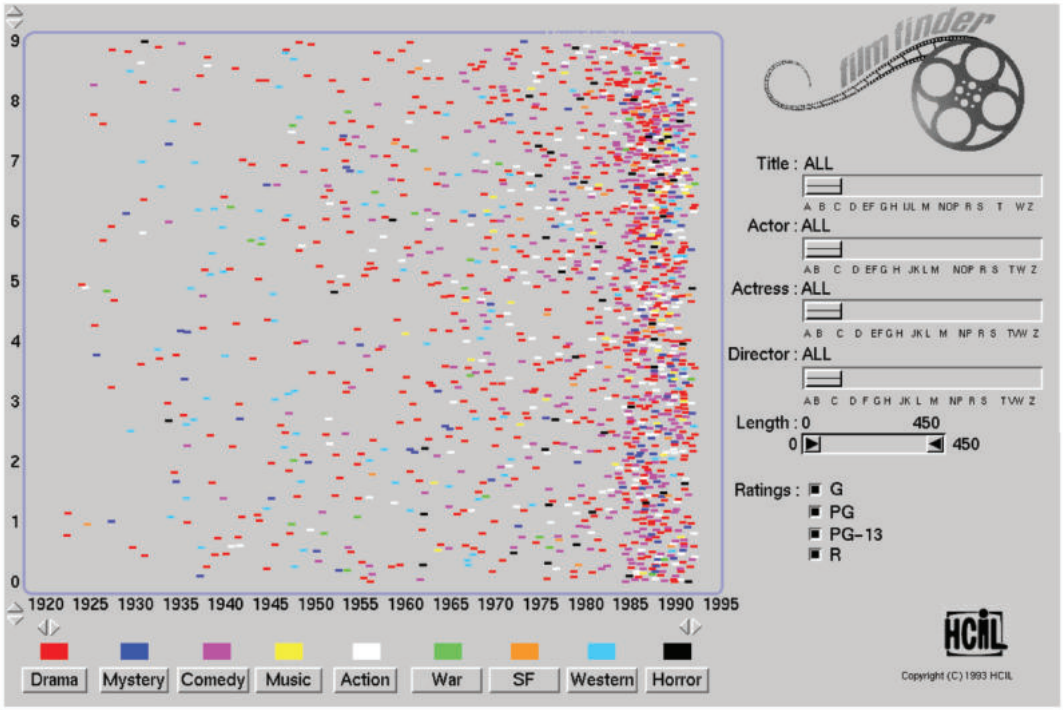
A timeline seeker bar enables users to go back and forth in time to see the readings of the exact year in question.

16 Rao, Ramana and Card, Stuart. The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus + Context Visualization for Tabular Information. Proc. of CHI 1994, ACM, New York, pp 318-322.

FilmFinder by *Christopher Ahlberg and Ben Shneiderman*[17] uses a star field display instead of a bubble-chart to provide a Visual Information Seeking System.

The FilmFinder starts with the principle of direct manipulation where rapid incremental filtering reveals progressively refined results which becomes a new search parameter. The progressive refinement through a reversible direct manipulative interactions yields results immediately and uses the concept of tight coupling to prevent null result-sets.

FilmFinder



Each rectangular block represents a movie in the zoomable area. The Canonical Adjust operation for zoom is used here also and as it is zoomed, the size of each rectangle changes according to the zoom level. A zoom range is set so as to prevent from an empty zoom area or too congested visualization. When fewer than 25 movies are visible on the screen, the titles of the movies are displayed automatically.

17 Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays, Christopher Ahlberg and Ben Shneiderman

Technology & Visualization Frameworks

During the process of the development of an IIE prototype, some of the technological frameworks or libraries are considered. The project does not intend to inhibit itself from any particular framework or library.

Over several frameworks, the following are some of the options IIE was attempted to be developed upon.

- Google Visualization APIs
- PhiloGL
- D3
- jQuery Plugins (jqPlot, jqWidgets, TufteGraphs)
- Raphael
- Processing

The existing prototype however is developed with a blend of Vanilla HTML5, CSS3 and JavaScript coupled with the D3 visualization library. The choice is not governed by any technological specificity of these libraries but rather is made entirely on the convenience of developing the prototype as well as prior knowledge of the languages.

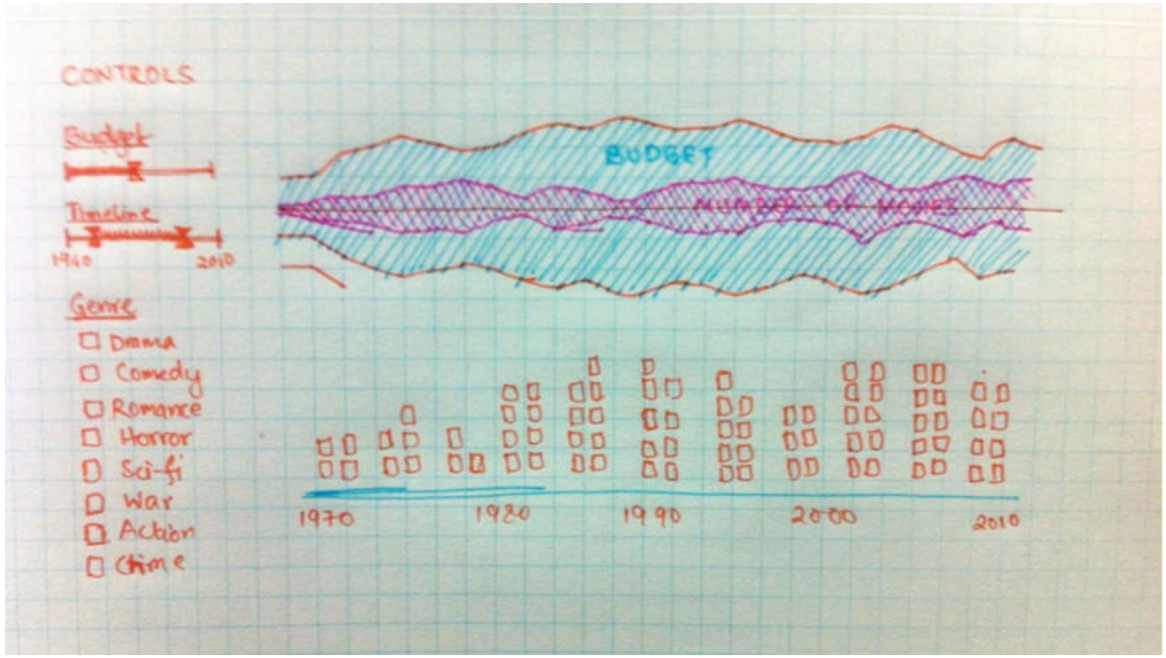
Concepts:

Concept 1:

Concept 1 has on 3 interaction panels:

- Tools (left)
- Visualization (top)
- Timeline (bottom)

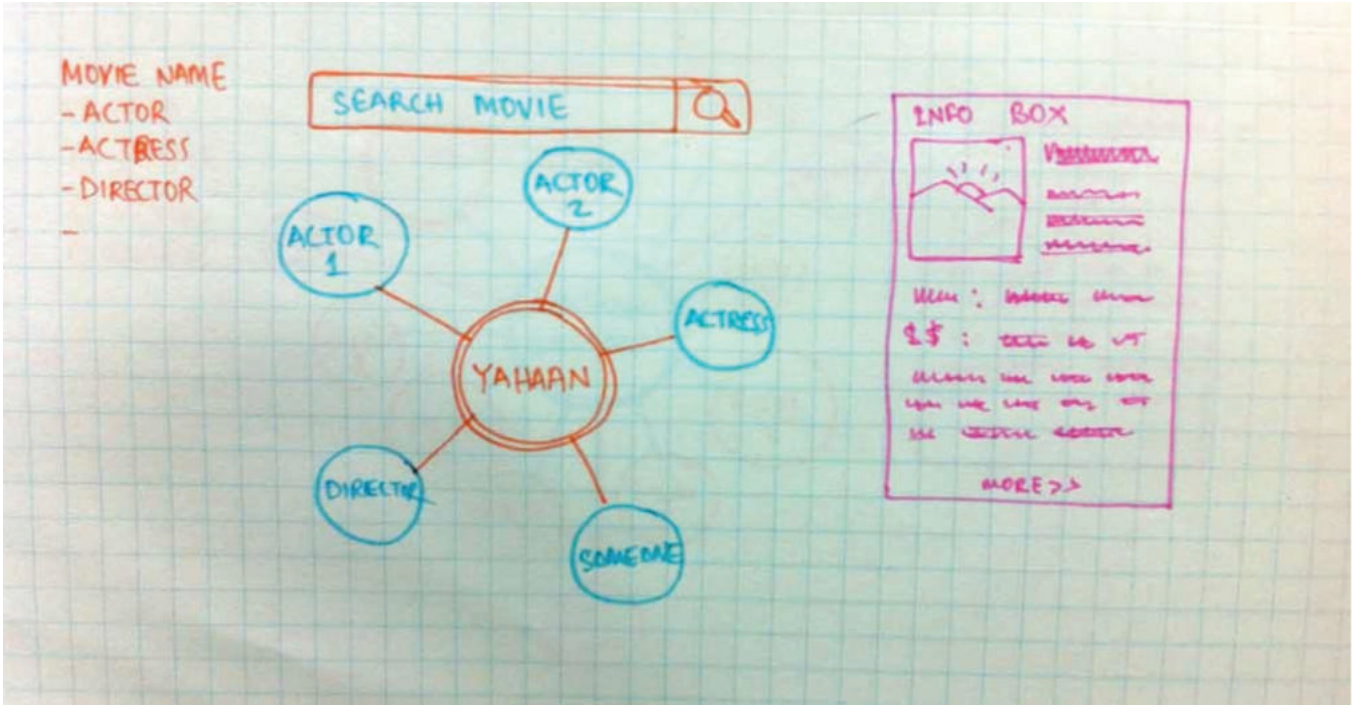
The tools provide basic interaction level filters according to which the visualization on top section is generated. The visualization may change from mode to mode. The effective visualization is generated from the timeline as shown in the bottom of the screen. The timeline is arranged in the chronological order.



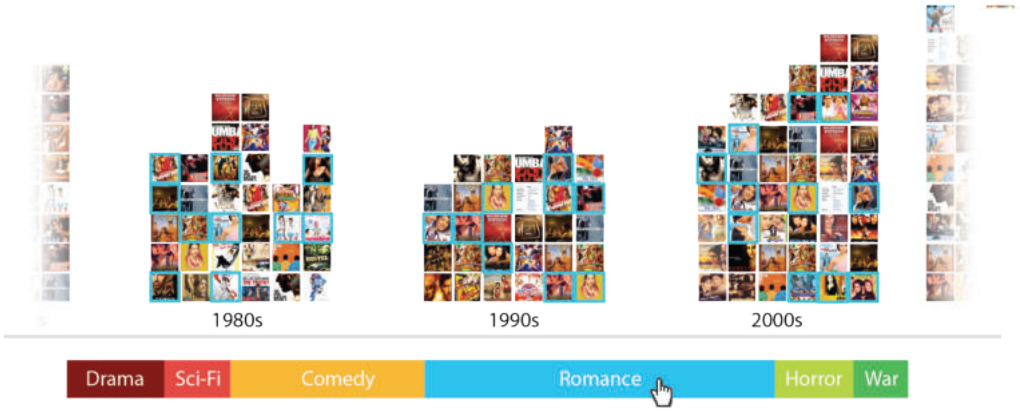
Concept 1 has the added advantage of direct manipulative capabilities. With the use of sliders and drop-down menus, choosing specific data or ranges of data values are easier and accurate. The use of traditional Interaction model in the sense that most of the UI elements implemented are conventional elements, users which have only basic knowledge about how computer or tablet PC works should understand should find it easy to understand IIE.

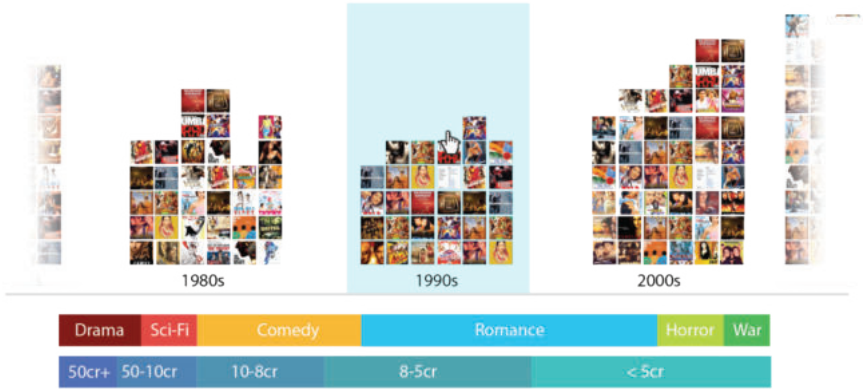
Concept 2

The second concept is based on data relationships. Each entity of the data-set is considered a node in the mess-network. The relation between each node is traced and connected as such when larger sets of data are considered, several interconnection between nodes are revealed as hidden insights.



A small information box also displays basic meta data about the active node.



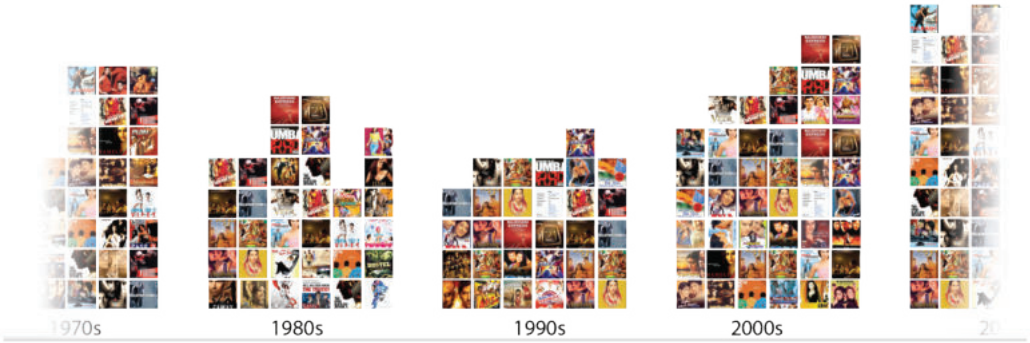


Concept 3

The third concept is again based on a timeline based systematic arrangement of the movie thumbnails. Here, the arrangement itself acts as an interactive element of the framework rather than a separate tools section as in the previous concept.

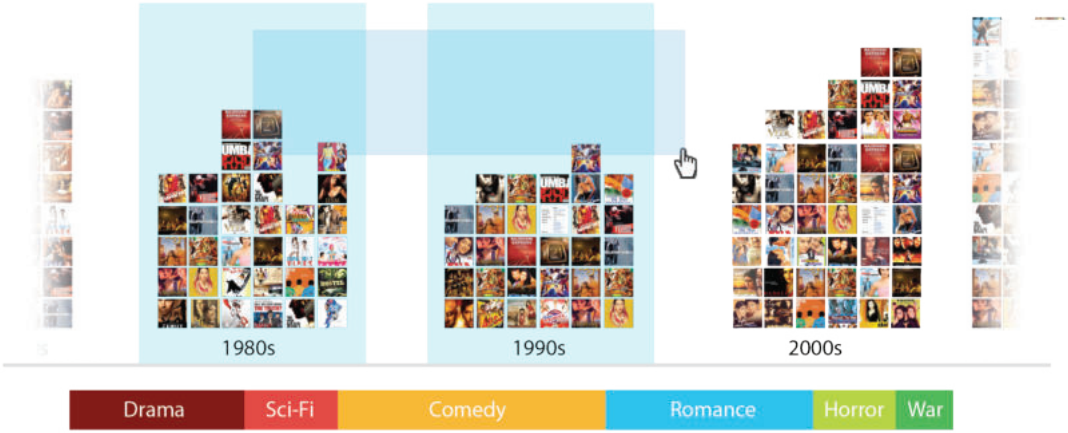
Primary interaction includes zoom-in or zoom-out the timeline, click on the thumbnail to find details of a movie, select groups of stacked thumbnails to visualise the distribution of selected parameters like genre and budget distributions etc.

Concept 3 Iterations



Iteration 1

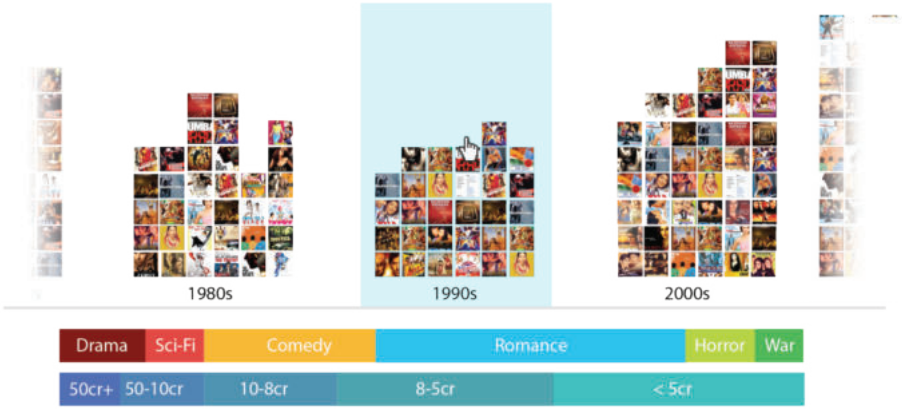
The movie thumbnails are arranged into a timeline format where horizontal scroll reveals more data on either side.



Iteration 2

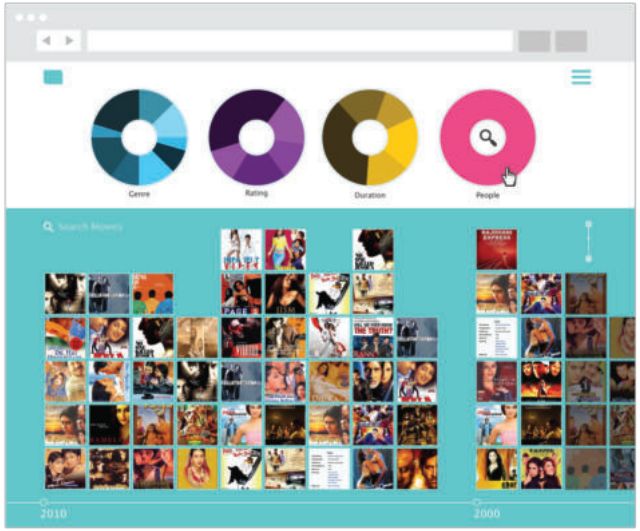
A contextual colored bar is introduced to visualize selected or currently active movie parameters such as genre & budget.

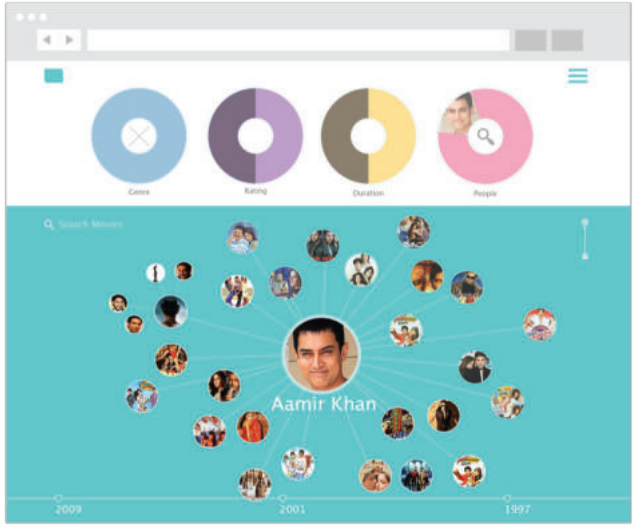
Concept 3 Iterations



Iteration 3

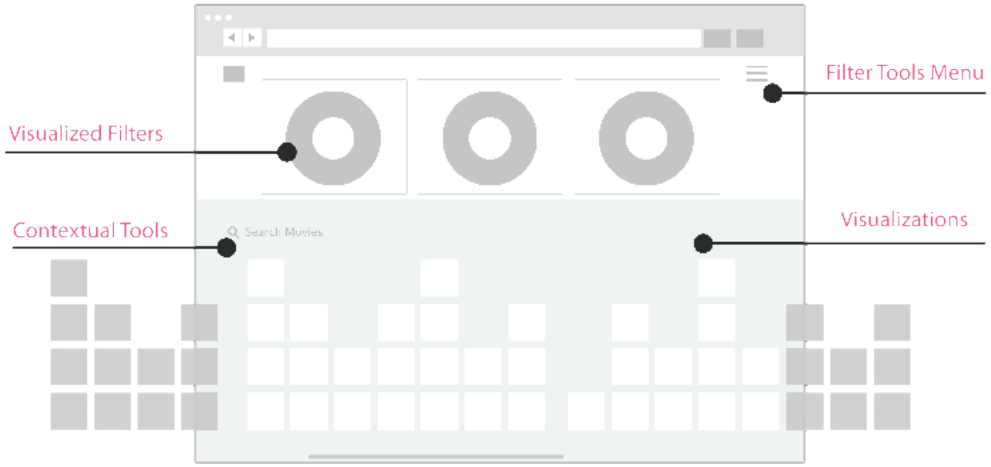
Additional bars are introduced that are replaceable as per user requirements. Currently it shows 2 bars, one genre & one for budget. More bars can be added or either one of the 2 bars can be replaced with some other parameters.

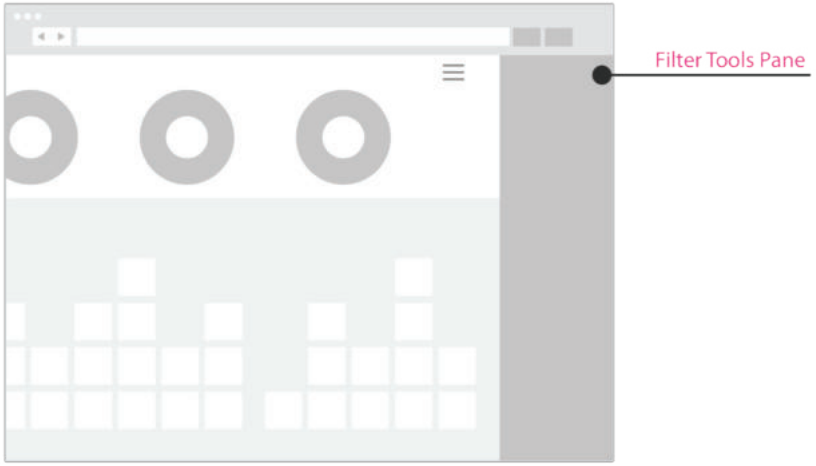




Iteration 4

Final concept UI is discussed in detailed under ‘Final Concept’ header.



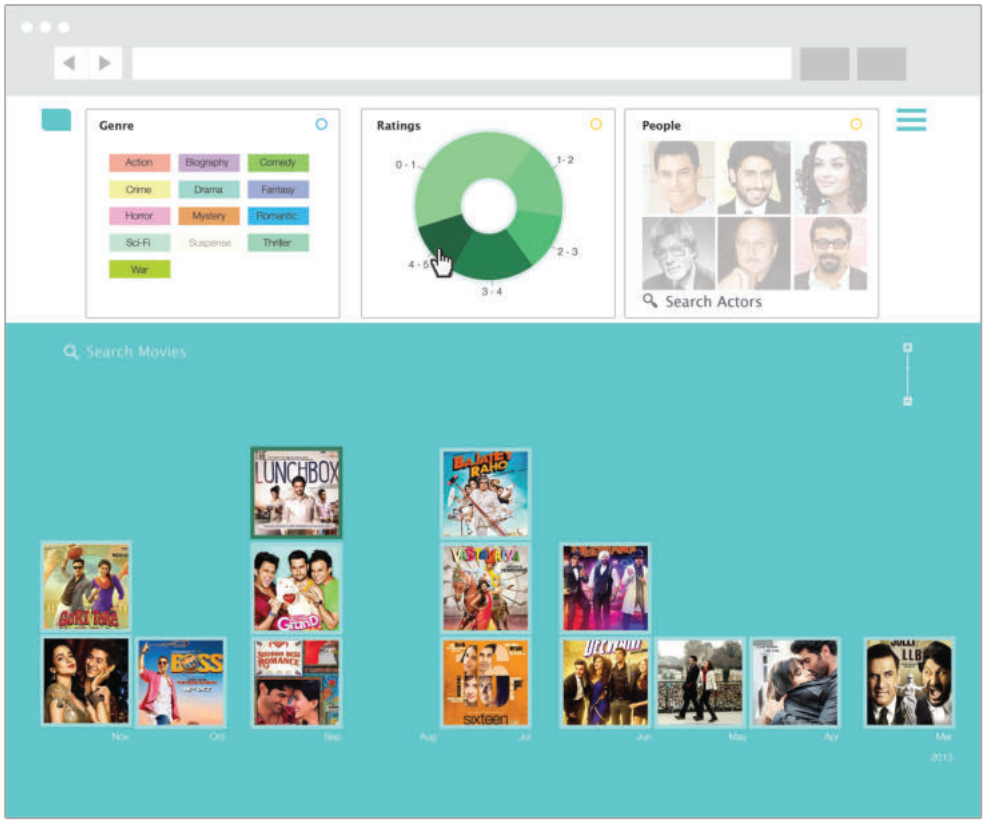


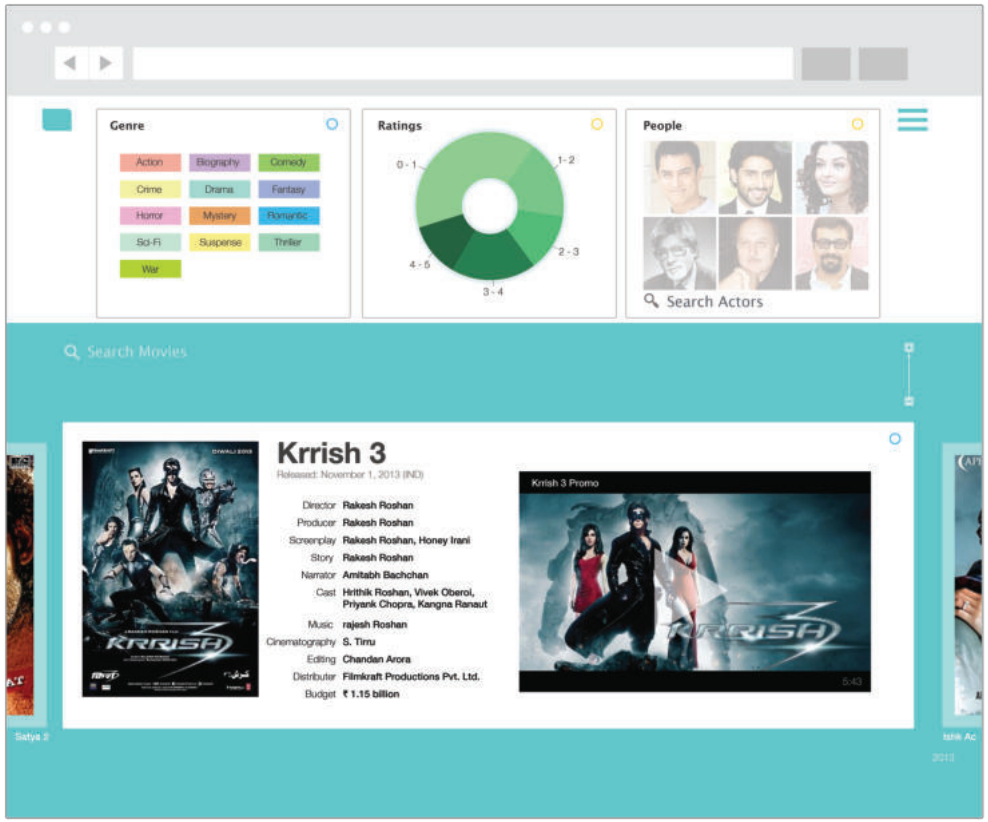
Final Concept

Implementing feedbacks and suggestions over the course of the project, the final concept is laid down. It roughly integrates concept 2 and concept 3 along with its own interface and interaction model.

The IIE's main interface consists of 4 main modules. The modules interconnects to each other tightly coupling to serve an integrated framework to provide the intended interaction model and give a smooth user experience.

1) Visualization Filters
2) Visualizations
3) Filter Tools pane
4) Cards


Genre
Action
Biography
Comedy
Crime
Drama
Fantasy
Horror
Mystery
Romantic
Sci-Fi
Suspense
Thriller
War
Ratings
0 - 1
1 - 2
2 - 3
3 - 4
4 - 5
People
Search Actors
Search Movies
LUNCHBOX
sixteen


Genre
Action
Biography
Comedy
Crime
Drama
Fantasy
Horror
Mystery
Romantic
Sci-Fi
Suspense
Thriller
War
Ratings
0 - 1
1 - 2
2 - 3
3 - 4
4 - 5
People
Search Actors
Search Movies
Krrish 3
Released: November 1, 2013 (IND)
Director Rakesh Roshan
Producer Rakesh Roshan
Screenplay Rakesh Roshan, Honey Irani
Story Rakesh Roshan
Narrator Amitabh Bachchan
Cast Hrithik Roshan, Vivek Oberoi, Priyank Chopra, Kangna Ranaut
Music rajesh Roshan
Cinematography S. Tiru
Editing Chandan Arora
Distributer Filmkraft Productions Pvt. Ltd.
Budget ₹ 1.15 billion
Krrish 3 Promo
5:43

Visualization Filters

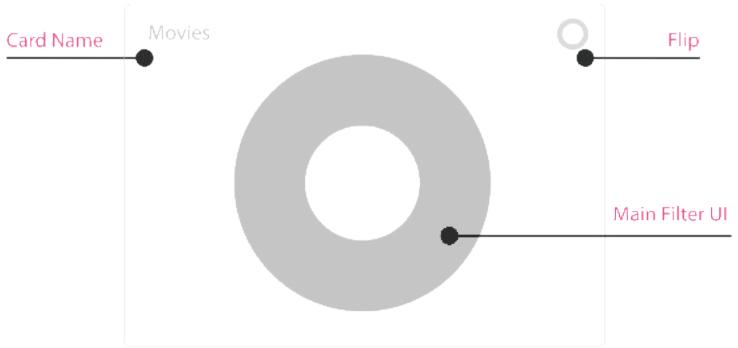
The visualization filters pane carries clickable UI elements that acts as filters to the data. These visualized charts carries the contextual representation of the data currently visible in the active region of the visualizations' pane below

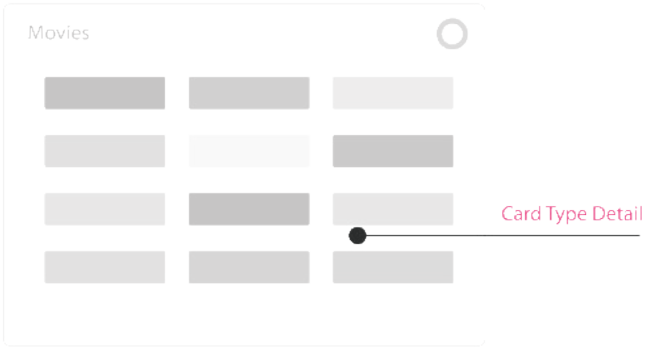
Visualization

The bottom panel where thumbnails are visible (in the images on the left) carries the main visualization. Each visualization may differ per data-set or mode of visualization. Each graphical element on this pane also are interactive and respond to click and scroll events.
The thumbnails are organized in a timeline manner where scroll goes horizontally from left-to-right in latest-to-oldest arrangement.
As one zooms-in or out, the thumbnail size increases or decreases and re-arranges itself within the specified pane area. The zoom level correspond to decade view, yearly view and monthly view of the data.

Filter Tools Pane

The filter tools pane carries additional filter. These filters can be dragged & dropped or interchanged with the visualization filters. Extra filters can be applied to an existing state and it may change the structure in totality.

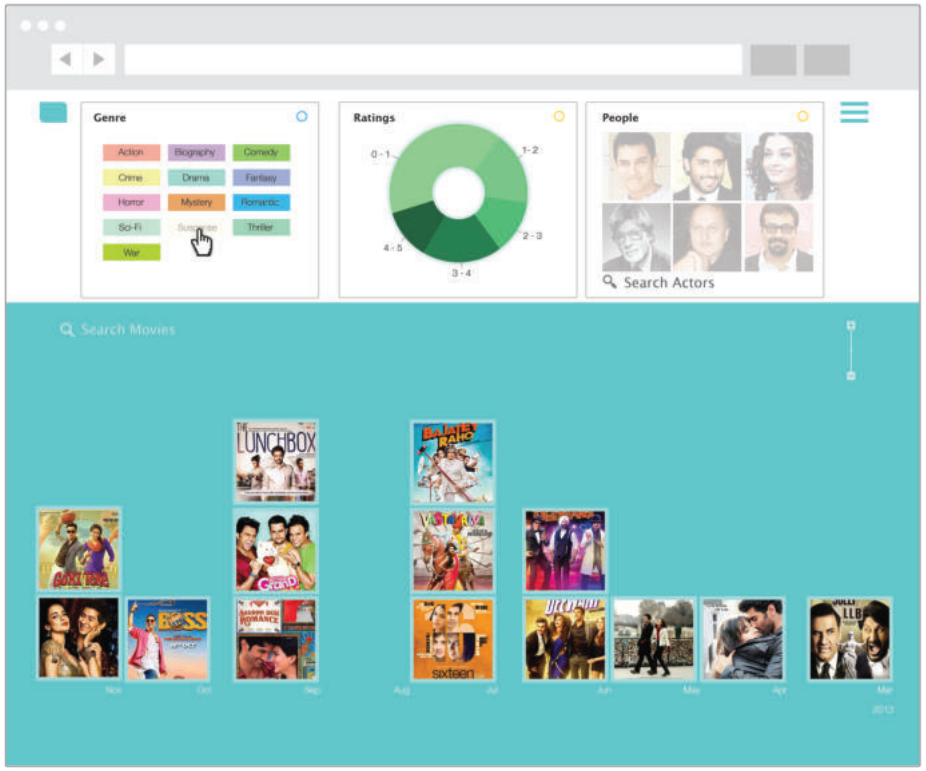


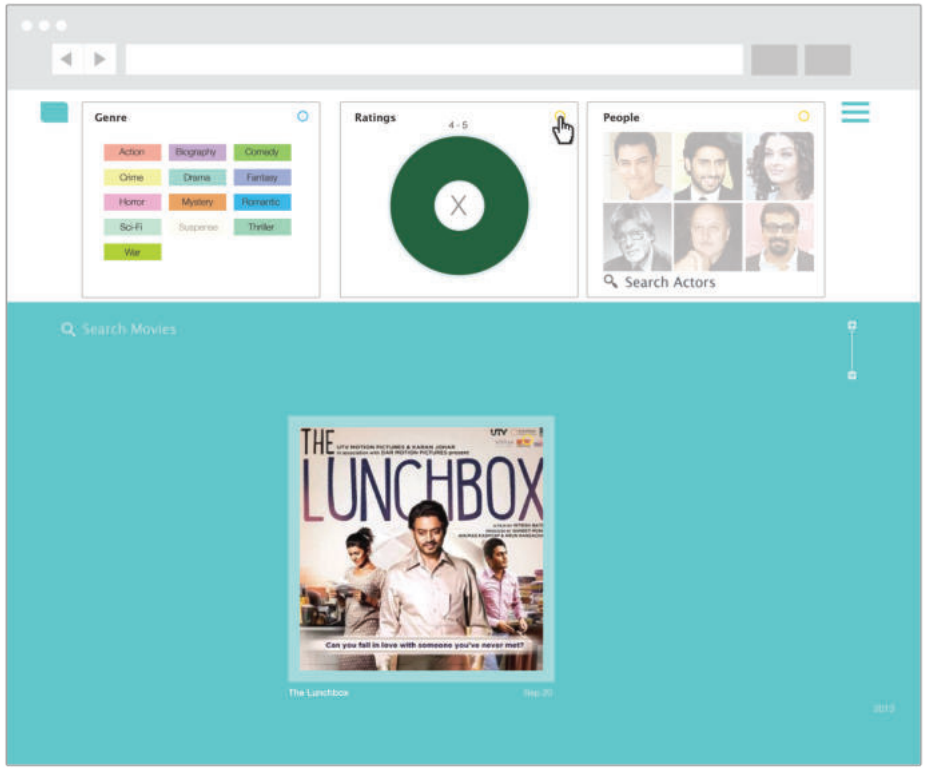


Cards

It is a flippable two-sided panel where a crude version of the information or interface is provided in the front panel. When the card is flipped, more contextualized information or detailed interaction elements can be accessed. For example, while accessing information about a movie, the front panel of the card will be shown with a thumbnail of the movie poster. When the card is flipped, detailed description of the movie will be displayed.

Cards can be considered modularised building block to the interface. The front panel may carry basic interface to control a coarse filter like a date range while the the back side may provide a more directly manipulative interface of the same filter type.





The cards entitled ‘Genre’, ‘Ratings’, ‘People’ at the top acts as filters to screen data shown in the visualization below. As such, in the screenshot on the left, a user can hover and click on any filter item, for instance on ‘suspense’ genre in this case. As one clicks on the filter, it gets applied to the visualization below. If a donut chart is the filter, the whole donut chart gets converted into the filter type applied as a filter.

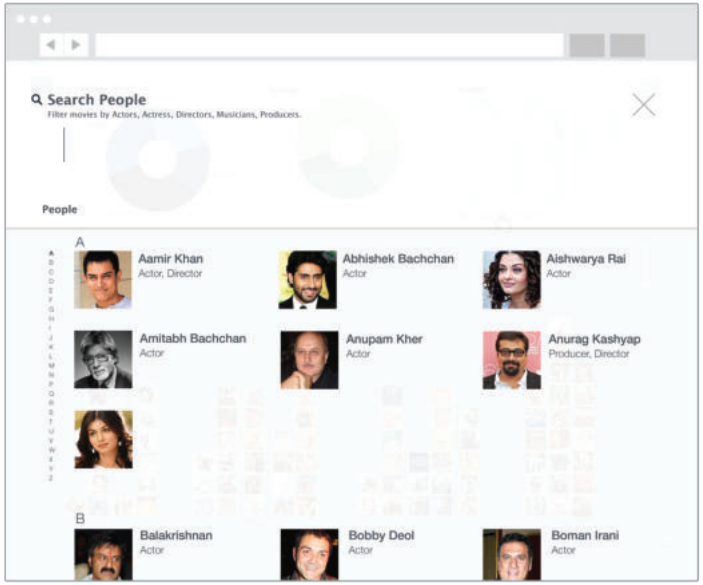
As a result, only movies which belong in that category will load in the area below.

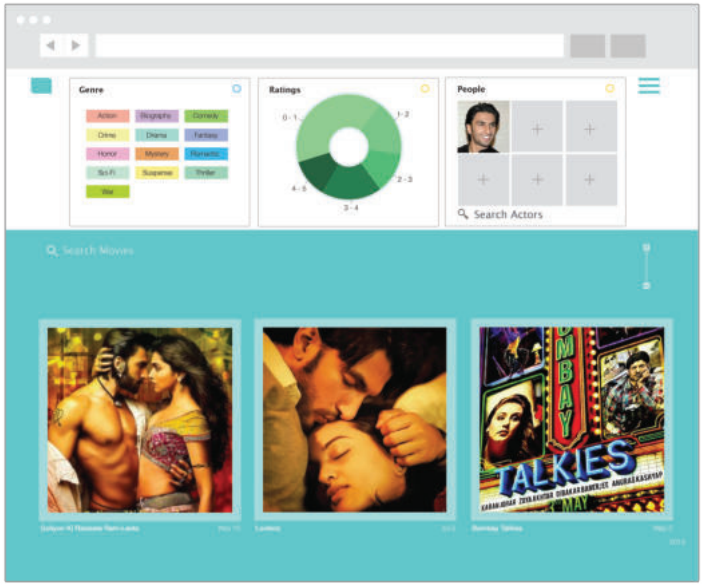
As such, all the mutually exclusive filters can be applied or remove and the data loaded below can be drilled down to the most basic. As such, this acts as a progressive refinement or broadening of the search criteria.

Taking a scenario 1:
A user wants to find a movie released in a certain given time.

The user can simply click on the zoom button and use the horizontal scroll to get to display the appropriate time period in the visualization area. Once the visualization is zoomed properly enough and time period visible on the screen is the specified time, the movie(s) which satisfy the filter criteria will be visible on the screen.

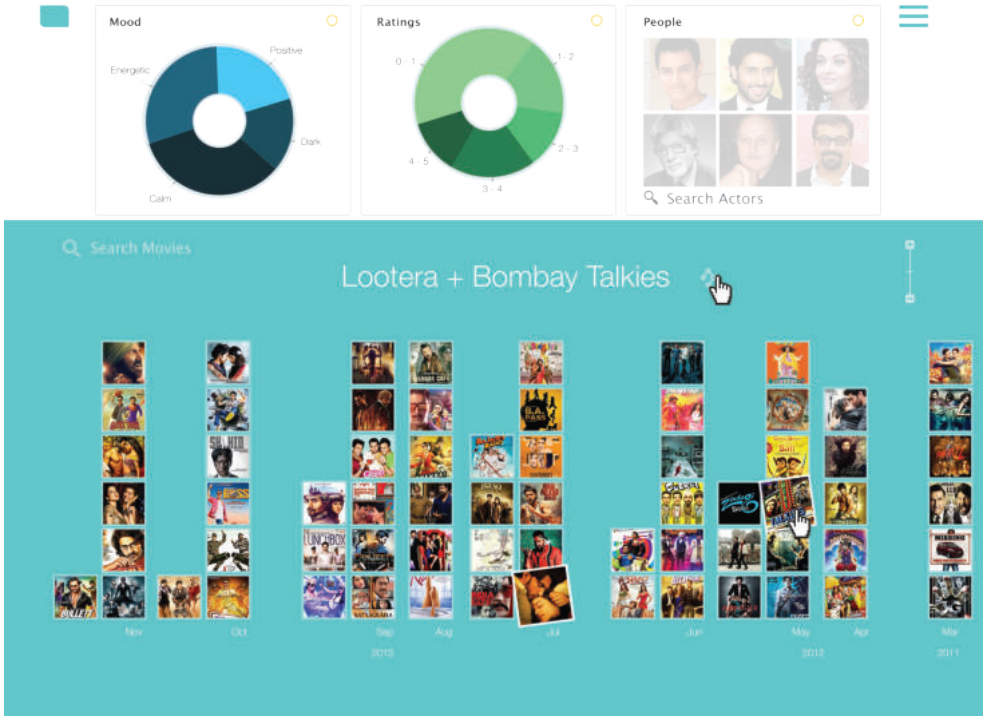
This finally empowers the user with an exploratory tool.

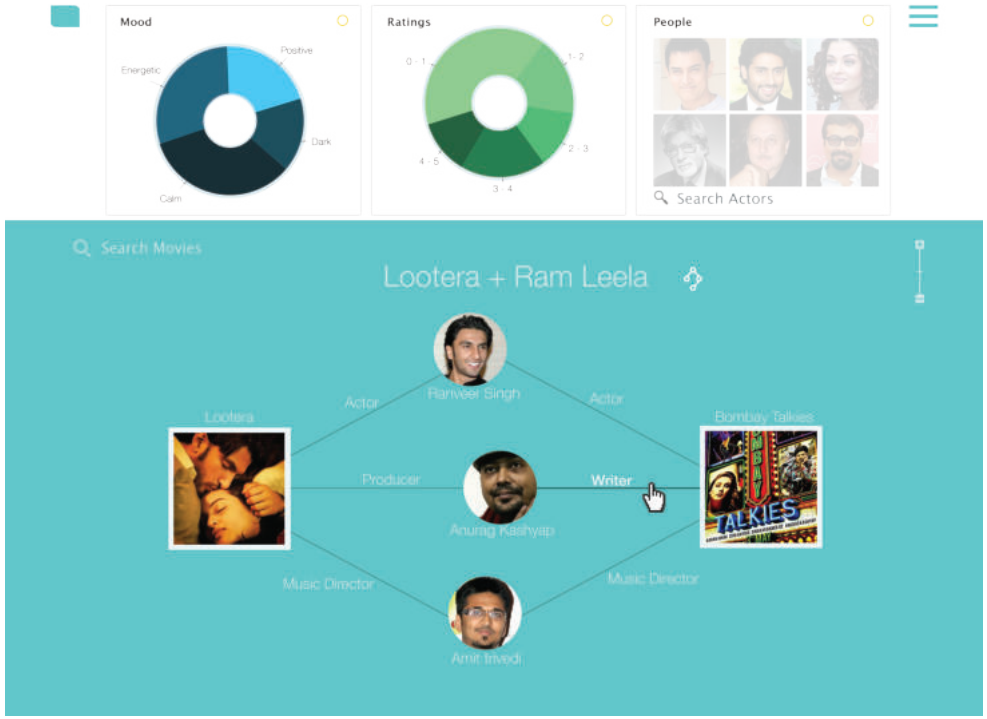




Taking a second scenario:
Apart from the basic filter capabilities as in scenario 1, If a user wants to watch the movie which casts a specific actor, or is directed or produced by a specific person, a user can use the people filter (which is search).

When a user clicks on search, a new panel will appear with a search interface. From the search results, the item can be picked and as obvious results, data will be shown which satisfies the filtere criteria including the search results.





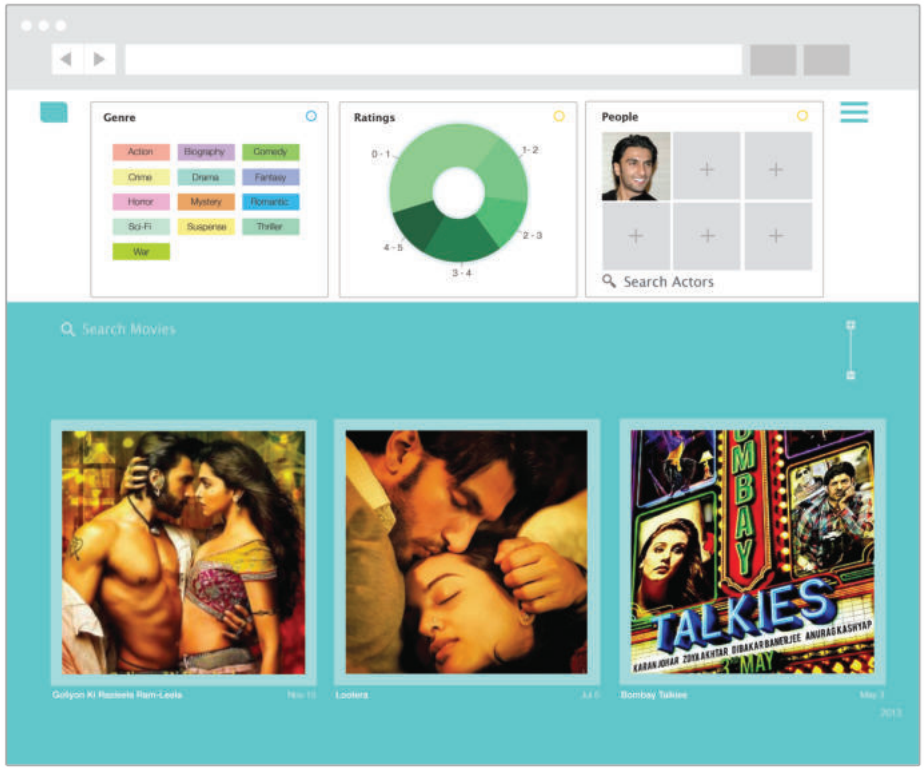
Movie-Movie relationship mess-chart - the relationship nodes displays how the movies are connected.

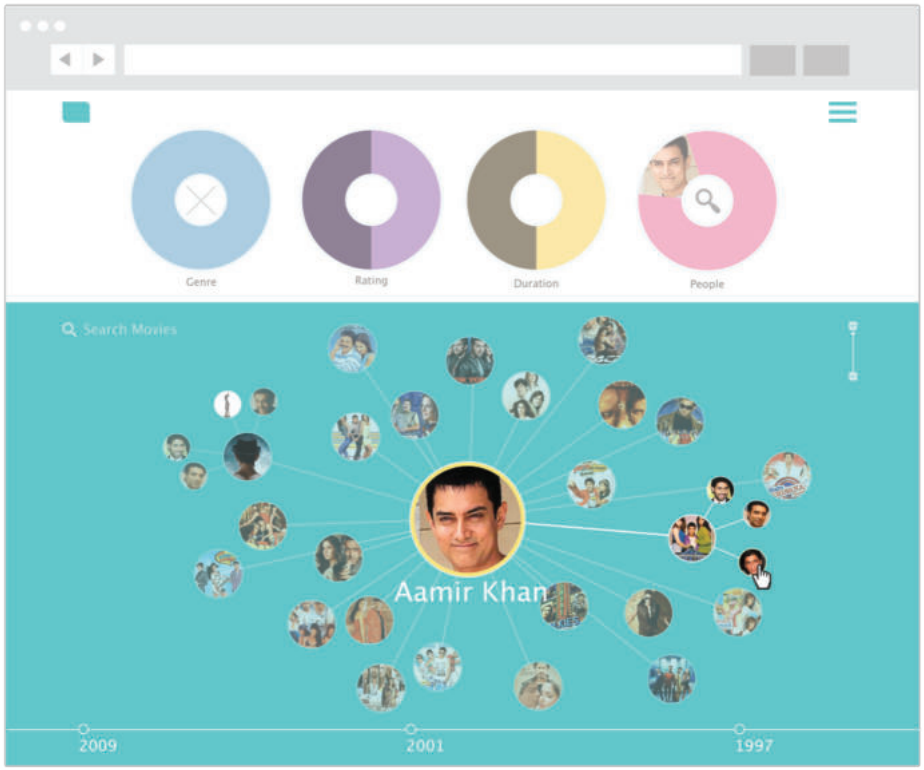
On an alternate view, the relationship between data elements are represented as nodes.

The mocks (on the previous page) displays the data node (movie) relationships which brings up the underlying connection between data in the information architecture.

The Movie-Movie relationship chart shows how the movies are connected through people in the film industry and their role to the movie.

Initially a people-people relationship chart was implemented to indicate the connections between individuals in the industry. However, the idea was dropped as the visual turns out to be just an alternative representation of applying the people filters with multiple individuals in the filter.





Comparision between newer mocks and older mocks - both mocks display movies connected to a particular individual. The second however is discarded as it becomes only an alternative representation.

The right mock was considered in the earlier version to be implemented. It shows the movies connected to the individual in the center - in this case, Aamir Khan. However, it is again dropped considering that the new mock on the left serves it purpose. Once the people filter on the top is applied for the specific individuals intended, the results that throws up in the visualization panel is the exact same information - the only difference is how the data is visualized. The visual on the right does not in any way enhance the data available to the project.

The visual on the right would probably have been useful to indicate the importance of the movie to the indivual through the distance of the nodes from the central node. The distance can also represent time-lenght of the release of the movie from the present. However, since it is scattered over 360 degrees, the comparision between two elements are not represented best in this form.

Evaluation Criteria

IIE aims to help users make better decisions fast while choosing a movie to watch. Often times, traditional approach to database queries results in textual base results possibly along with a thumbnail. If the users are unsatisfied, new queries are made which means restarting the whole search process. IIE however uses filters that progressively refine or broaden the search parameters.

To establish this, the product will be evaluated based on 3 simple tasks assigned to the users. Time taken to achieve each task as well as number of steps involved will be measured separately for the traditional method of searching movies as well as while using IIE.

In traditional method, users will asked to perform the task whichever way he or she is comfortable with. The user is free to use any tool or method he or she wishes - ranging from Google searches to movie database sites or so on. For IIE, the user will be asked to perform a similar task using IIE.

The three tasks are as follows:

Find a movie released in a specific time?
In this task, a specific date (or a time period) will be given to the user and the user will then be asked to find a movie released on that

particular date (or period). The steps involved as well as total time taken to complete the task will analyzed to decide the effectiveness of the IIE framework.

Find a specific movie with a certain rating, which belongs to a certain genre or other attributes, from a specific time period?
A certain criteria involving the rating of a move, the genre as well as other attributes like mood or budget etc. and a time period will be given to the user and the user will be asked to find the movies which meets the criteria.

Find the number of movies of a particular actor in a specific time?
The user will be asked to find certain movies of a specified individual (an actor, actress, director, musician etc).

While designing IIE, the following ideologies have also be considered.

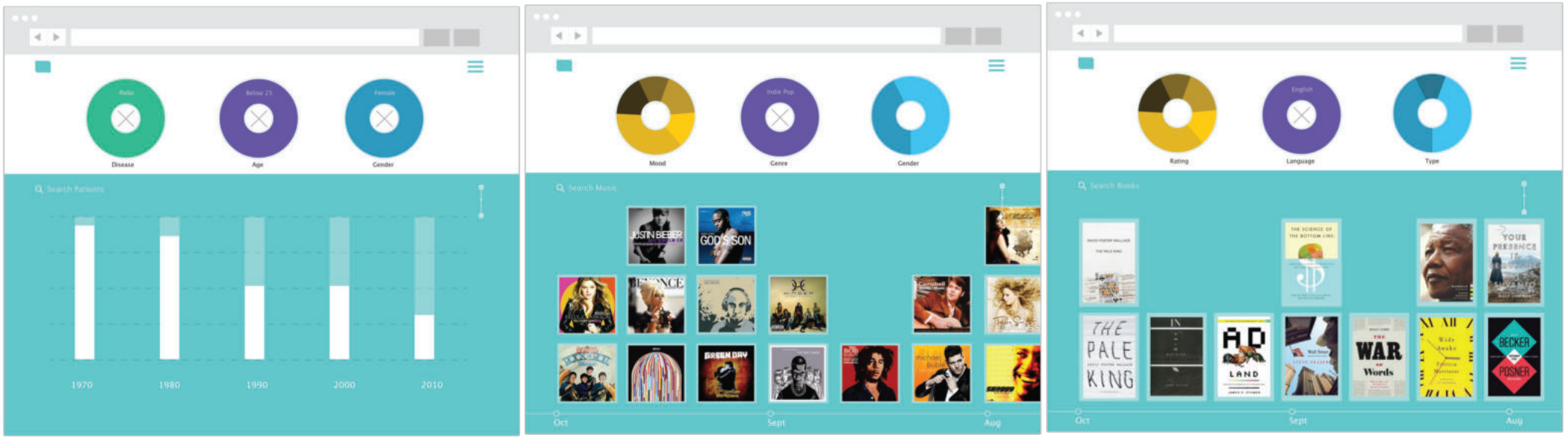
Scalability
The Information Architecture of the system is expected to be scalable as one of the primary ideology being IIE is the generalized data

exploration that inherently requires the handling of multiple data-set.

Provision of a holistic overview
The system is expected to provide a holistic overview of the whole data and help users see the overall picture. The criteria doesn't involve giving manipulative UI to handle detailed information visualization.

Explore data and understand the connections
The framework should support an exploratory experience which 'might' lead to 'serendipitous' discovery of likely 'unexpected' but 'interesting insights' which otherwise would have been hard to notice.

Access to data
The system should enable users to access data acting as a data explorer.



Scope & Conclusion

The major take on the essence of creating IIE is to look for a generalized way of exploring differnt variety of data-sets.

The basic framework may likely support exploration of other similar data than movies. For instance music album data, books data and even medical data can be implemented in IIE.

As such, the filter pane will adopt different user-created filters. New user generated filters thus can be added or removed. This will further enhance ecosystem as a standalone product.

Reference

1 Microsot Pivot TED talk by Gary Flake
http://new.ted.com/talks/gary_flake_is_pivot_a_turning_point_for_web_exploration

2 Visualizing Data by Ben Fry

3 National Film Archives website | http://www.nfaipune.gov.in

4 Information is beautiful Hollywood data | http://goo.gl/2XD8hO

5 The Visual Mescellaneum by David McCandless

6 Feltron Report

7 Interactive Analysis of Big Data

8 Stat! – Interactive Analytics Environment

9 A Common Misconception About Designing With Data
http://www.smashingmagazine.com/2013/07/29/common-misconception-designing-data/

10 File Explorer | http://en.wikipedia.org/wiki/File_Explorer

11 Information Visualization | http://en.wikipedia.org/wiki/Information_visualization

12 Black Country Caravans | http://goo.gl/YQjEzw

13 Hollywood Data Explorer | http://indexity.net/vis/hw/

14 Gapminder World | http://goo.gl/UQVTe9

15 The best stats you've ever seen
http://new.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen

16 Rao, Ramana and Card, Stuart. The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus + Context Visualization for Tabular Information. Proc. of CHI 1994, ACM, New York, pp 318-322.

17 Visual Information Seeking
Tight Coupling of Dynamic Query Filters with Starfield Displays, Christopher Ahlberg and Ben Shneiderman