

ARTS1422 Data Visualization

Lecture 14

AI4VIS: Empowering Visualizations with AI

Quan Li
Spring 2024
2024.04.23

Data Visualization



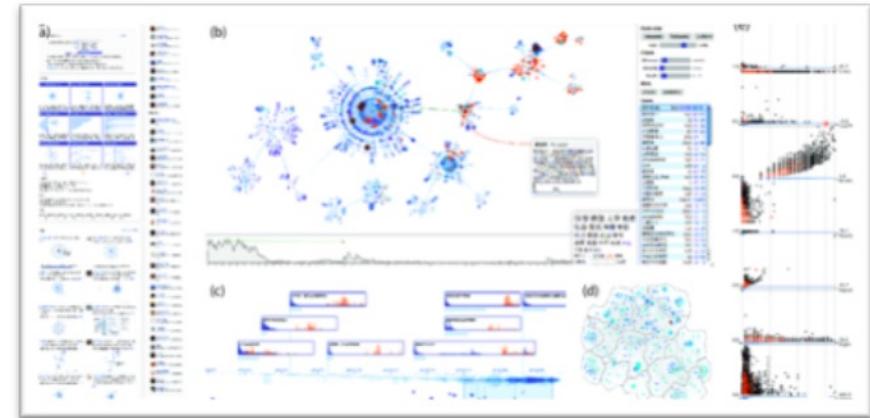
Data visualization is the graphical representations of data



Chart and Diagram



Infographics



Visual Analytic System

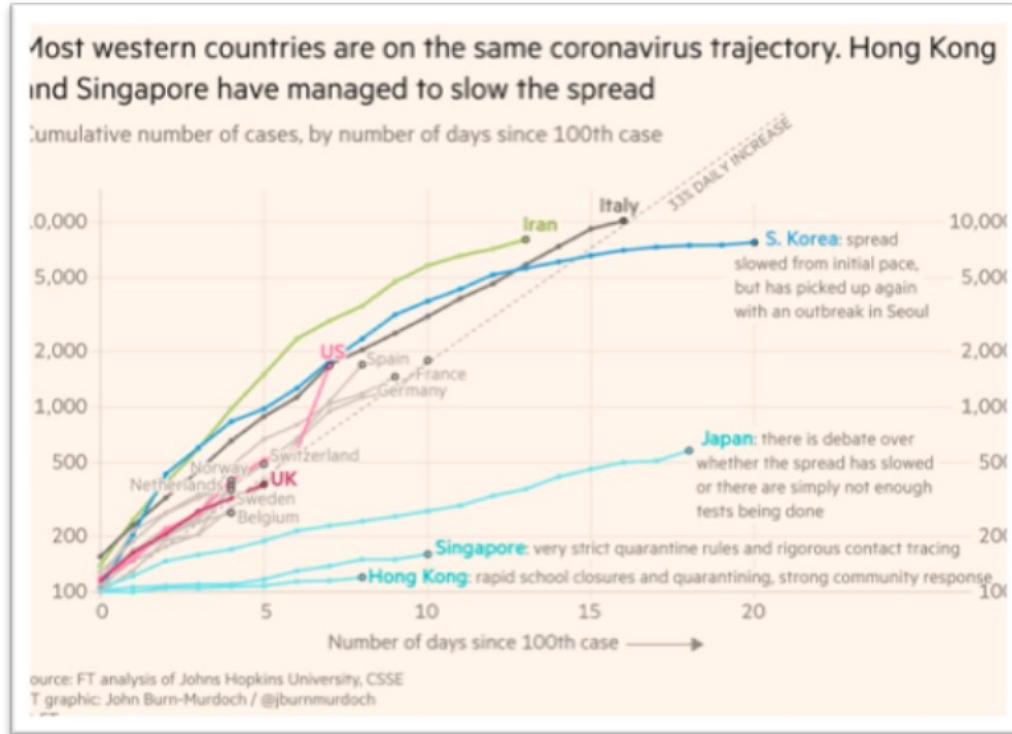
Basic

Advanced

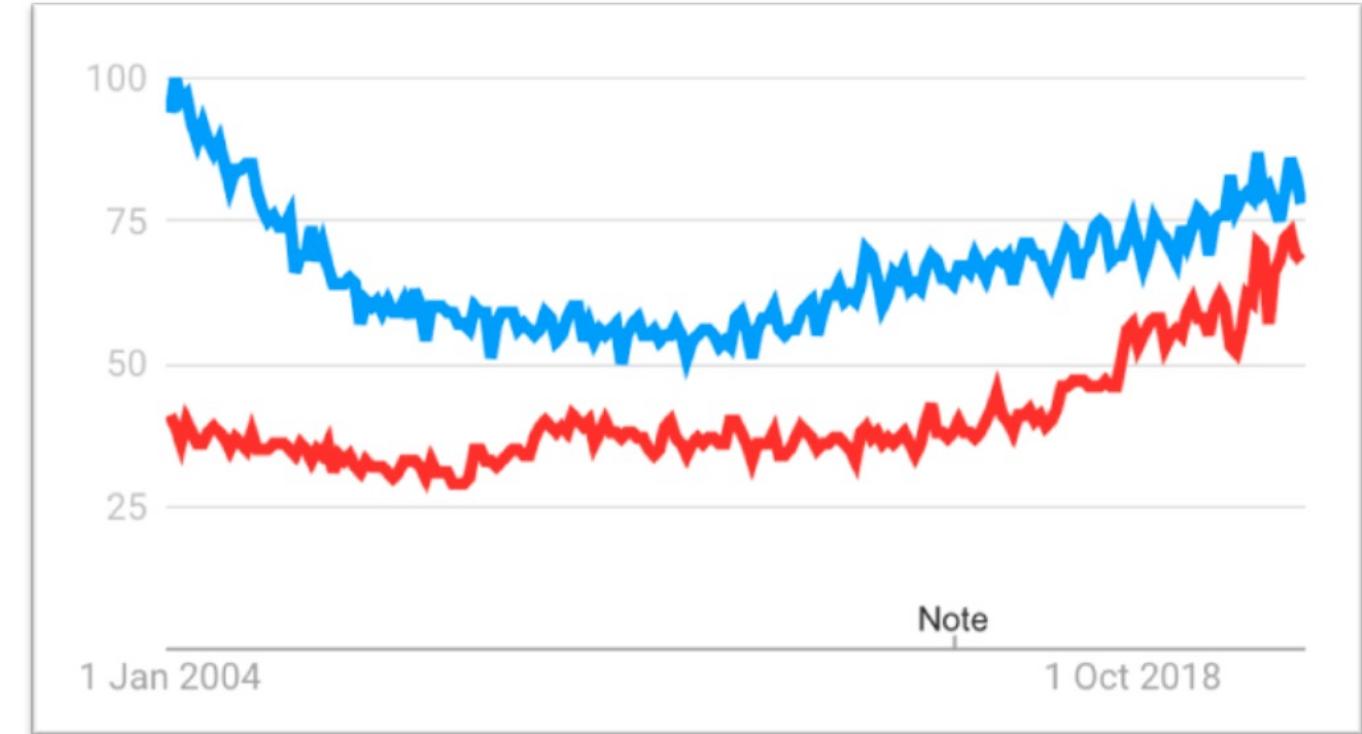


Why Visualize?

- Visualizations are easy to see and arguably one of the most easiest way for the masses to assess data



By far the most visited page in NY Times



Google Search Trend Globally: **Data** v.s. **Chart**



From Data Visualization to Visualization Data

- Visualizations are increasingly created, and therefore become a new type of data

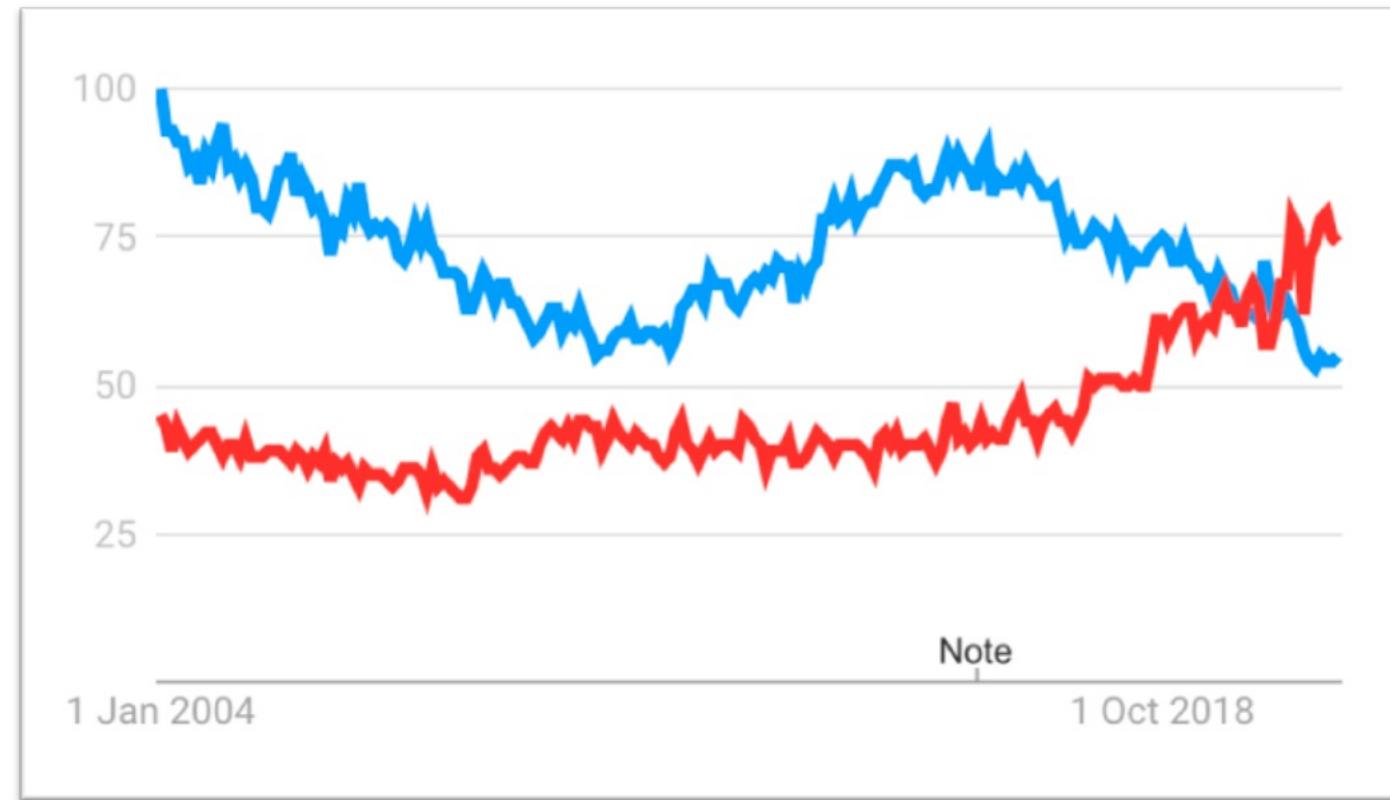
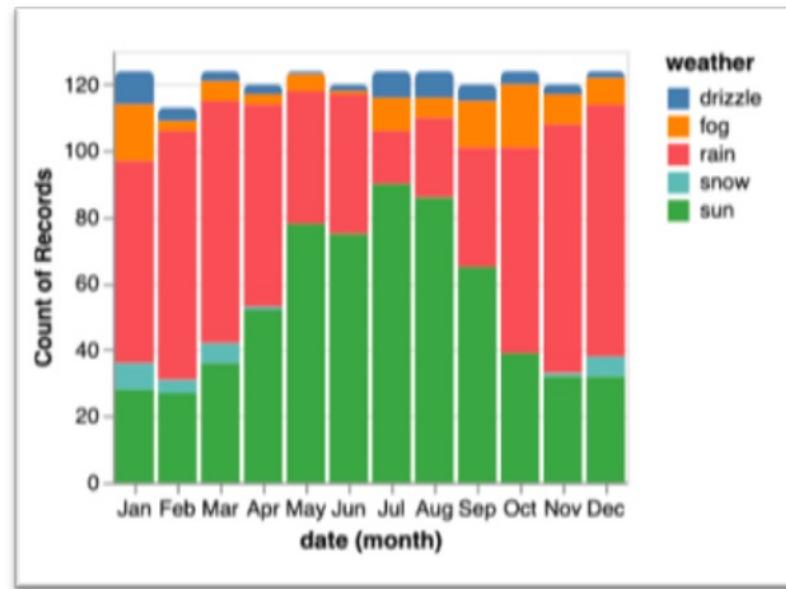


Chart has surpassed Image globally in Google Search Trend since Jan 2020

Visualization Data

- Visualization Data is defined as the digital representations of data visualizations.



Graphics

```
{  
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",  
  "data": {"url": "data/seattle-weather.csv"},  
  "mark": {"type": "bar", "cornerRadiusTopLeft": 3,  
           "cornerRadiusTopRight": 3},  
  "encoding": {  
    "x": {"timeUnit": "month", "field": "date", "type": "ordinal"},  
    "y": {"aggregate": "count"},  
    "color": {"field": "weather"}  
  }  
}
```

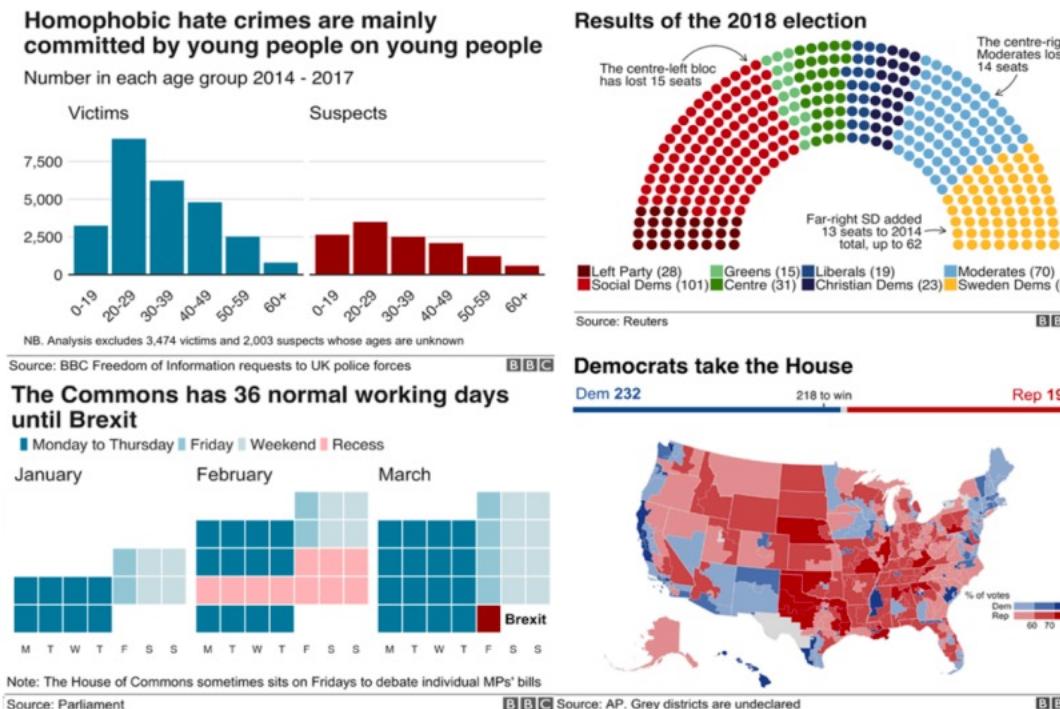
Program



Graphics



- Graphics are the most common format of visualization data:
 - Raster Image (e.g., JPEG, PNG)



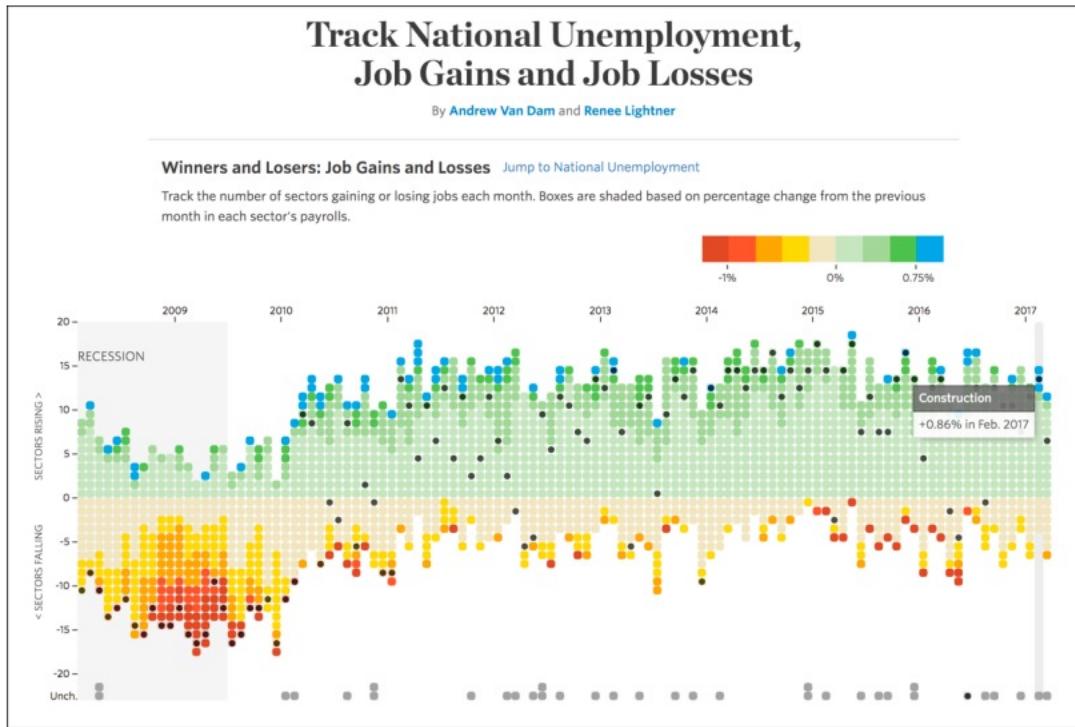
Raster images are generally what people think about when thinking of images.

- Common
- Easy-to-share
- Compatible

Graphics



- Graphics are the most common format of visualization data:
 - Raster Image (e.g., JPEG, PNG)
 - Vector Image (e.g., SVG)



Vector images have many benefits:

- Resolution-independent
- Interactivity
- Hierarchy with tags (in XML format)



Programs

- Researchers have developed many approaches to create and save visualizations as programs.
 - Imperative Programming



(D3, TVCG 2011)

```
d3.selectAll("circle")
  .data([4, 8, 15, 16, 23, 42])
    .style("r", function(d) { return d + "px"; });
```

Imperative programs are often:

- Unorganized
- Missing semantics (e.g. visual encodings)



- Researchers have developed many approaches to create and save visualizations as programs.
 - Imperative Programming
 - Declarative Programming

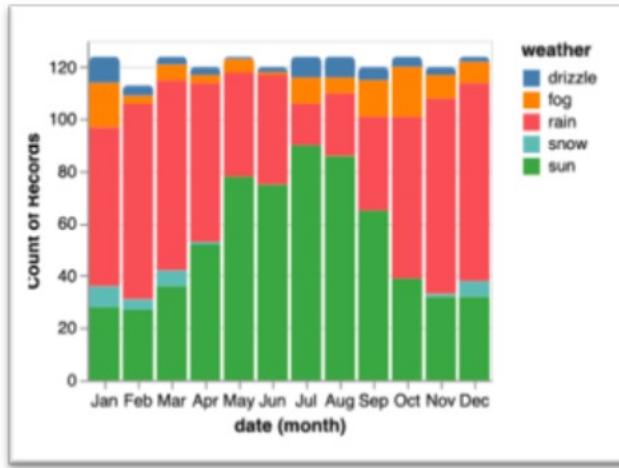


```
{  
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",  
  "data": {"url": "data/seattle-weather.csv"},  
  "mark": {"type": "bar", "cornerRadiusTopLeft": 3,  
           "cornerRadiusTopRight": 3},  
  "encoding": {  
    "x": {"timeUnit": "month", "field": "date", "type": "ordinal"},  
    "y": {"aggregate": "count"},  
    "color": {"field": "weather"}  
  }  
}
```

Vega & Vega-Lite VISUALIZATION GRAMMARS
(Vega, TVCG 2015; Vega-Lite, TVCG 2017)



Taxonomy of Visualization Data



Graphics



Human-friendly
Easy-to-share

Hybrids

Program

```
{  
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",  
  "data": {"url": "data/seattle-weather.csv"},  
  "mark": {"type": "bar", "cornerRadiusTopLeft": 3,  
  "cornerRadiusTopRight": 3},  
  "encoding": {  
    "x": {"timeUnit": "month", "field": "date", "type":  
    "ordinal"},  
    "y": {"aggregate": "count"},  
    "color": {"field": "weather"}  
  }  
}
```

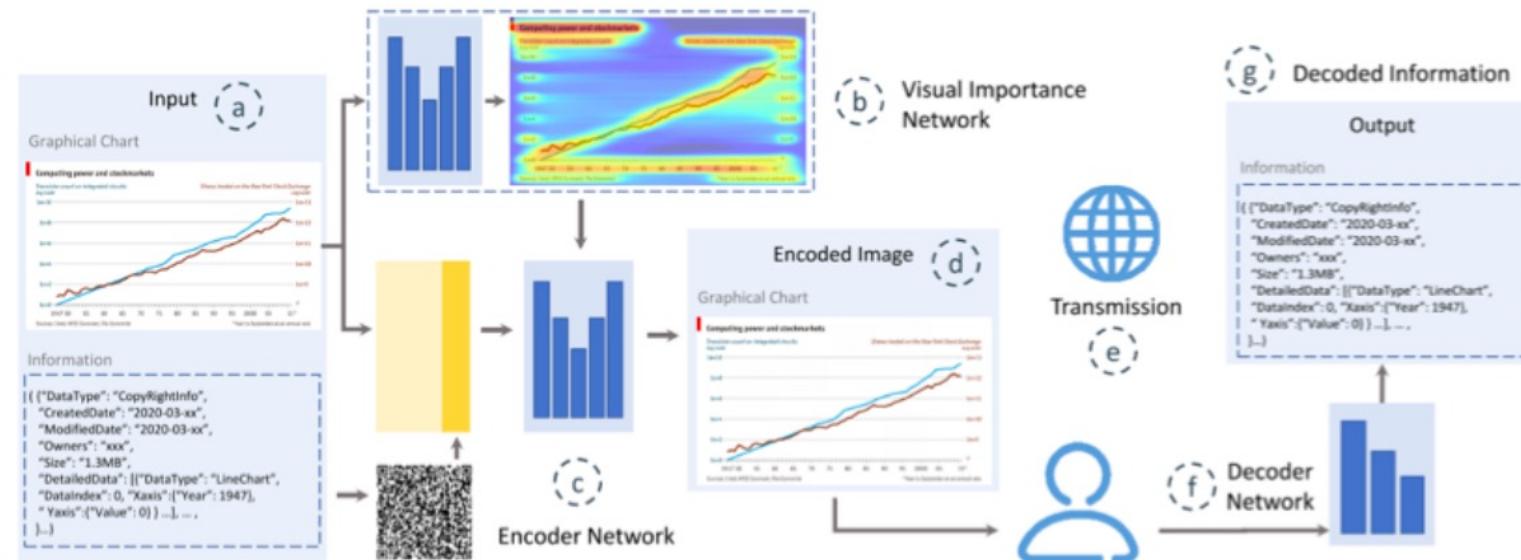
Machine-friendly
Easy-to-process



Hybrids



- Several hybrid formats have been recently proposed for the combined benefits of graphics and programs



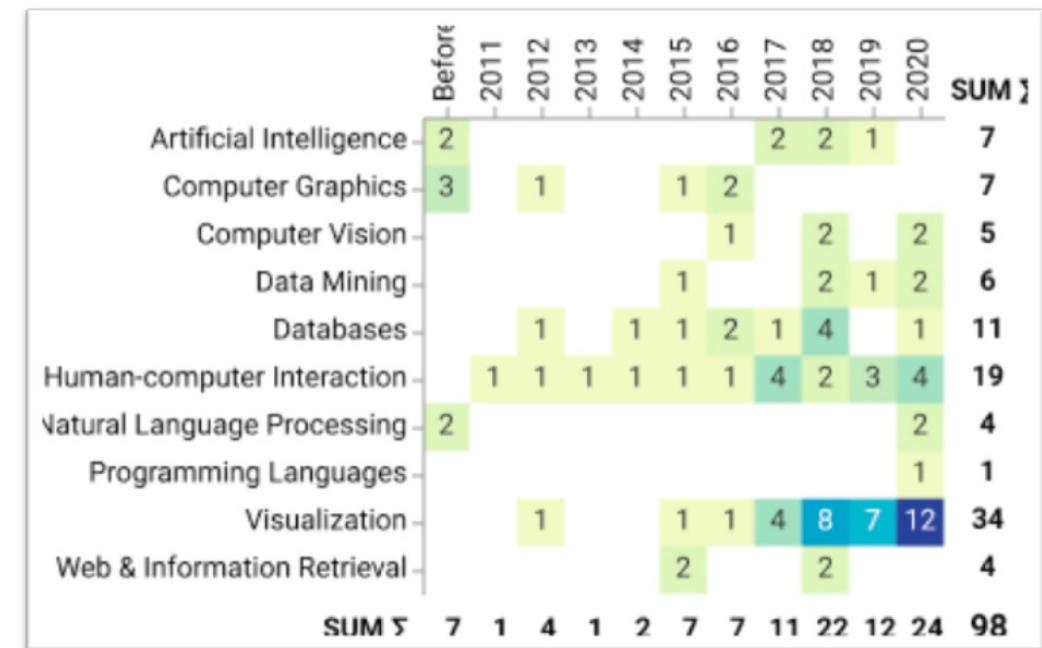
(VisCode, TVCG 2020)

A long way to go before such standards can be adopted



AI4VIS

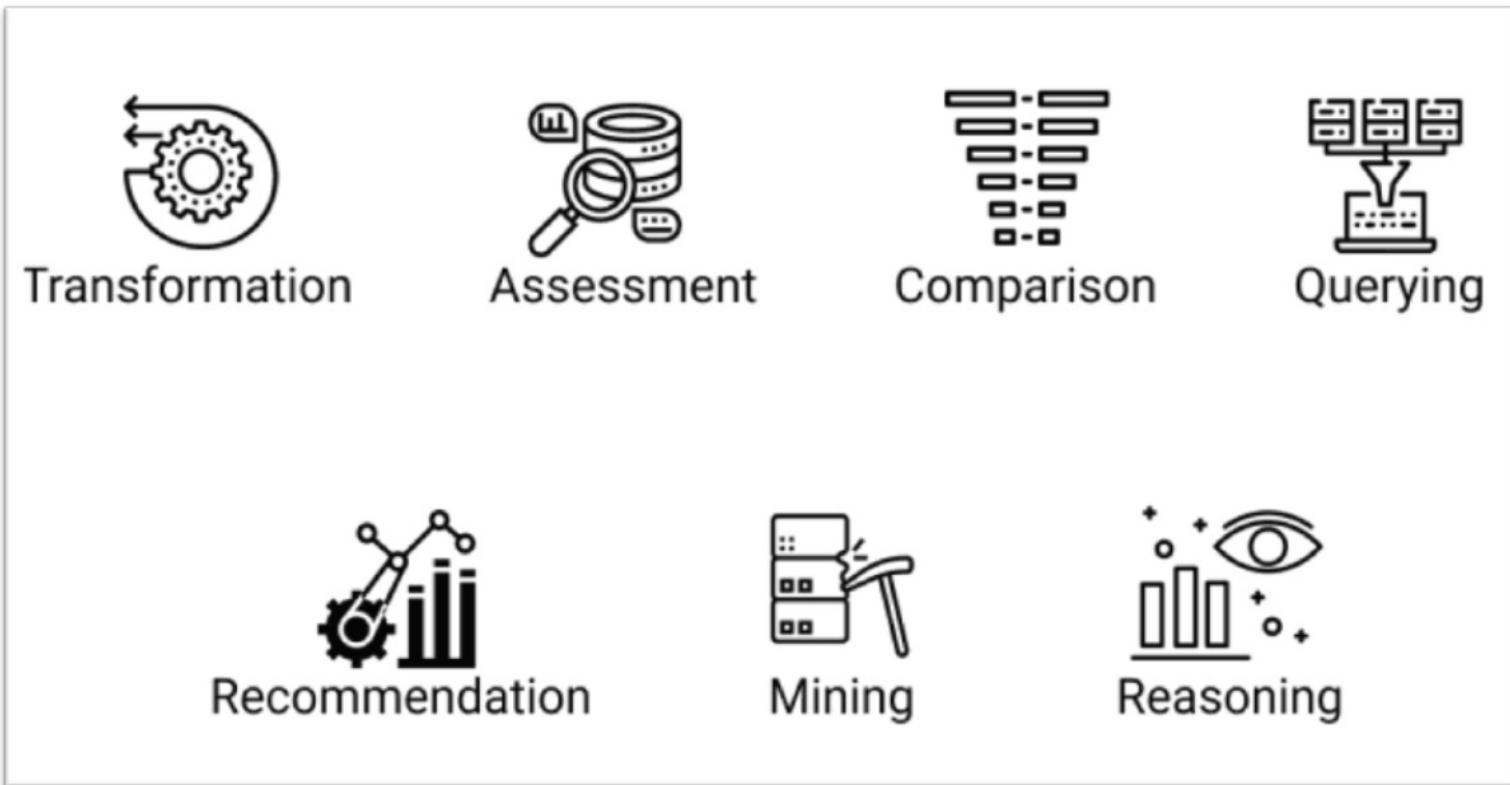
- AI4VIS has a diversity of not only its roots (visualization data), but also its research community and research problems



Surveyed paper by the research area and year

Taxonomy of Research Problems

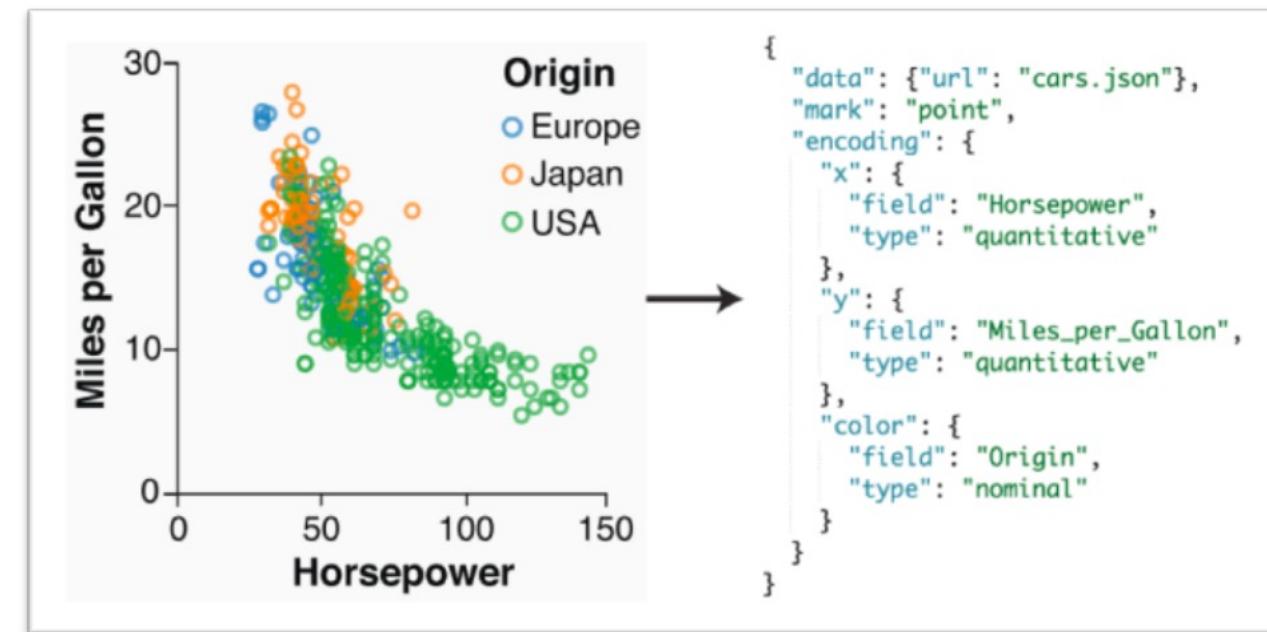
- AI4VIS research primarily focuses on 7 problems



Transformation



- Transformation converts visualizations from one format to another (frequency: 35/98)

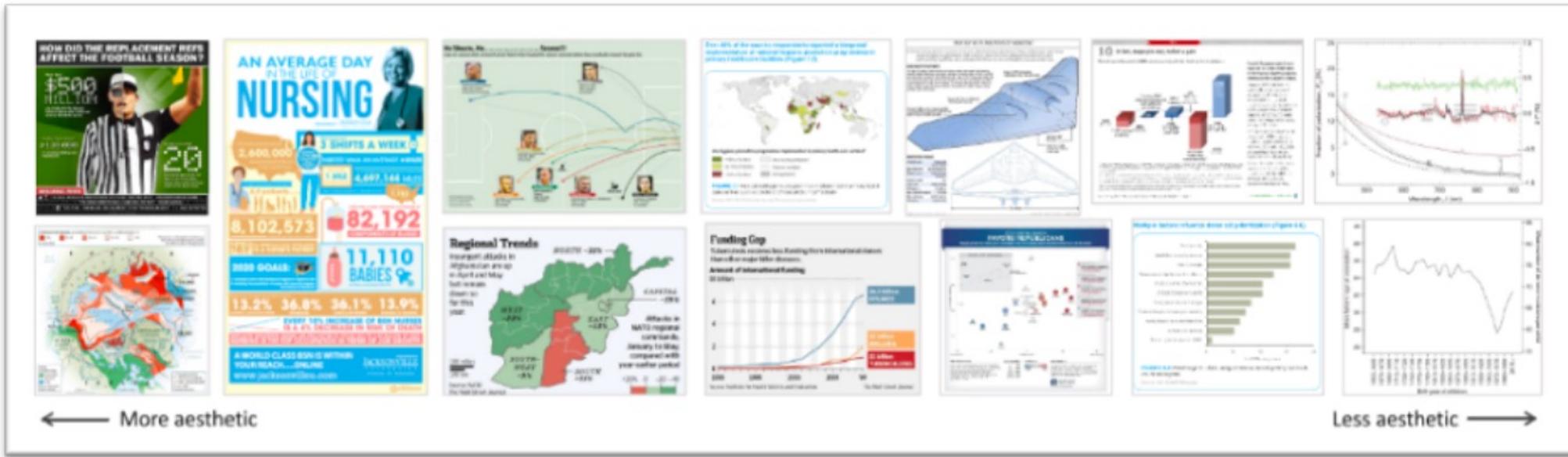


Reverse-Engineering Visualizations: Recovering Visual Encodings from Chart Images.
(EuroVIS 2017)



Assessment

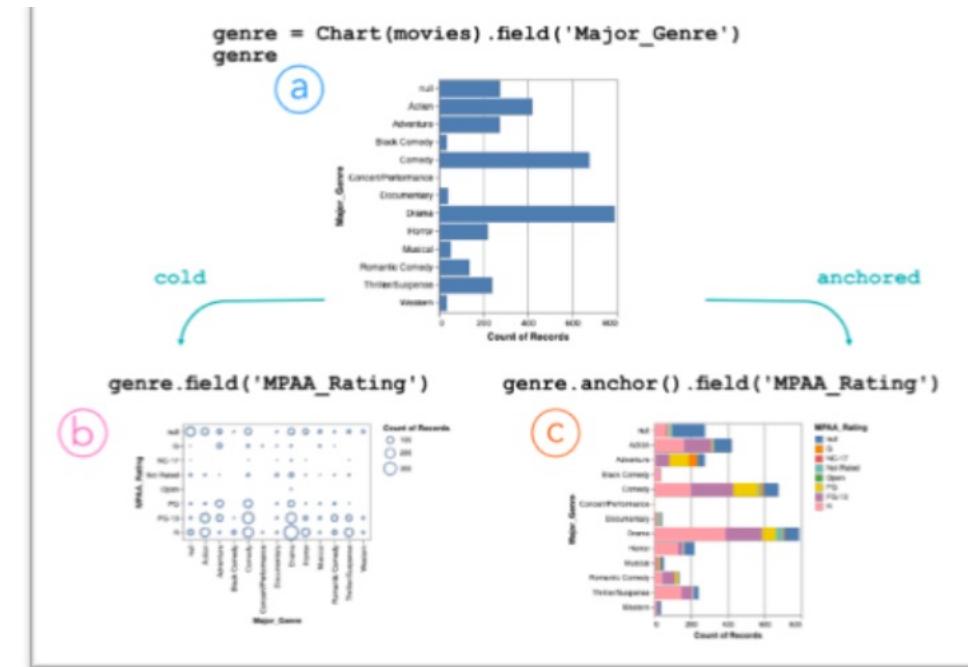
- Assessment measures the absolute or relative quality of a visualization in terms of scores or ranking (frequency: 19/98)



Visualization Assessment: A Machine Learning Approach
(VIS 2019)

Comparison

- Comparison estimates the similarity or other metrics between two visualizations (Frequency: 12/98)

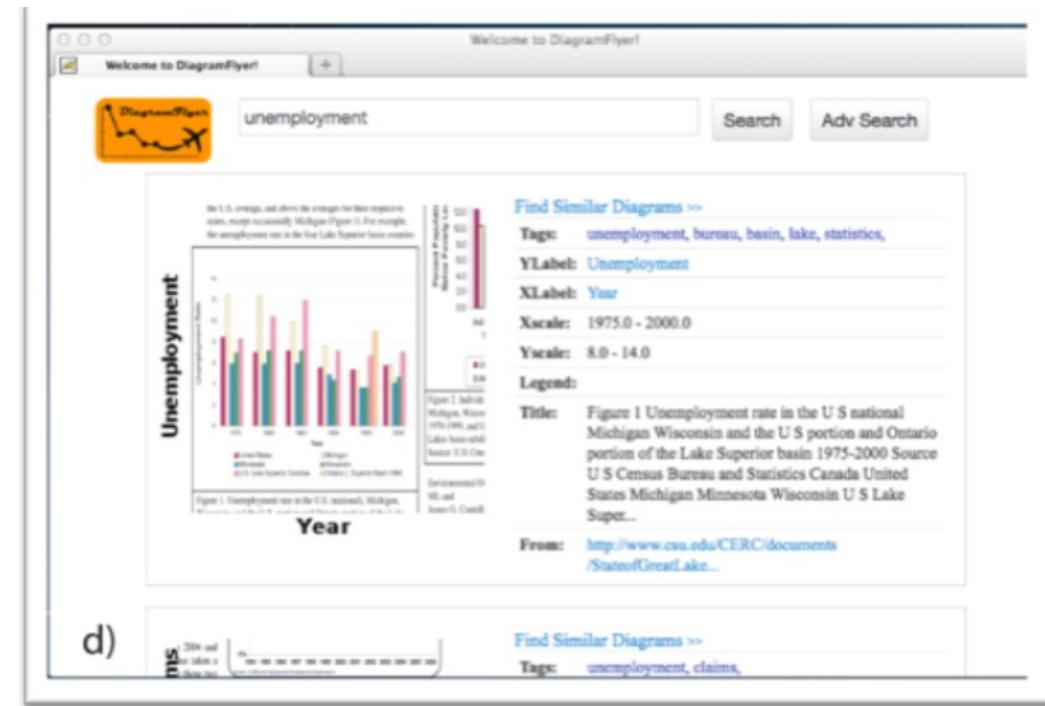


Dziban: Balancing Agency & Automation in Visualization Design
via Anchored Recommendations
(CHI 2020)

Querying



- Querying finds the target visualization relevant with a user query within visualization collections (frequency: 10/98)



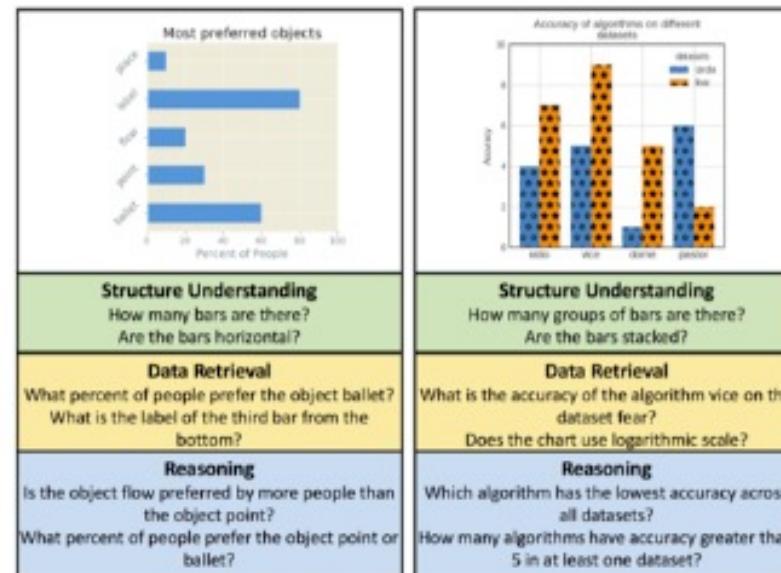
DiagramFlyer : A Search Engine for Data-Driven Diagrams
(WWW 2015)



Reasoning



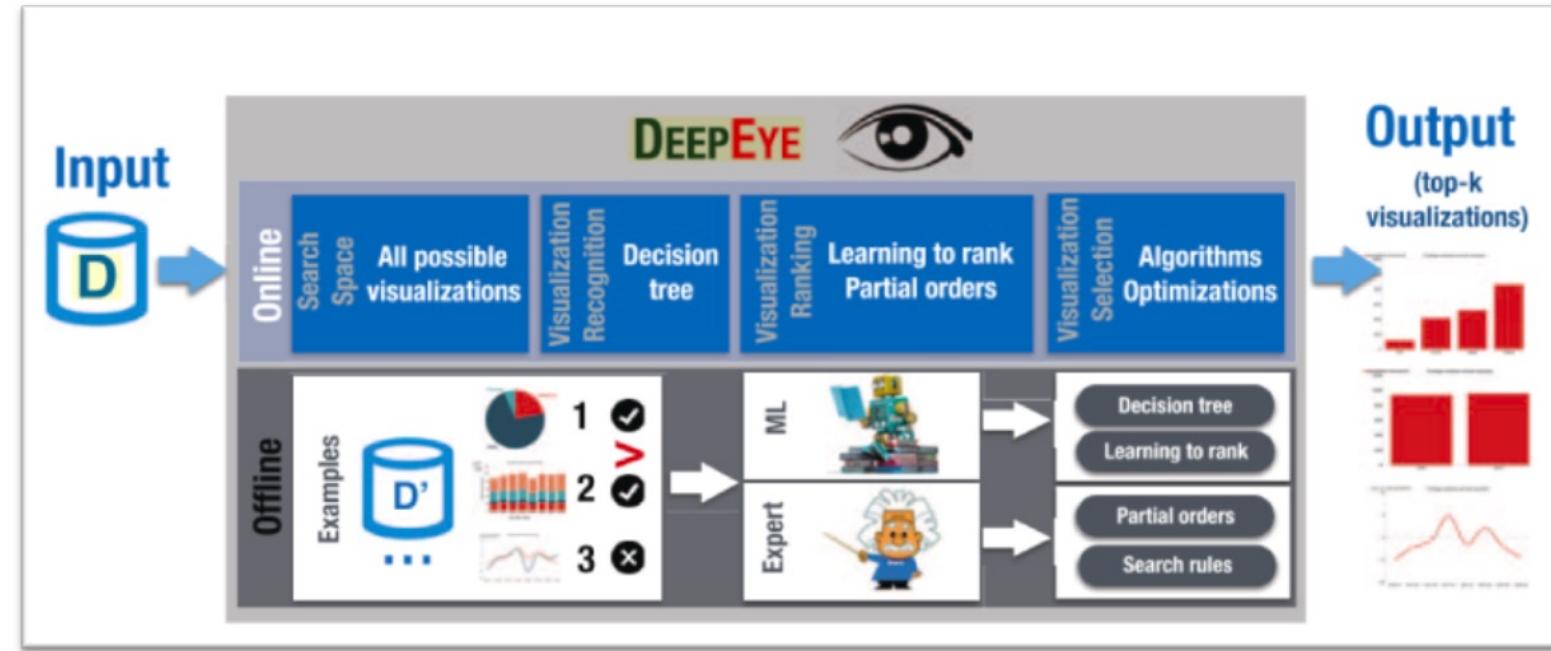
- Reasoning challenges AI to interpret visualizations to derive high-level information like insights and summaries (frequency: 22/98)



DVQA: Understanding Data Visualizations via Question Answering
(CVPR 2018)

Recommendation

- Recommendation automates the creation of visualizations by suggesting data and/or visual encodings (Frequency: 38/98)



Deepeye: towards automatic data visualization
(ICDE 2018)

- Mining discovers insights from visualization databases (Frequency: 7/98)



Composition and Configuration Patterns in Multiple-View Visualization
(TVCG 2020)



Datasets

A	B	C	D	E	F
Year	Brand	Category	Model	Sales	
1/1/2007	BMW	Compact	BMW 3-Se	142490	
1/1/2008	BMW	Compact	BMW 3-Se	112464	
1/1/2009	BMW	Compact	BMW 3-Se	90960	
1/1/2010	BMW	Compact	BMW 3-Se	100910	
1/1/2011	BMW	Compact	BMW 3-Se	94371	
1/1/2007	BMW	Compact	BMW 5-Se	54142	
1/1/2008	BMW	Fullsize	BMW 7-Se	14773	
1/1/2009	BMW	Fullsize	BMW 7-Se	12276	
1/1/2010	BMW	Fullsize	BMW 7-Se	9254	
1/1/2011	BMW	Fullsize	BMW 7-Se	12253	
1/1/2008	BMW	Midsize	BMW 5-Se	45915	
1/1/2009	BMW	Midsize	BMW 5-Se	40109	
1/1/2010	BMW	Midsize	BMW 5-Se	39488	
1/1/2011	BMW	Midsize	BMW 5-Se	51491	
1/1/2007	BMW	Sporty	BMW 6-Se	9033	
1/1/2008	BMW	Sporty	BMW 6-Se	6533	
1/1/2009	BMW	Sporty	BMW 6-Se	3549	
1/1/2010	BMW	Sporty	BMW 6-Se	2418	
1/1/2011	BMW	Sporty	BMW 6-Se	3903	
1/1/2007	BMW	Sporty	BMW Z4	10097	
1/1/2008	BMW	Sporty	BMW Z4	5879	
1/1/2009	BMW	Sporty	BMW Z4	3523	
1/1/2010	BMW	Sporty	BMW Z4	3804	

Visualizations





Data Analysis

Visual Design



Data Analysis



Data contents should be
Informative, interesting, insightful
Explore data, find, and organize



Visual Design

Visual Design should be aesthetically pleasing

Design visual components, layouts and styles



How to ease visualization design and creativity



Data Analysis



Visual Design



Professional tools can ease the process, but still very hard for novice users.



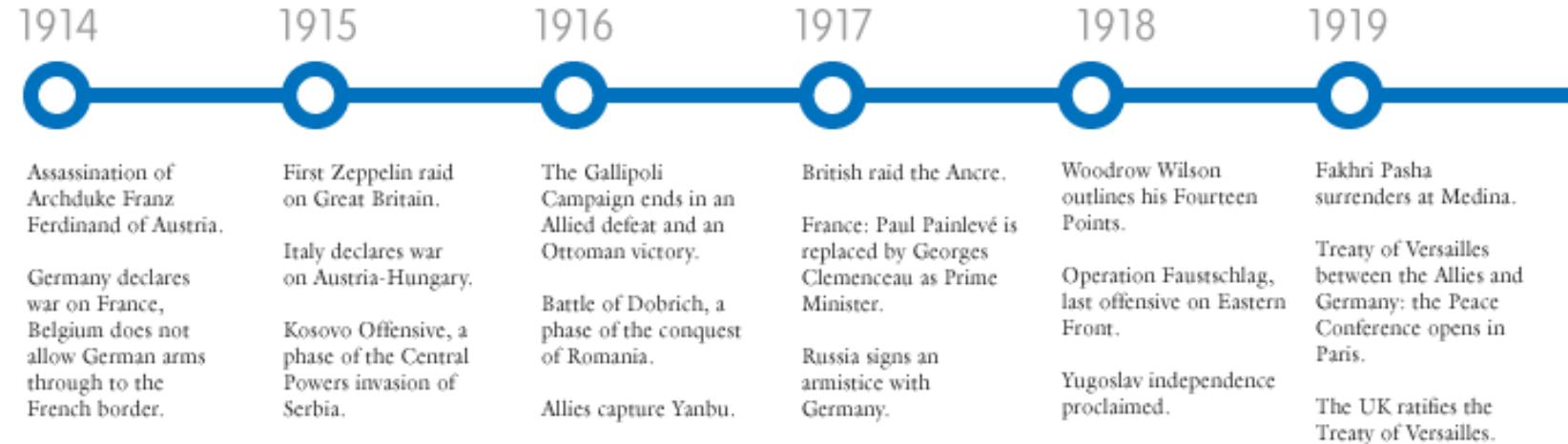
Visualization Design vs. General People

- A typical paradigm to create a visualization
 - Option 1: Ask a designer to ...
 - Communication overhead
 - Money cost
 - Option 2: Do it myself
 - Unfamiliar with graphic tools
 - Lack of design ideas
- Ad-hoc, case-by-case, and time consuming
 - Prevent it to be first-class method for presentation and communication
- Our Solution: (1) reuse existing designs and (2) generate new designs



Research Case 1

Auto-Extraction of Extensible Timeline



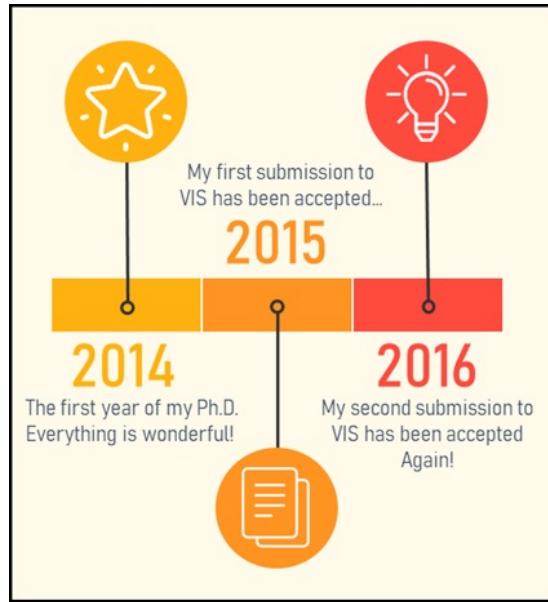
Timeline of World War I



Motivation: How to Reuse the Timeline Design?

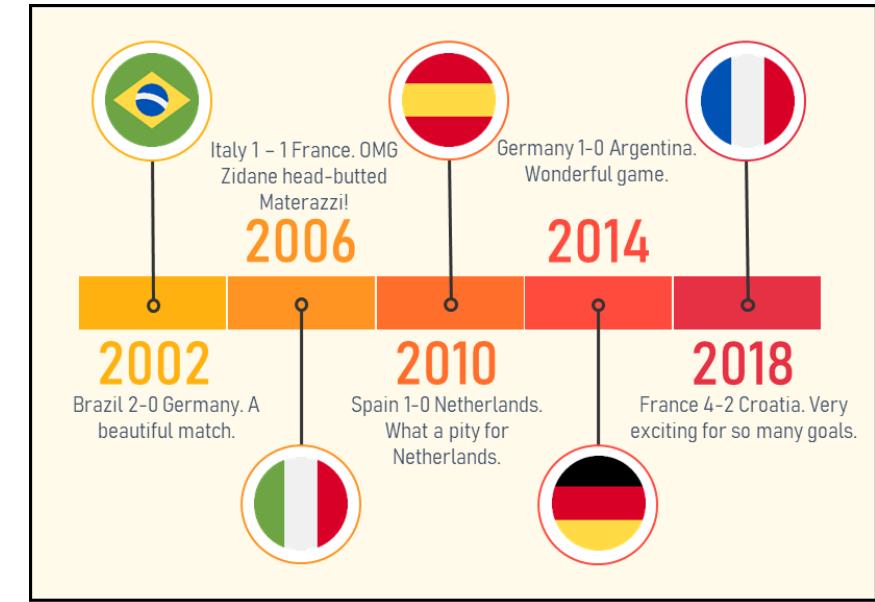
2002	Brazil 2-0 Germany. A beautiful match.
2006	Italy 1-1 France. OMG Zidane head-butted Materazzi!
2010	Spain 1-0 Netherlands. What a pity for Netherlands.
2014	Germany 1-0 Argentina. Wonderful game.
2018	France 4-2 Croatia. Very exciting for so many goals

+



Timeline data

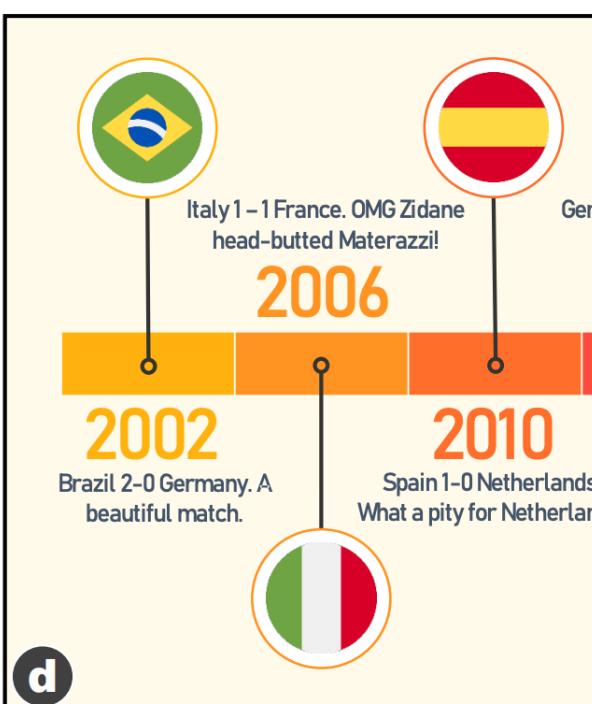
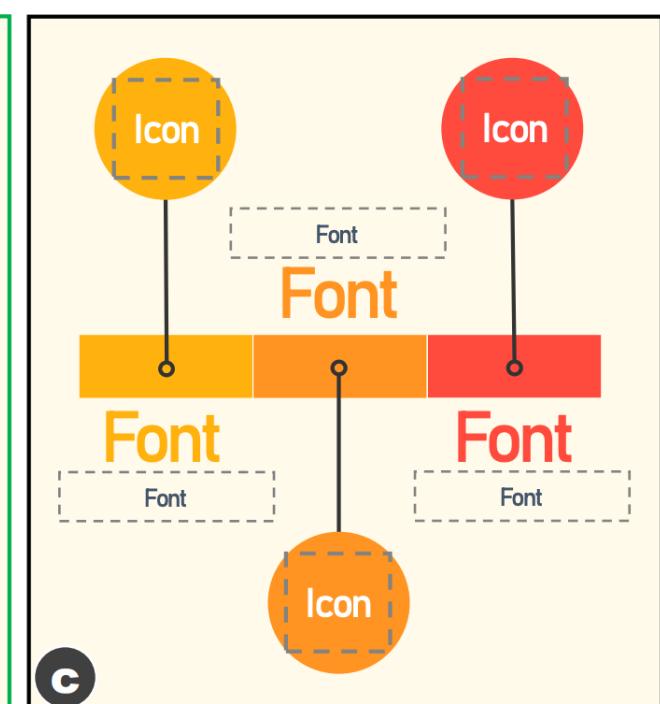
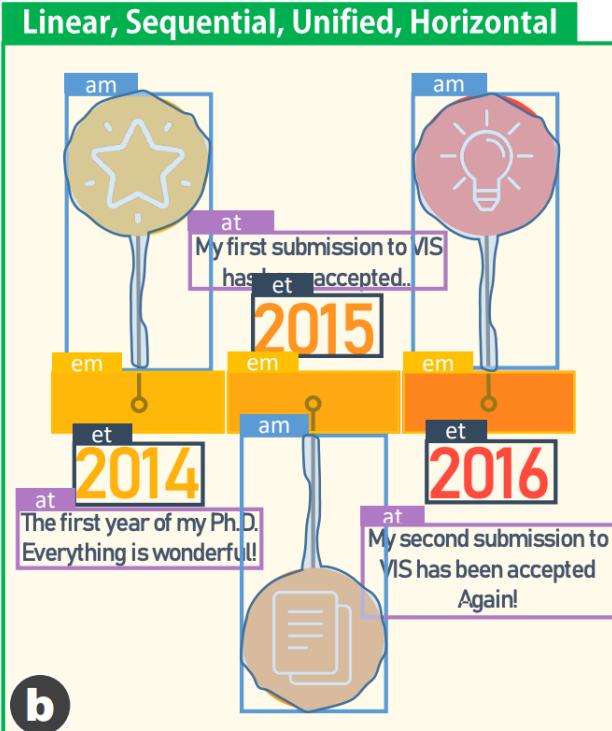
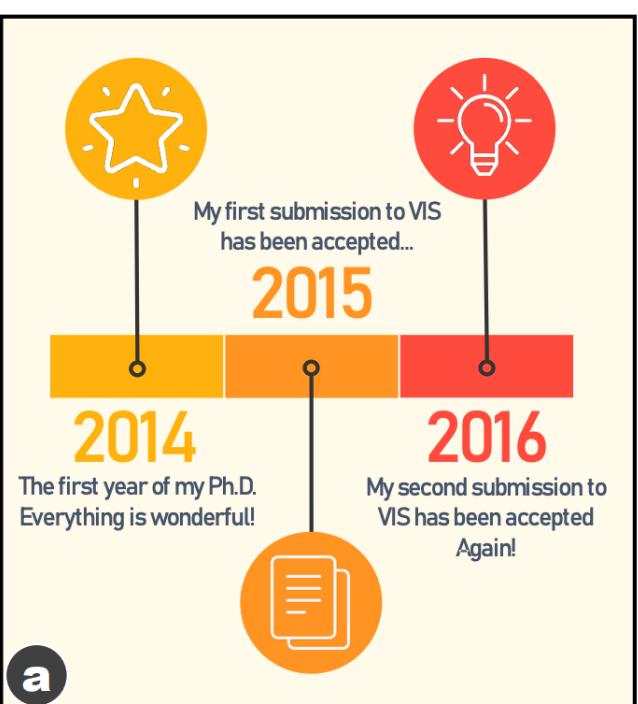
=



Timeline image

Updated timeline

Zhutian Chen, Yun Wang, Qianwen Wang, Yong Wang, Huamin Qu. Towards Automated Infographic Design: Deep Learning-based Auto-Extraction of Extensible Timeline (InfoVis '19)



Bitmap Image

Deconstruct

Content
Understanding

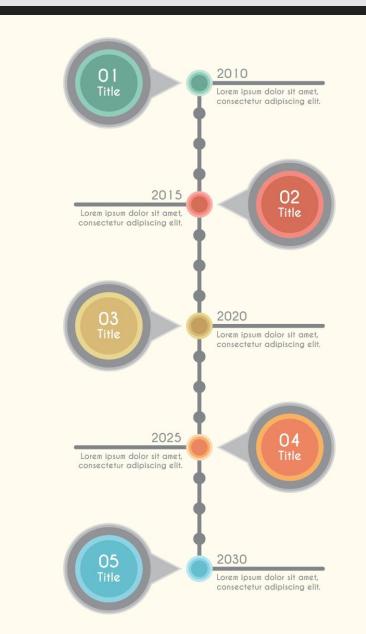
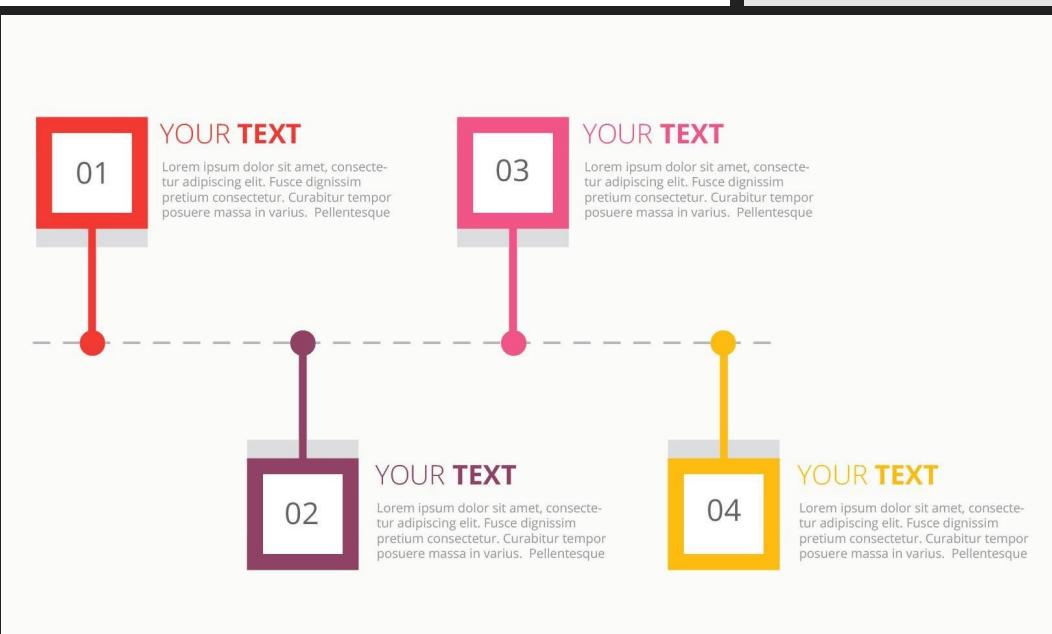
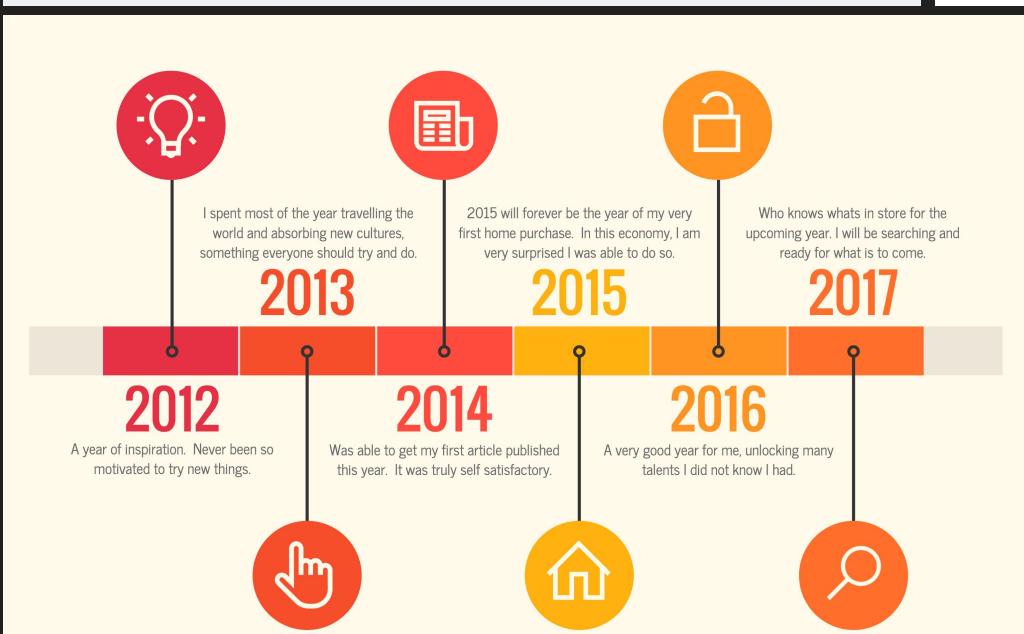
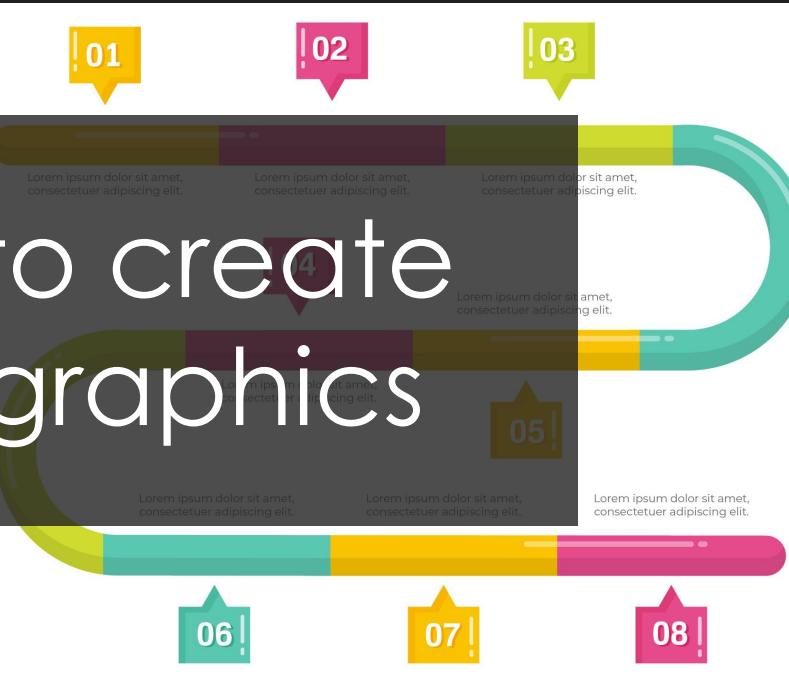
Reconstruct

Editable
Template

Render

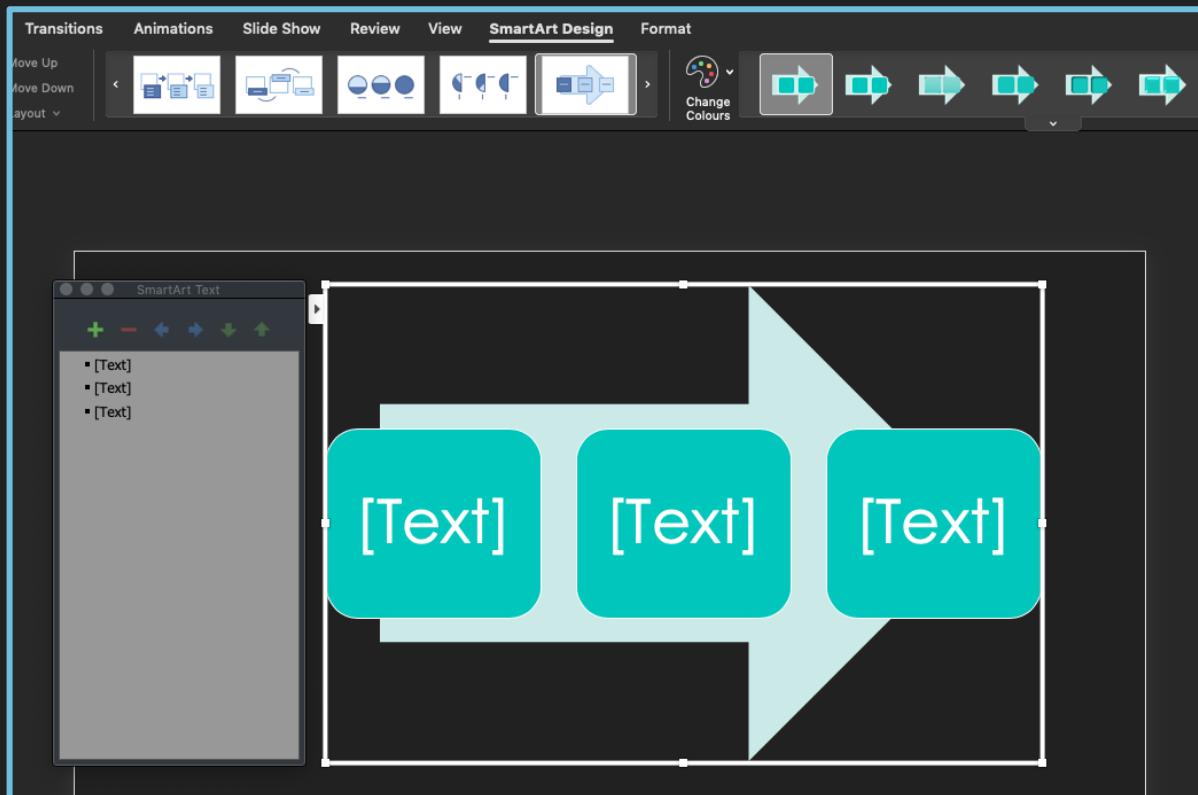
New Infographic

Towards Automated Infographic Design: Deep Learning-based Auto-Extraction of Extensible Timeline

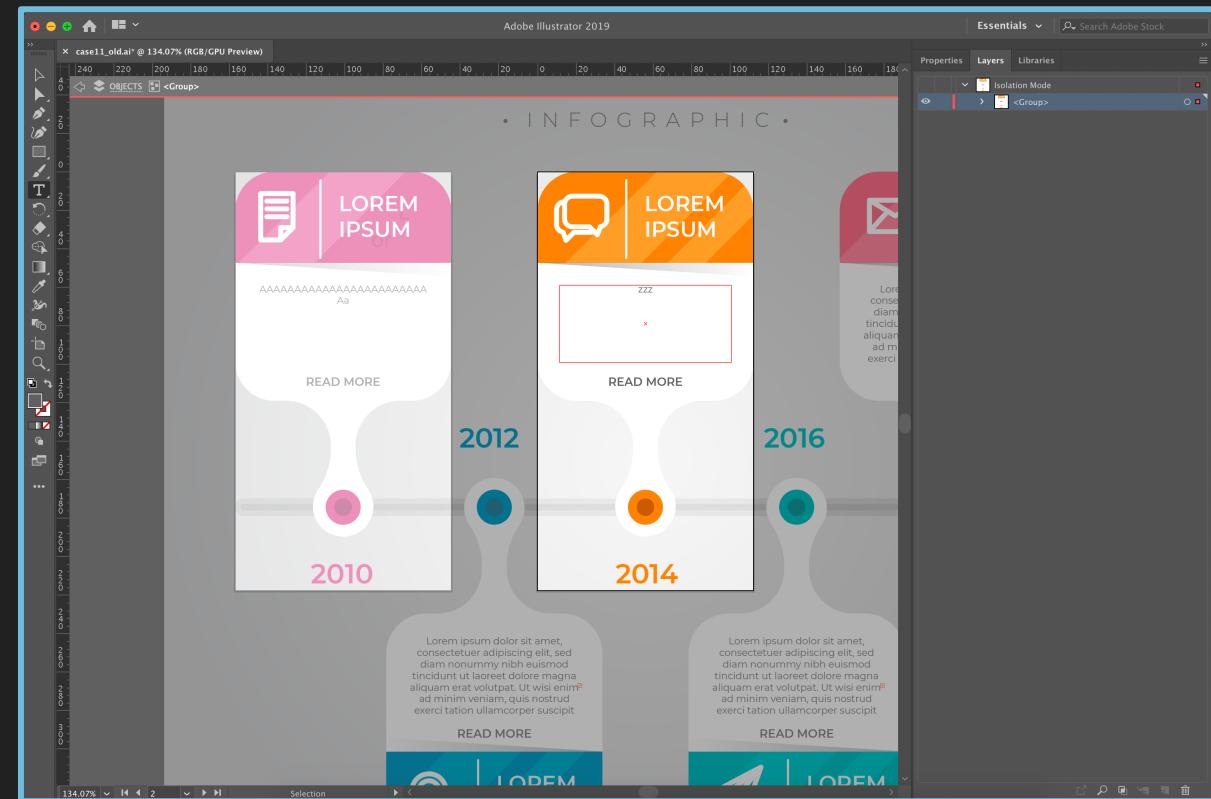


Template

PowerPoint

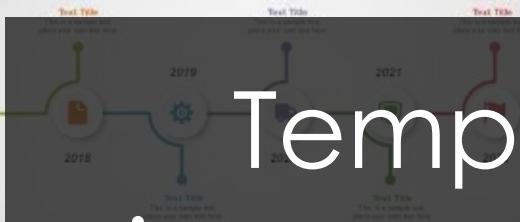


Illustrator



Template can automate the process

Presentation Timeline – 5 Years

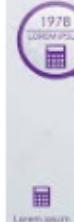


Free Timeline Template for PowerPoint ...
slidemodel.com



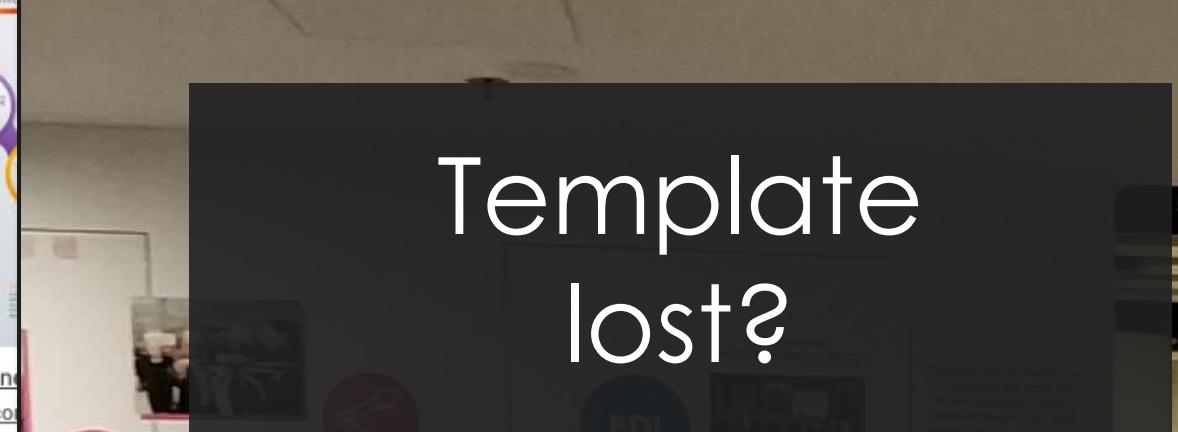
Timeline infographics design templa...
vectorstock.com

Timeline info



Timeline
123rf.co

Template lost?



timeline infographic template Vector ...
freepik.com



How Do I Create a Timeline Inf...
venngage.com



Timeline Infographics Th...
pinterest.com



line Infographic with charts ...
emarket.com



Timeline Images, Stock Photos & Vector...
shutterstock.com



Timeline Infographics Templates
showeet.com

File Home Insert Design Transitions Animations Slide Show Review View Add-Ins 福昕PDF 福昕阅读器 Storyboarding Format Tell me what you want to do... chen zhtian Share

No template?

A timeline graphic showing personal milestones from 2012 to 2017. The timeline is represented by a horizontal bar divided into segments for each year. Above the bar are three icons: a lightbulb for 2013, a document for 2015, and a lock for 2017. Below the bar are three icons: a hand cursor for 2012, a house for 2014, and a magnifying glass for 2016. Each icon has a corresponding text description:

- 2013: I spent most of the year travelling the world and absorbing new cultures. something everyone should try and do.
- 2015: 2015 will forever be the year of my very first home purchase. In this economy, I am very surprised I was able to do so.
- 2017: Who knows what's in store for the upcoming year. I will be searching and ready for what is to come.
- 2012: A year of inspiration. Never been so motivated to try new things.
- 2014: Was able to get my first article published this year. It was truly self satisfactory.
- 2016: A very good year for me, unlocking many talents I did not know I had

Click to add notes

Slide 1 of 13 English (United States)

Drawing Editing

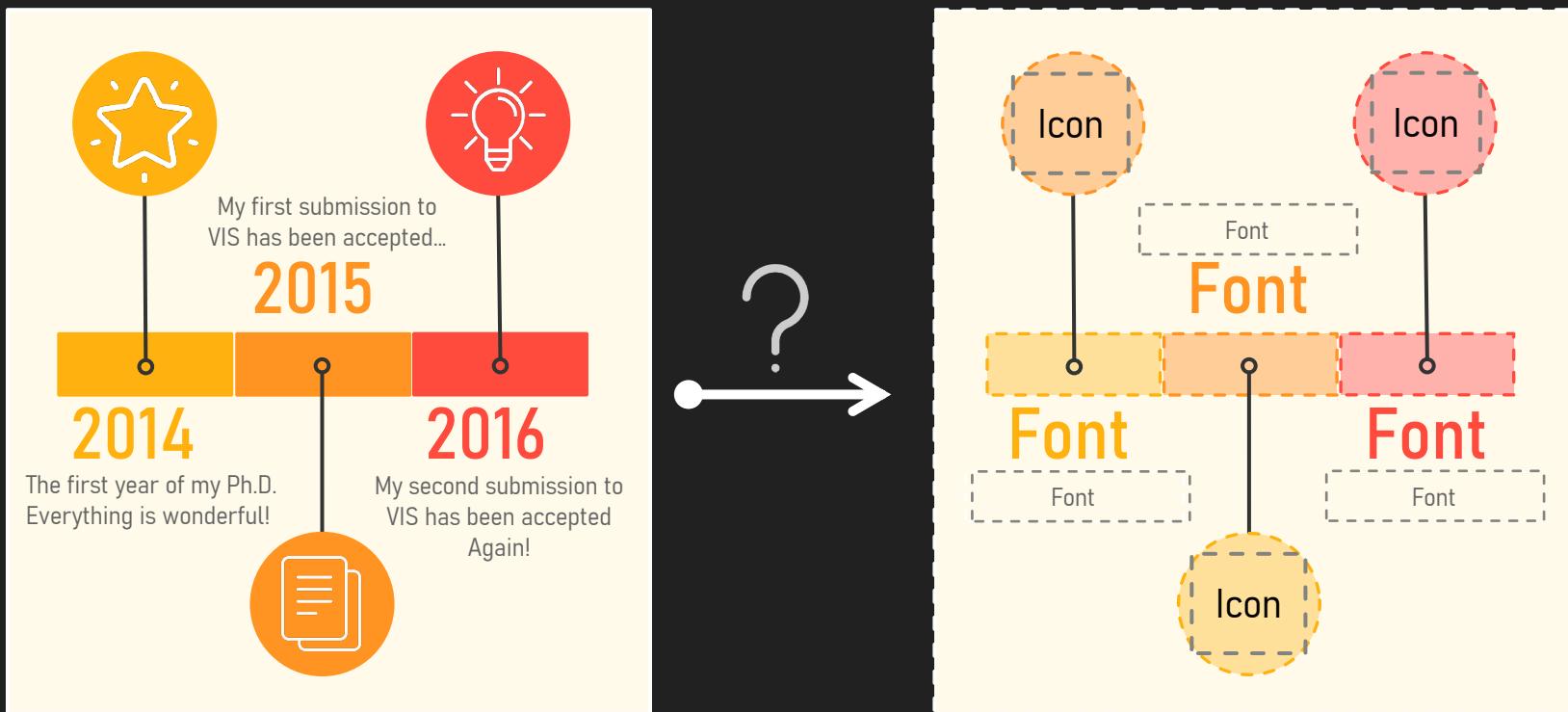
Clipboard Slides Font Paragraph Drawing Editing

Calibri (Body) 18 A A A S abe AV Aa A Text Direction Layout Reset New Slide Section Convert to SmartArt

Format Painter Paste

Arrange Quick Styles Find Replace Select

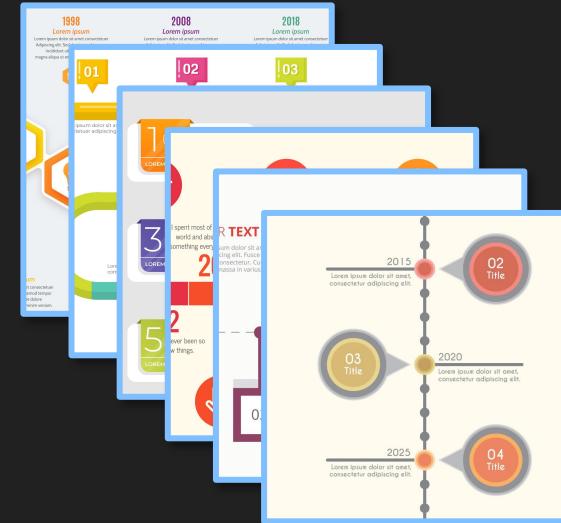
Goal



How to extract the template from a bitmap timeline infographic? automatically?

Challenge

1. How to automatically understand the content of a bitmap timeline infographic?
2. How to automatically convert a bitmap timeline to be extensible?



Immutable



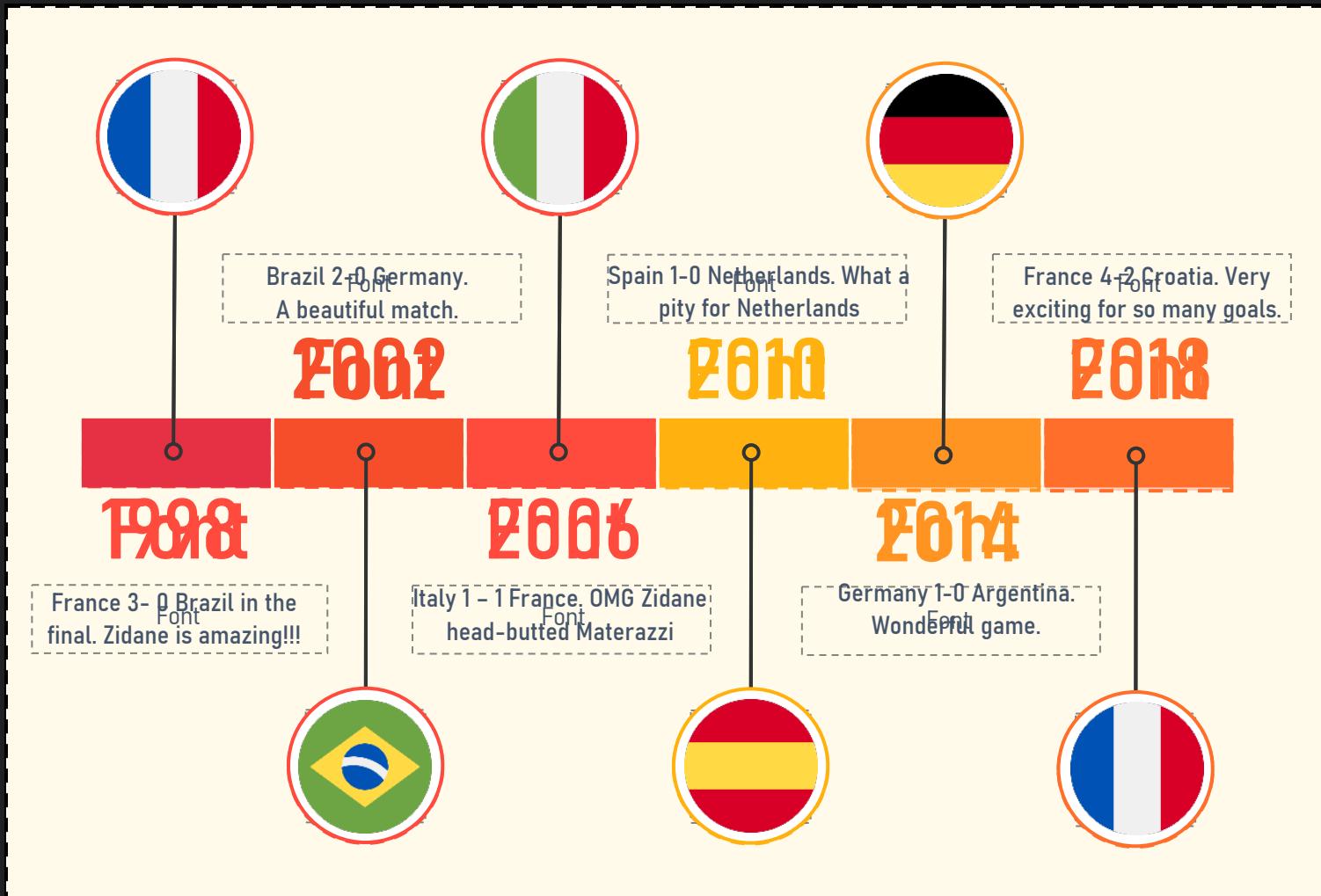
Editable



Extensible

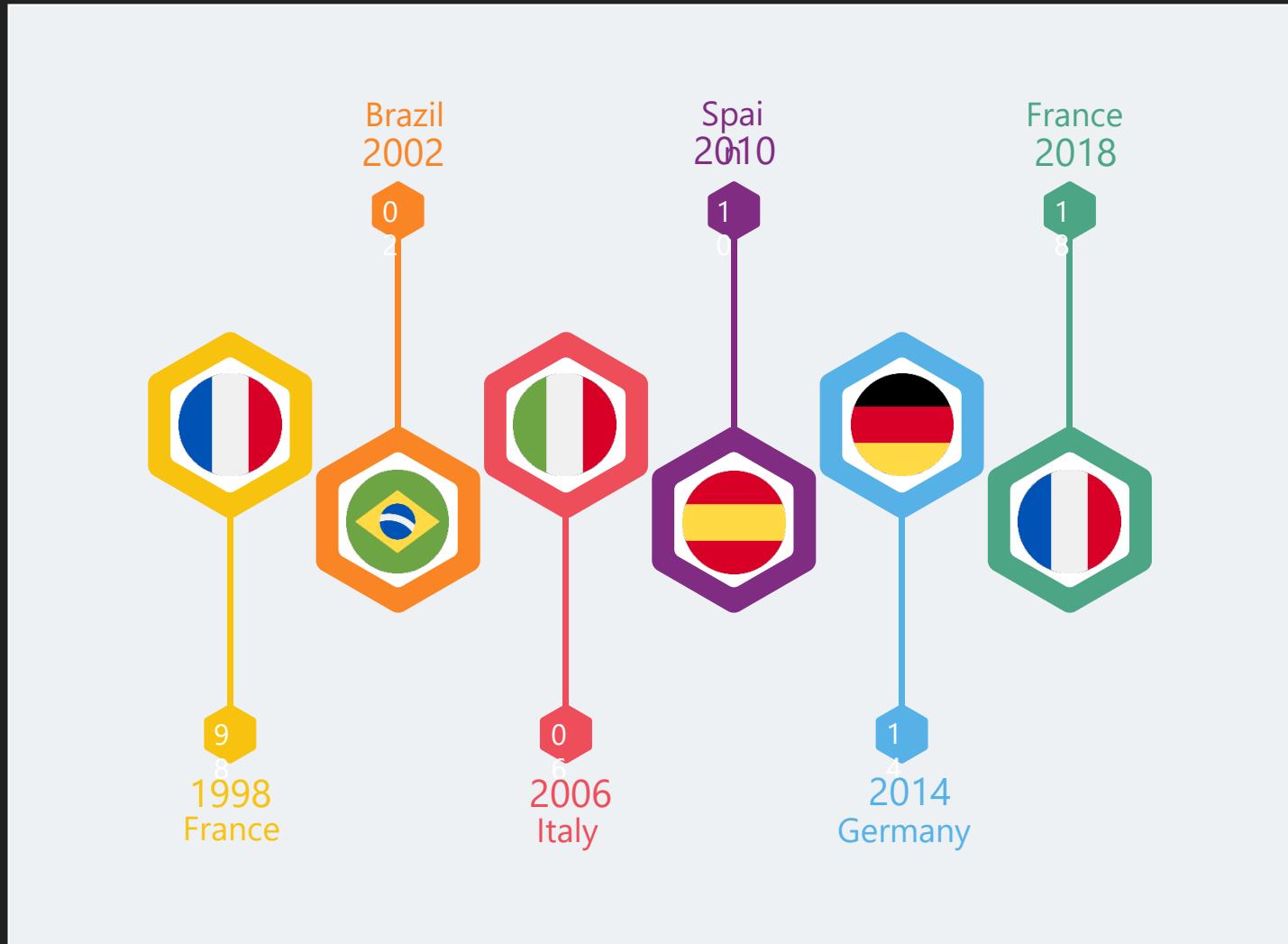
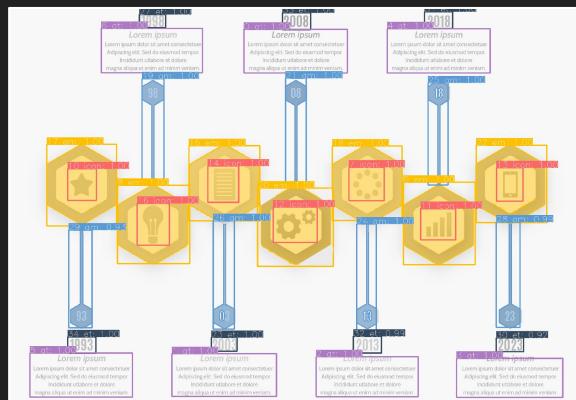
Applications

Input the data to the template...



Applications

Or transfer the style to other template...

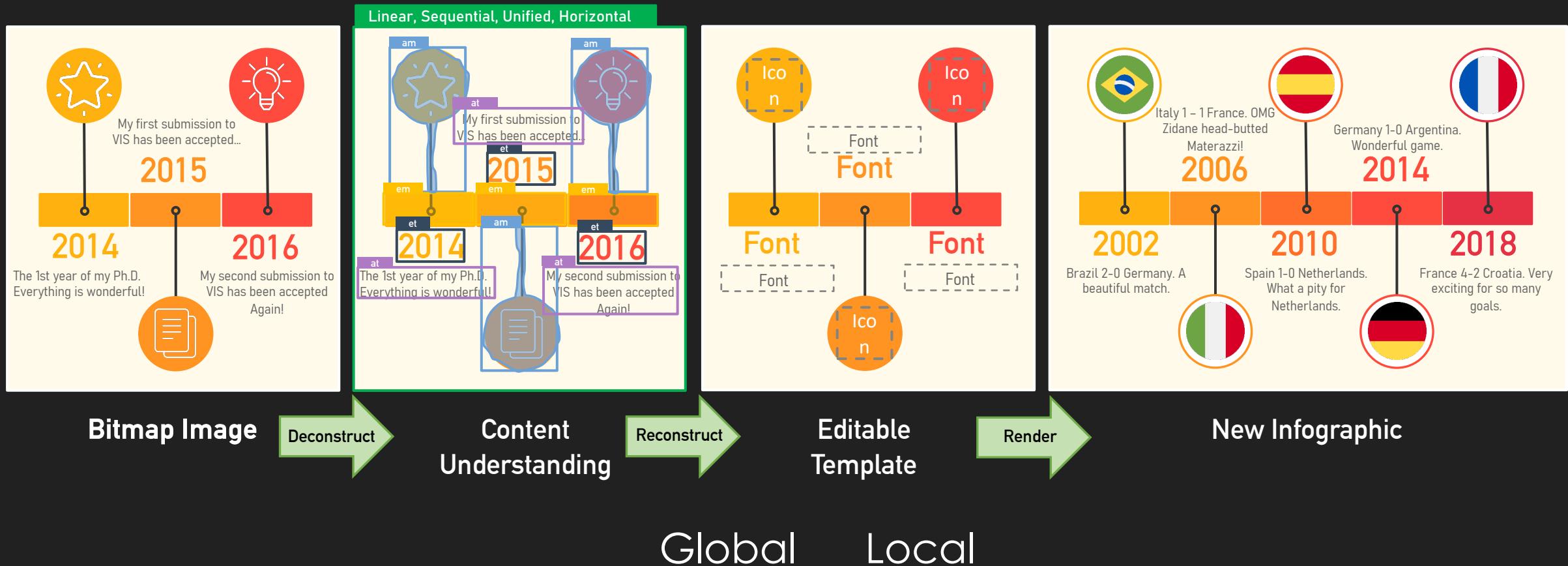


Applications

Or even create
AR visualization
in-situ...



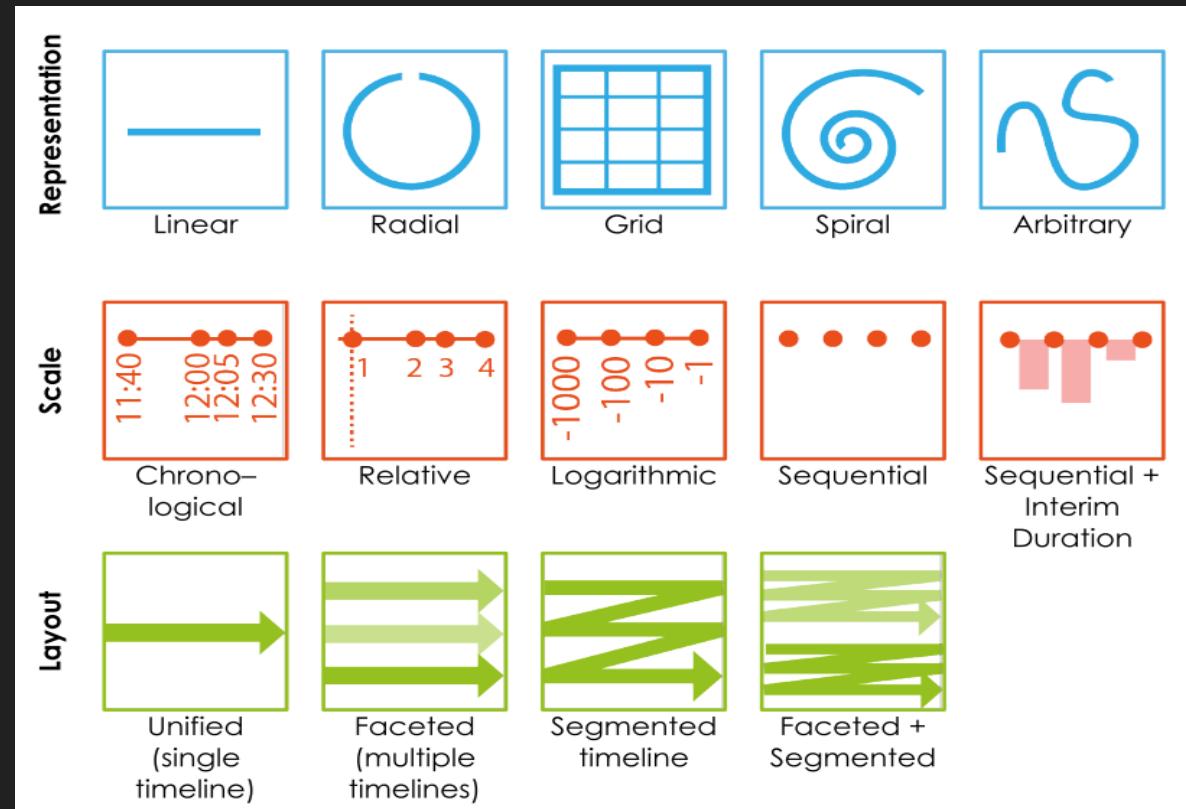
Pipeline – Deconstruction



Deconstruction – Global

About the whole timeline:

1. Representation
2. Scale
3. Layout
4. Orientation

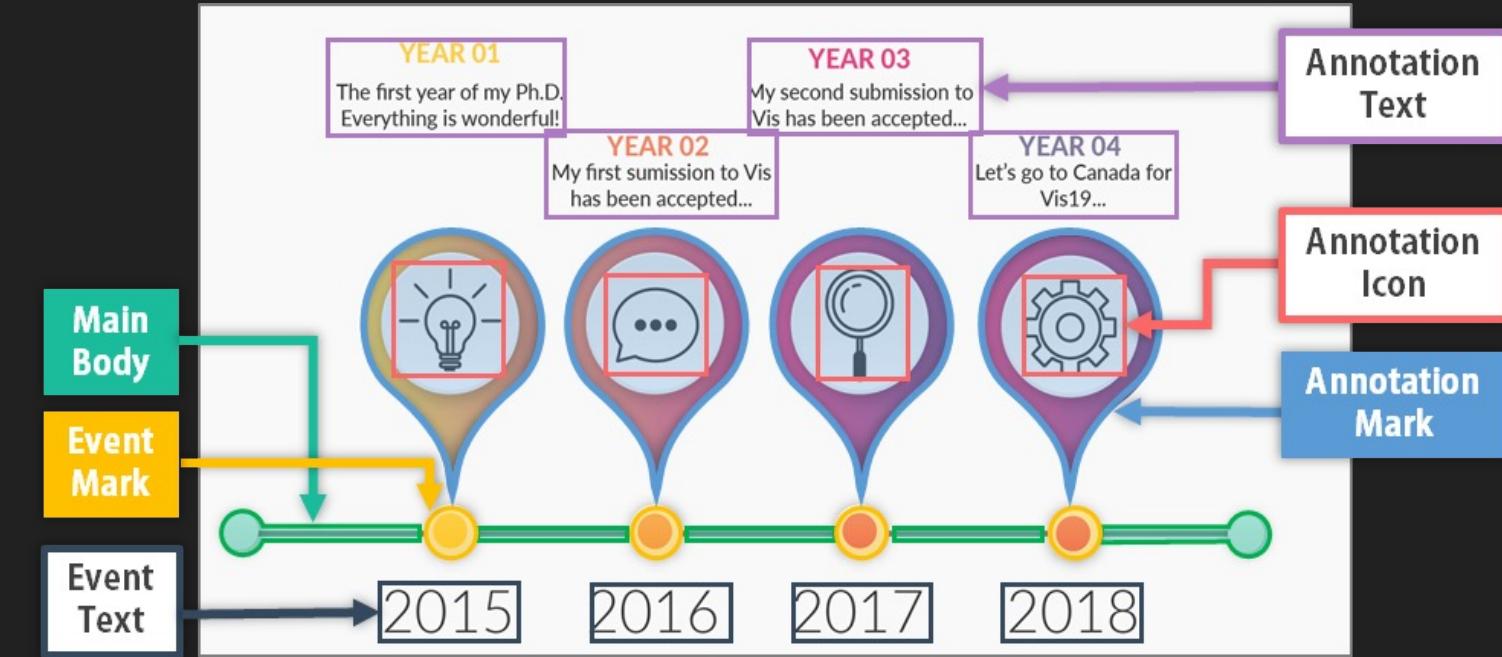


M. Brehmer, B. Lee, B. Bach, N. H. Riche, and T. Munzner. Timelines Revisited: A Design Space and Considerations for Expressive Storytelling. IEEE TVCG

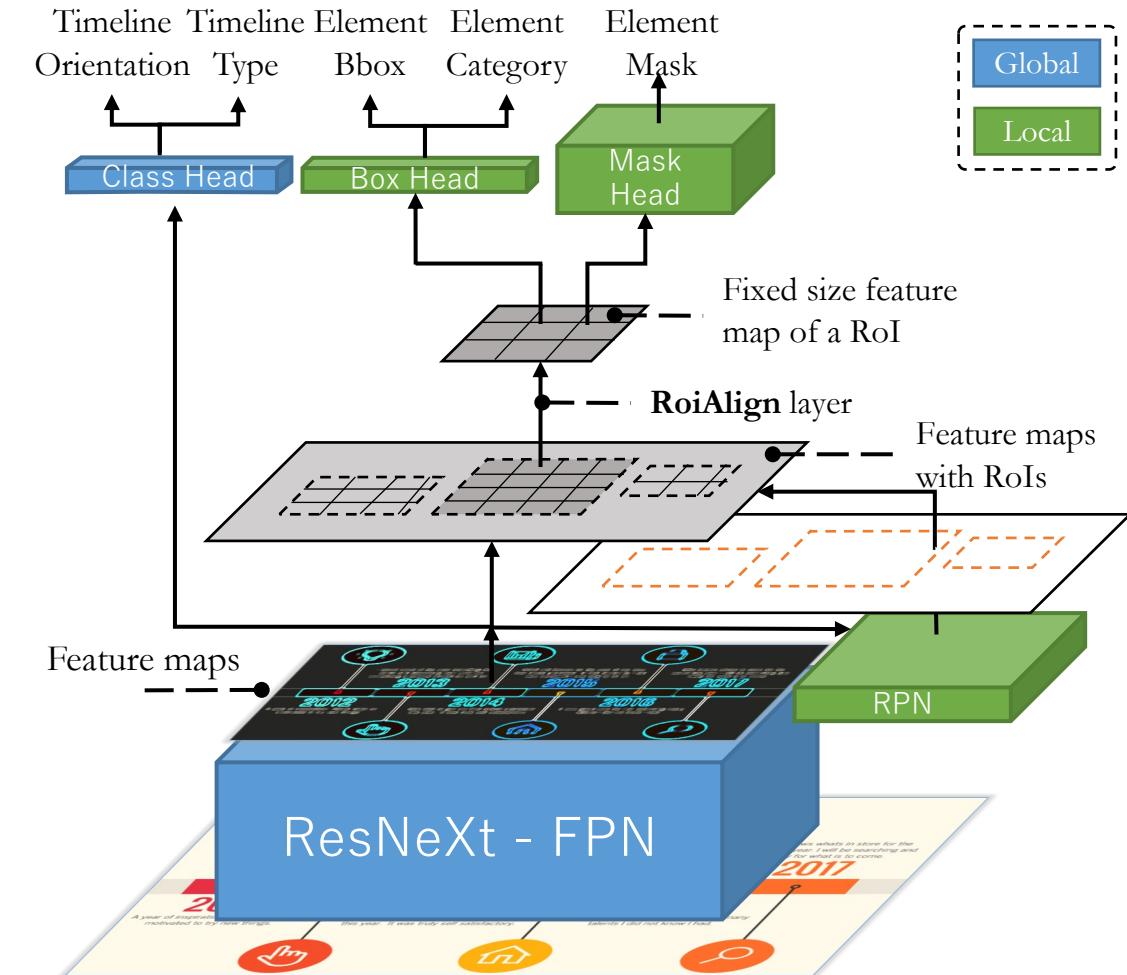
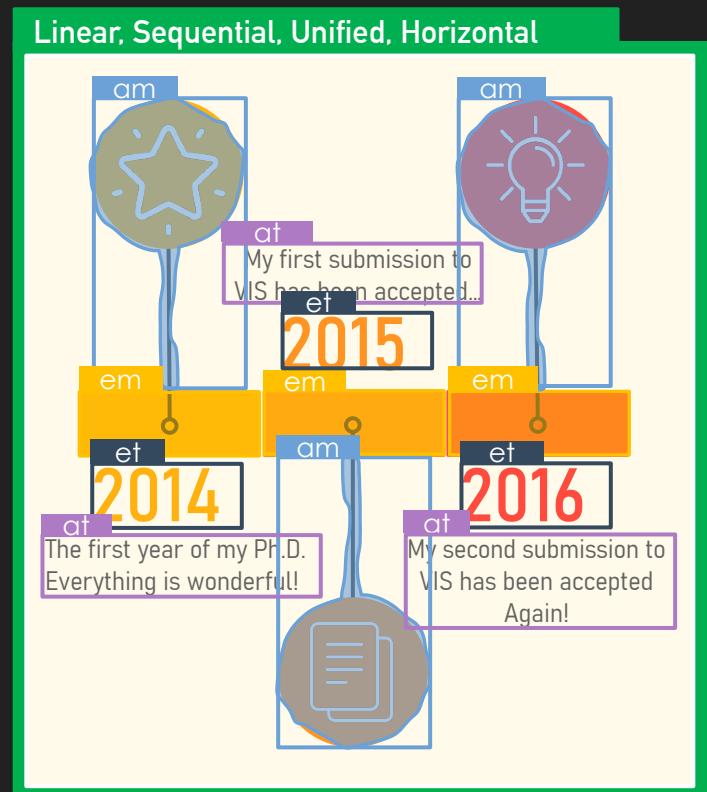
Deconstruction – Local

About the elements:

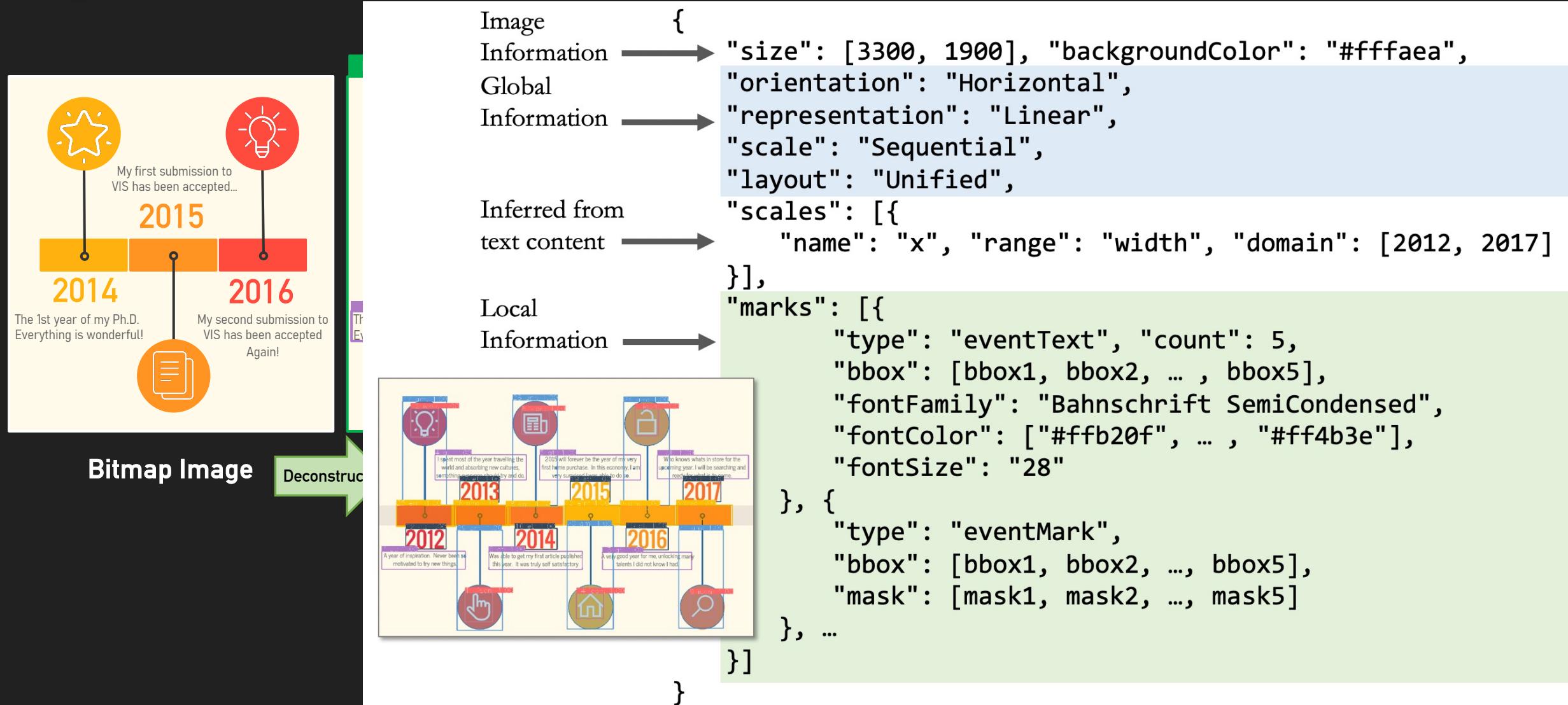
1. Location
2. Category
3. Mask



Deconstruction – Model

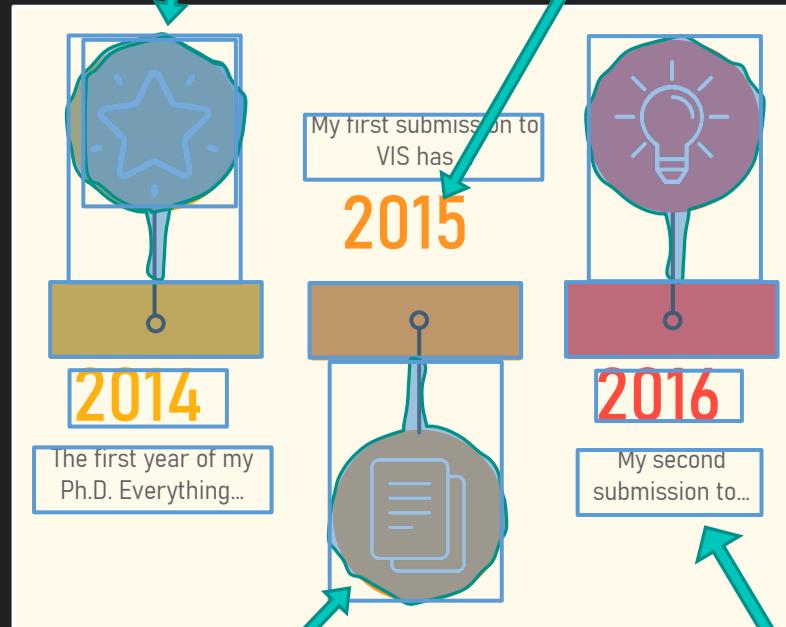


Pipeline – Reconstruction



Reconstruction

Repeated Undetected



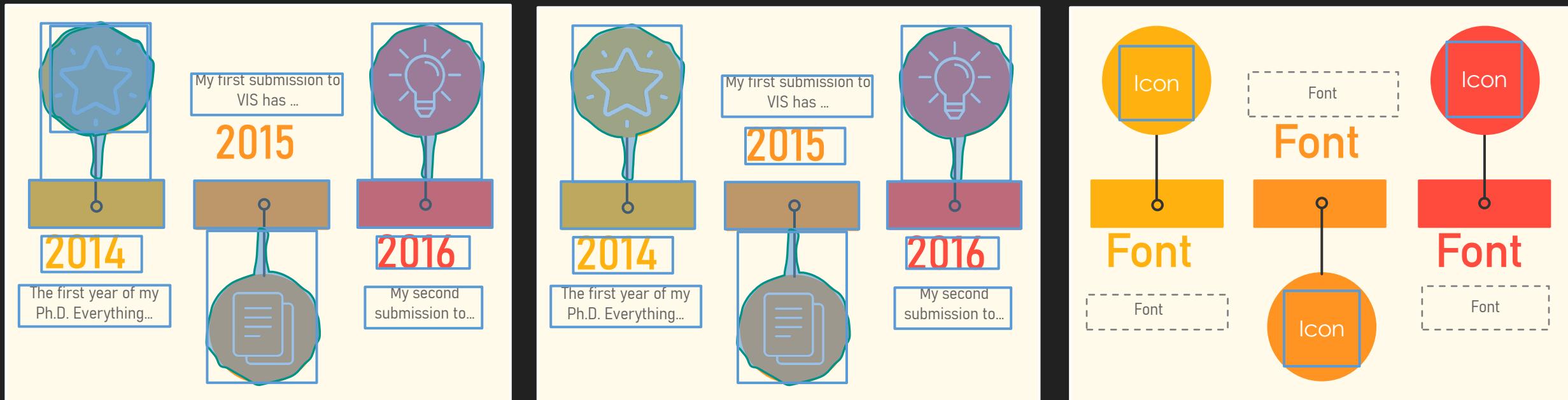
Coarse mask

Unknown font

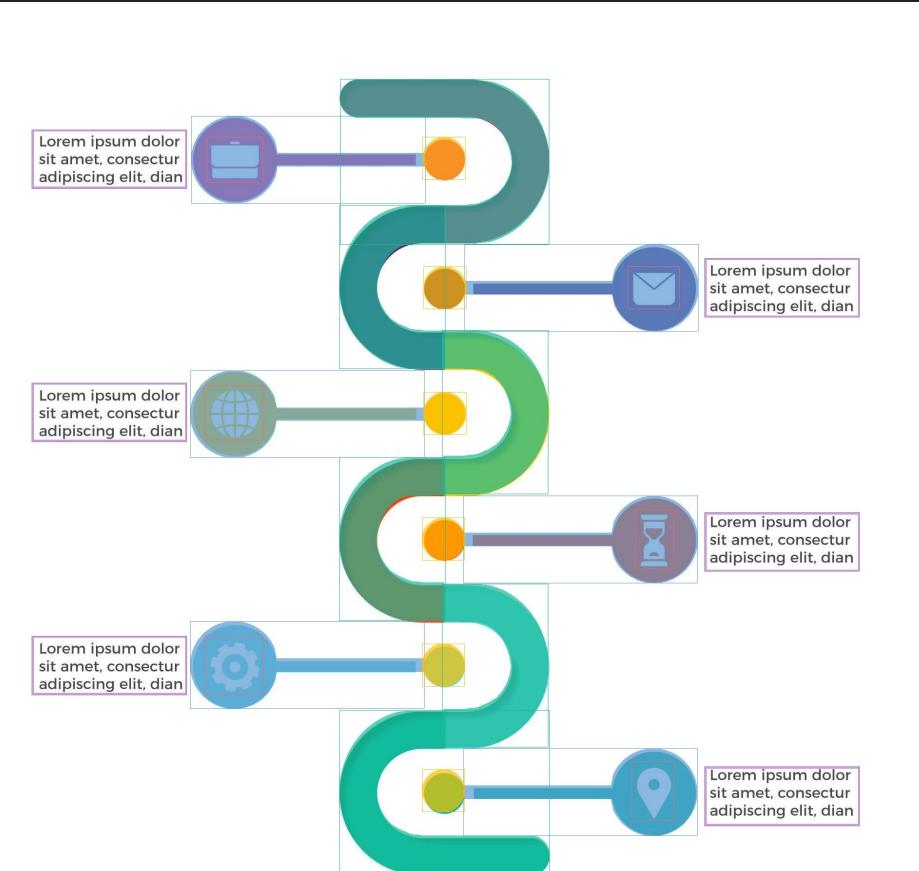
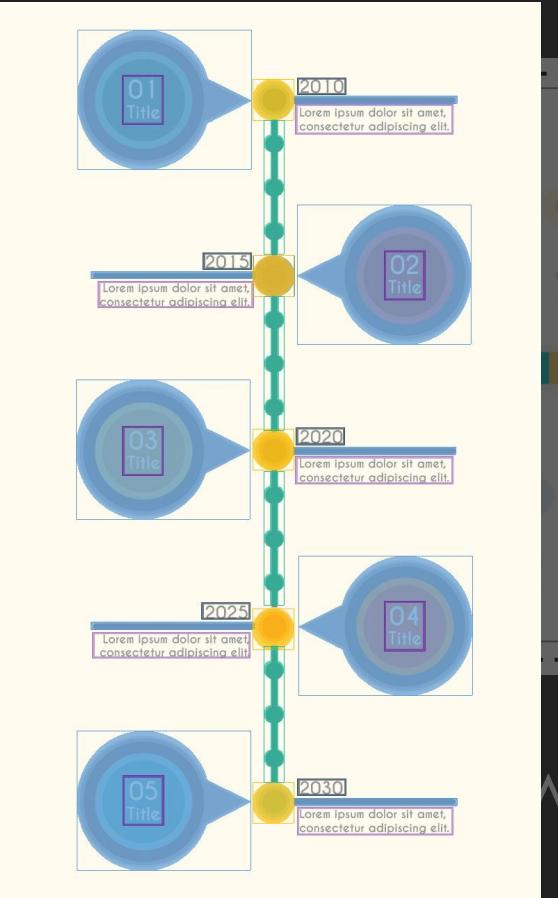
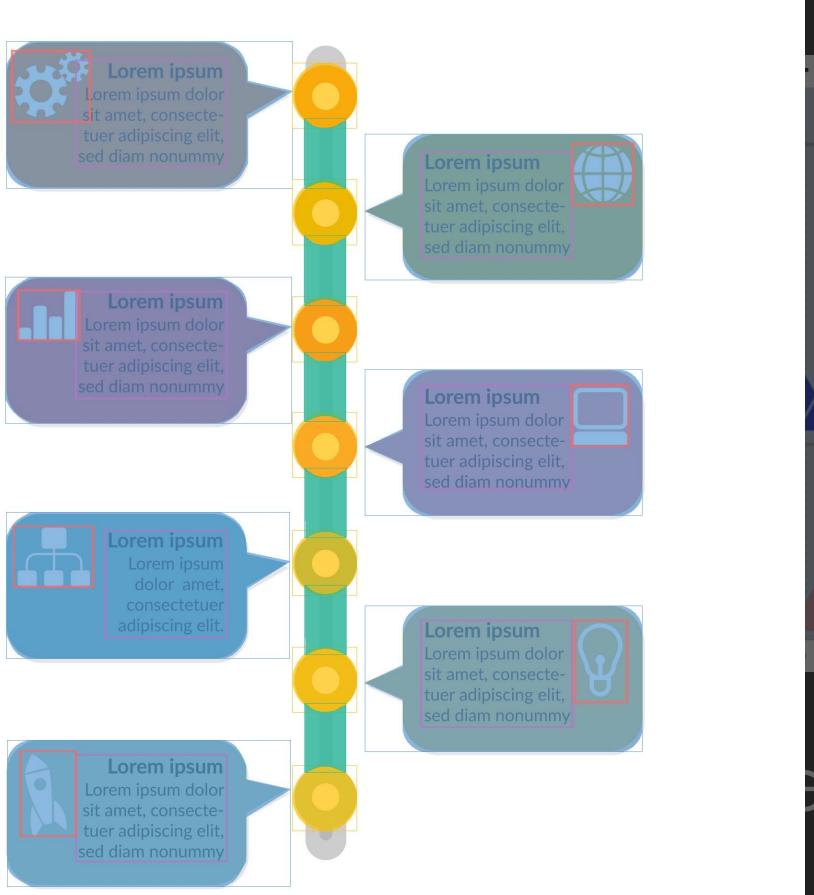
Reconstruction

Repeated
Undetected
Coarse mask

- Non-maximum merging
- Redundancy Recovery
- DL GrabCut



Evaluation – Dataset



TimelineStoryteller: <https://timelinestoryteller.com/>

Evaluation – Deconstruction

Table 3: Classification of timeline types and orientations.

Dataset / Backbone	Type			Orientation		
	Pre.%	Rec.%	F1%	Pre.%	Rec.%	F1%
$D_1 / R50$	99.1	99.1	99.1	100.0	100.0	100.0
$D_1 / R101$	99.5	99.5	99.5	100.0	100.0	100.0
$D_2 / R50$	88.7	86.4	87.5	97.7	97.1	97.4
$D_2 / R101$	92.2	90.9	91.5	97.7	97.1	97.4

Table 4: Average Precision of parsing local information.

Dataset / Backbone	BBox			Mask		
	$AP_{50:95}$	AP_{50}	AP_{75}	$AP_{50:95}$	AP_{50}	AP_{75}
$D_1 / R50$	79.0	93.6	88.0	79.8	96.4	91.6
$D_1 / R101$	81.9	93.9	89.1	79.9	96.9	91.1
$D_2 / R50$	53.4	79.3	61.8	56.9	80.1	61.6
$D_2 / R101$	56.4	81.7	64.9	59.1	82.5	65.1
COCO*	39.8	62.3	43.4	37.1	60.0	39.4

*A state-of-the-art performance on COCO dataset reported by [24].

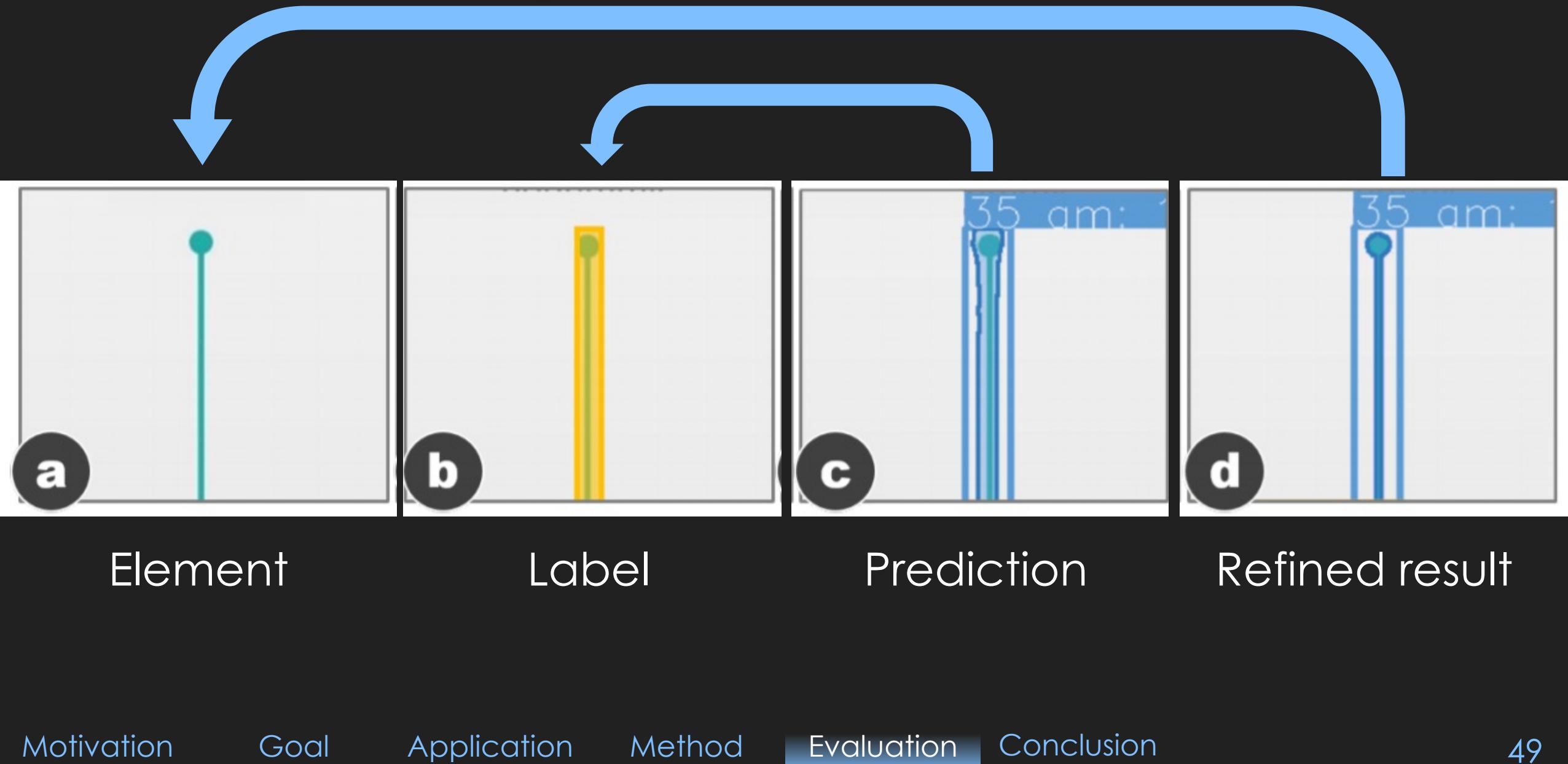
[24] Mask R-CNN. Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick

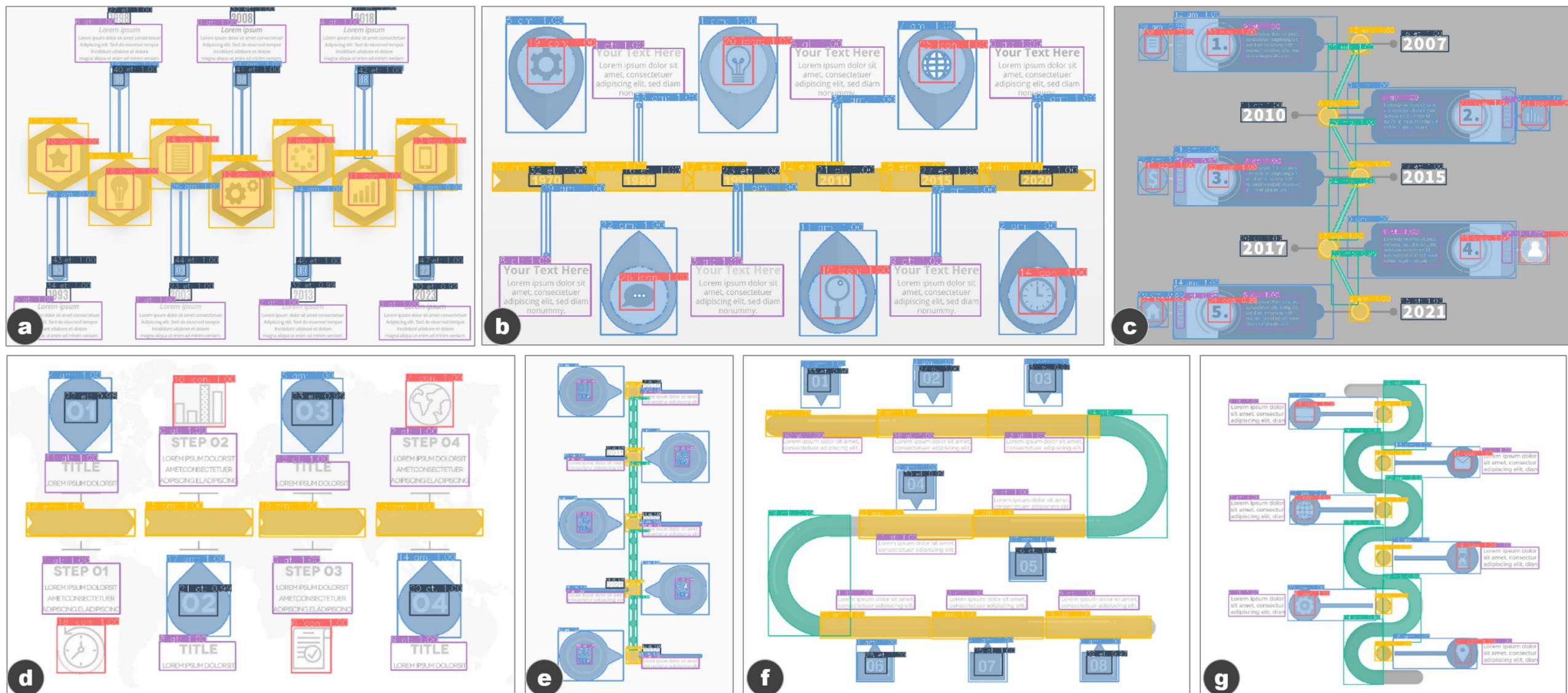
Evaluation – Reconstruction

Table 5: Gains come from Reconstruction at IoU 0.5 and 0.75.

	BBox				Mask			
	Pre_{50}	Rec_{50}	Pre_{75}	Rec_{75}	Pre_{50}	Rec_{50}	Pre_{75}	Pre_{50}
Raw	82.9	80.8	74.0	72.1	85.7	81.5	75.8	72.2
+NMM	+2.3	+1.0	+2.3	+0.8	+1.9	+0.6	+1.8	+0.3
+RR	+1.6	+2.5	+1.6	+2.3	+2.3	+2.1	+2.3	+2.0
+DLGC	0.0	0.0	0.0	0.0	-2.8	-5.5	-4.1	-3.8
Total	86.8	84.1	77.9	75.4	84.0	78.4	75.9	71.0

Evaluation – Imperfect label





Motivation

Goal

Application

Method

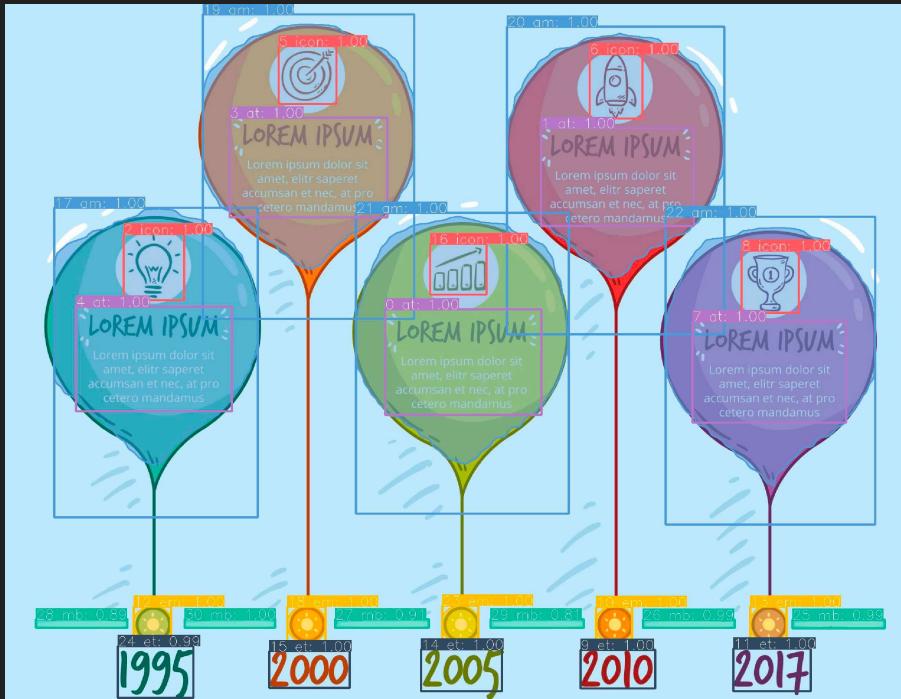
Evaluation

Conclusion

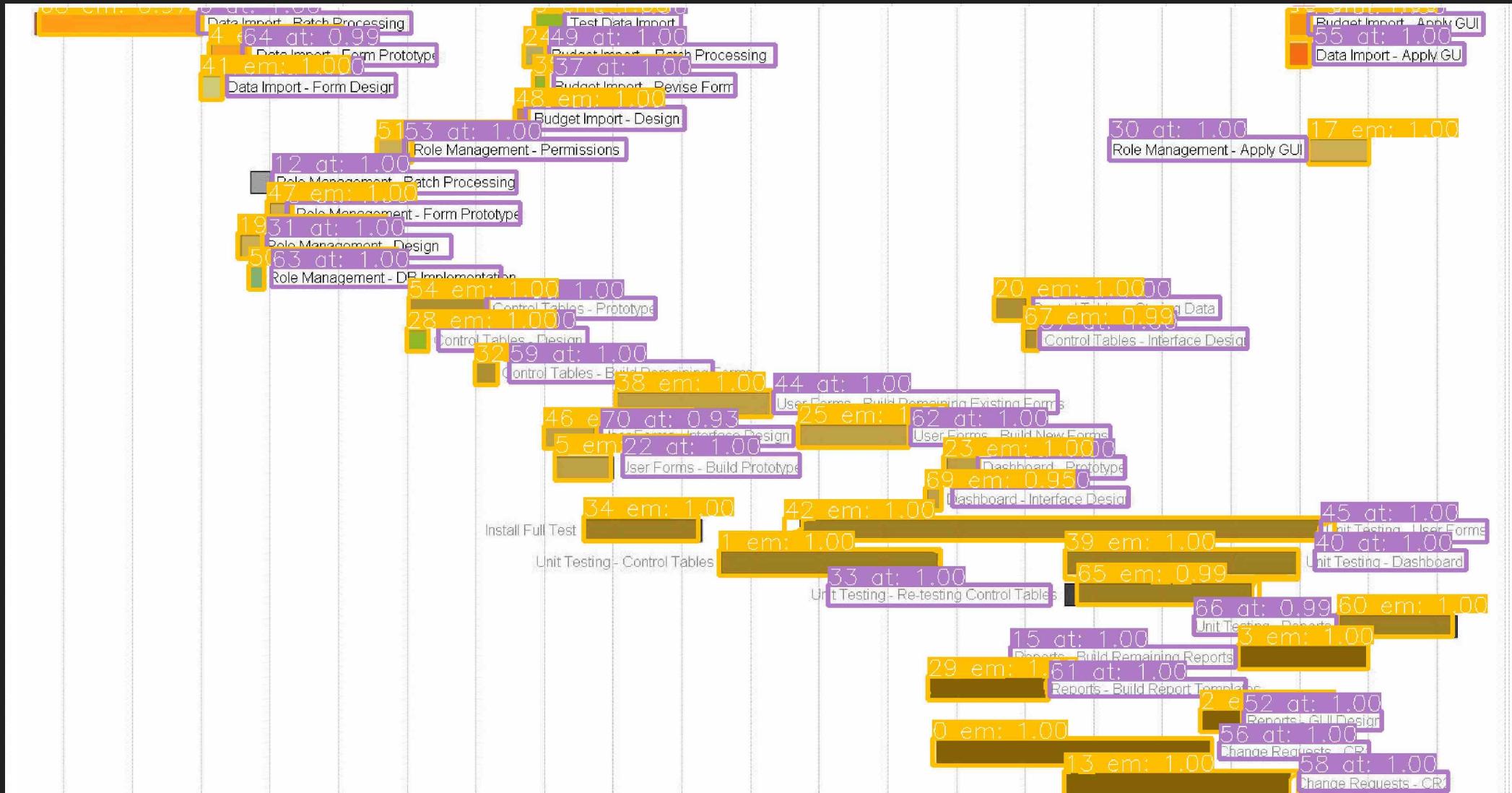
Limitation

Data! Data! Data!

Limitation – Failed Cases

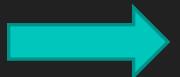
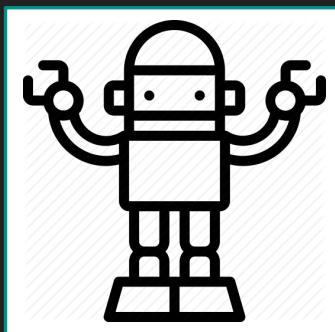
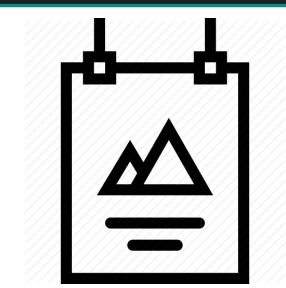
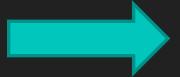
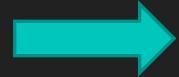
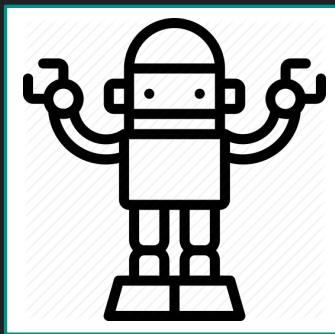


Limitation – Failed Cases



Future work

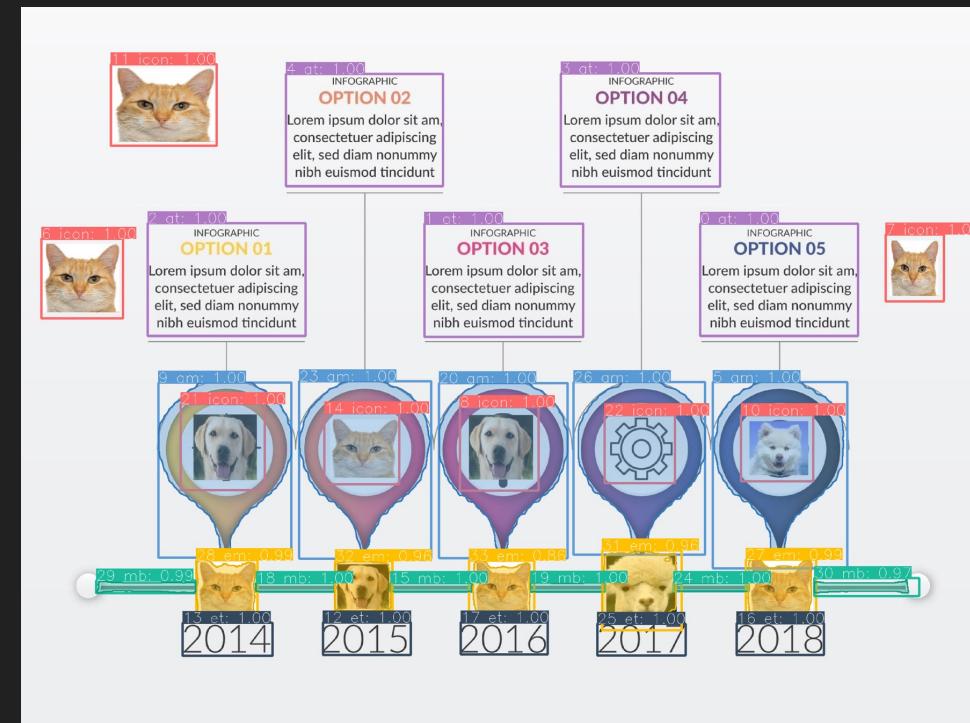
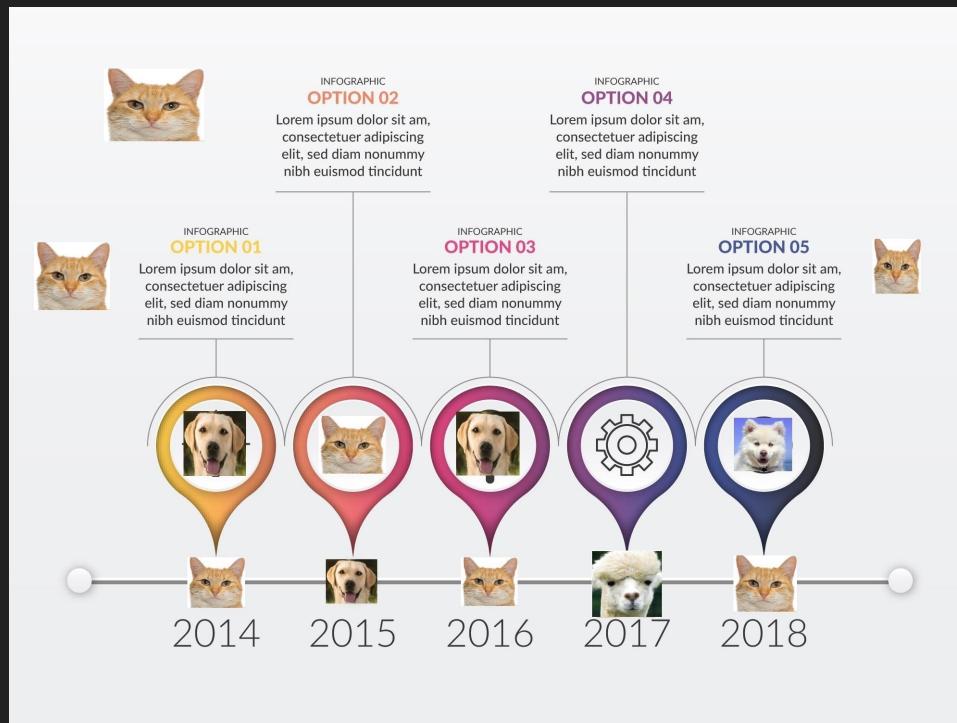
Human-ML Collaborative Authoring tool



Future work

Human-ML Collaborative Authoring tool

Visualization-driven deep learning



Motivation

Goal

Application

Method

Evaluation

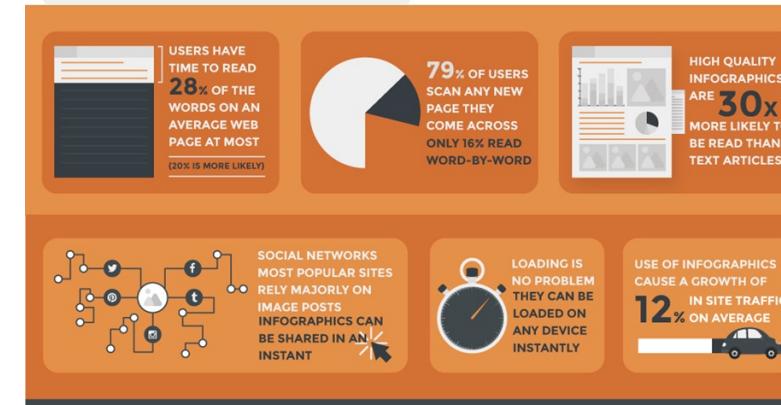
Conclusion

Research Case 2

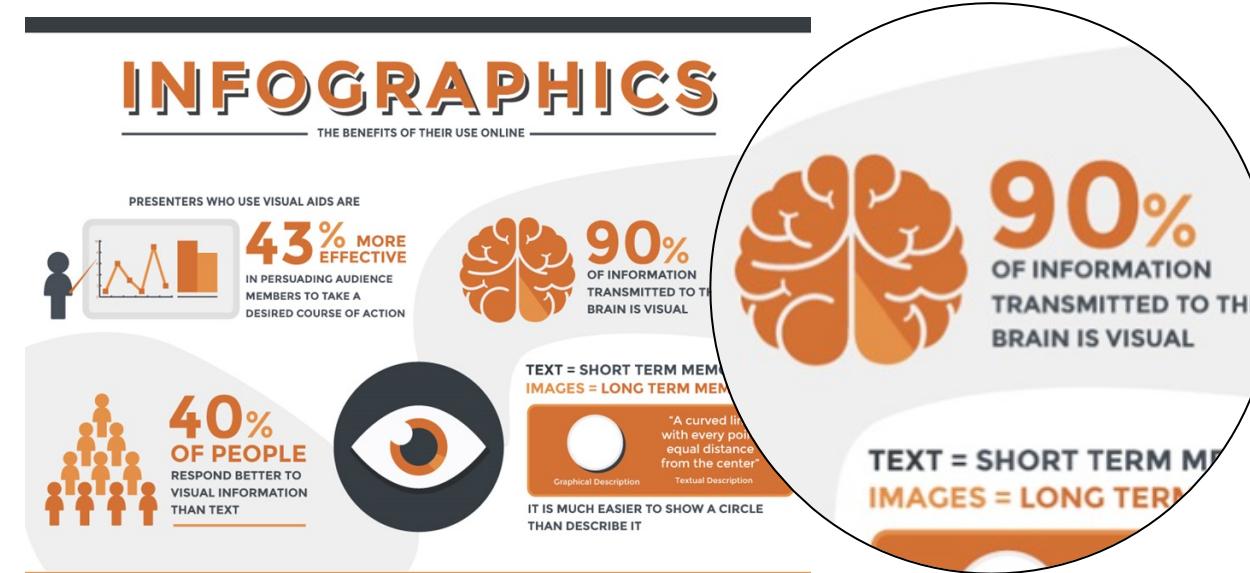
Text-to-Viz: Automatic Generation of Infographics



Fact sheets



Infographic components



Weiwei Cui, Xiaoyu Zhang, Yun Wang, He Huang, Bei Chen, Lei Fang, Haidong Zhang, Jian-Guang Lou, and Dongmei Zhang.
Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements

2024/4/23



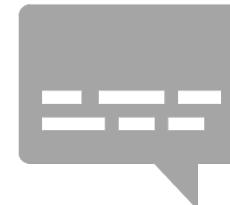
Challenges



Text

How to understand
statements?

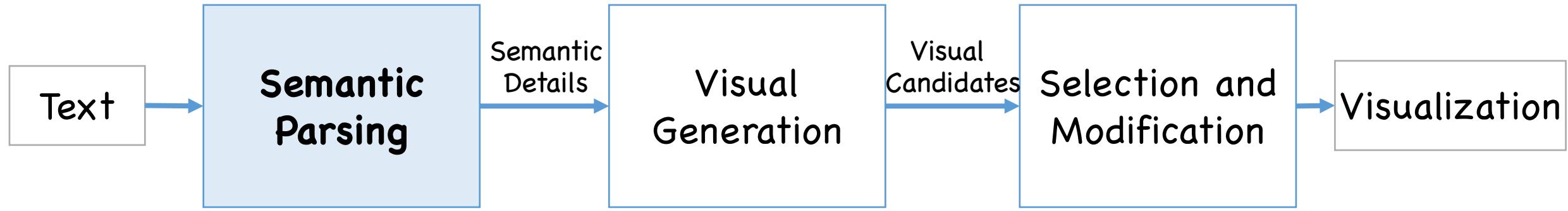
How to extract the



Visualization

How to transform
statements into
visuals?

Text-to-Viz Pipeline





Text Space of Proportion-Related Info

- Understand how people provide information using their daily language
 - 5,562 statements from ~100k PowerPoint slides
 - Using regular expression: n%, m in n, m out of n, half of, etc.
- Three key messages: Number, Whole, and Part
 - E.g., 40% of the students like football best





Semantic Parsing

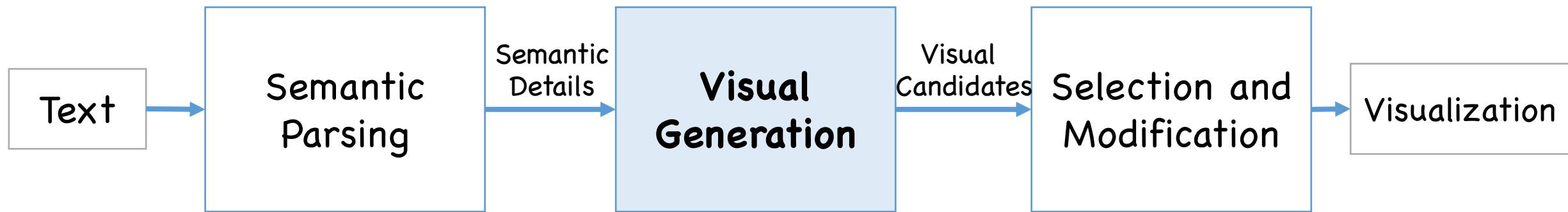
- Sample 800 valid statements
- Annotated key components (number, part, whole, modifier, etc.)
e.g.,

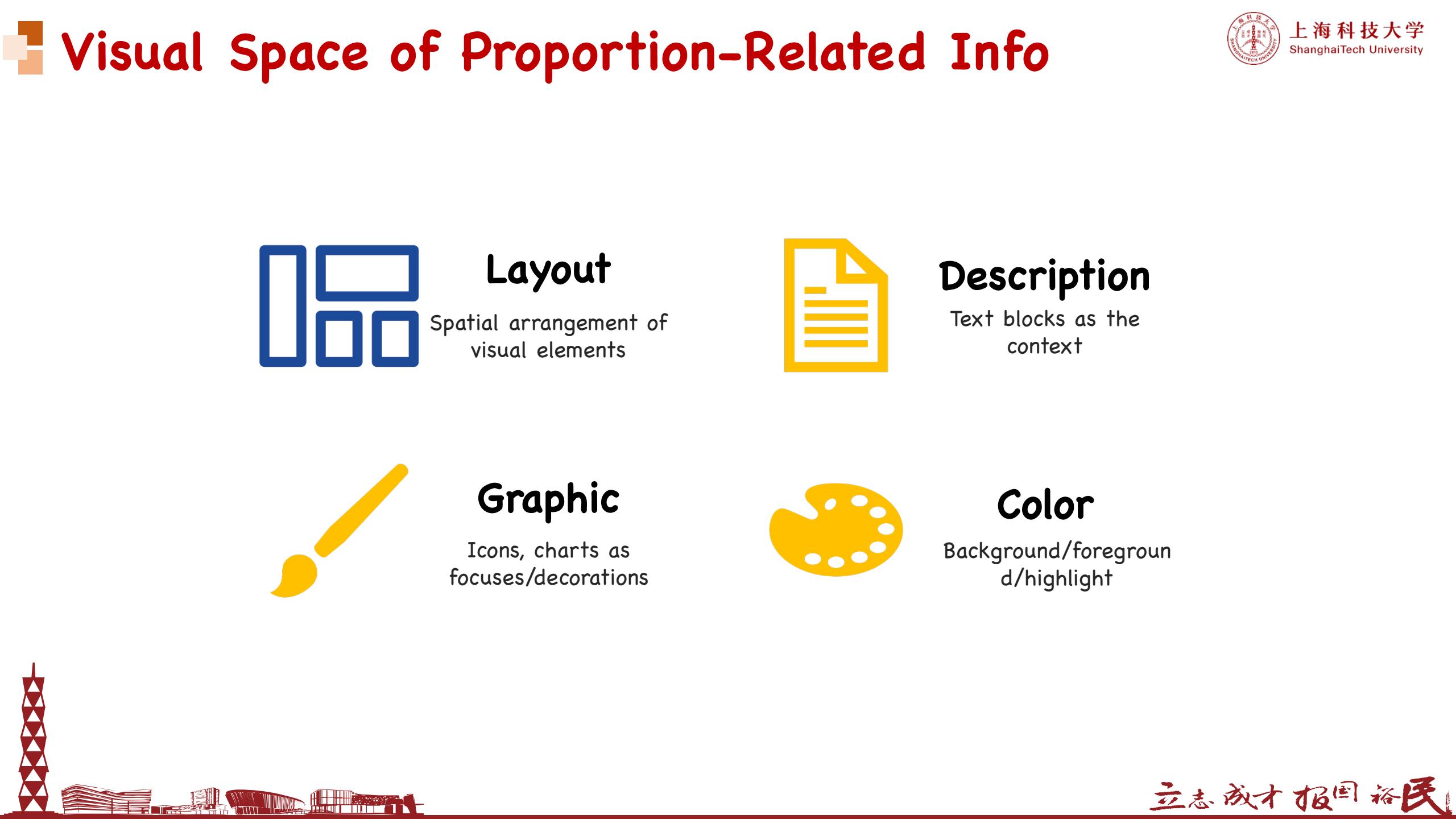


- Three-step machine-learning model
 - Tokenization -> Featurization -> CNN+CRF layer



Text-to-Viz Pipeline

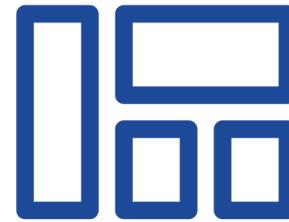




Visual Space of Proportion-Related Info



上海科技大学
ShanghaiTech University



Layout

Spatial arrangement of visual elements



Description

Text blocks as the context



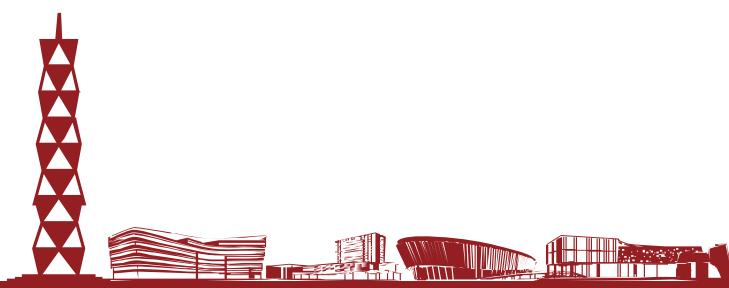
Graphic

Icons, charts as focuses/decorations



Color

Background/foreground/highlight

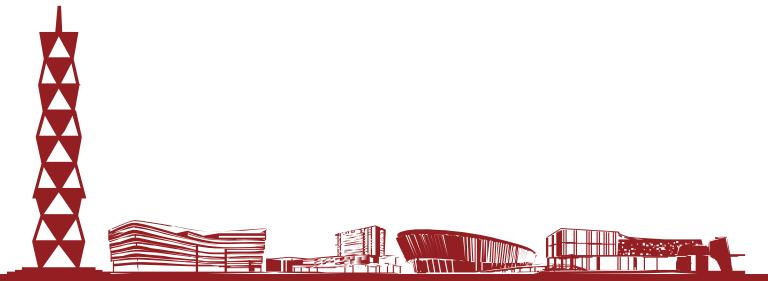


立志成才报国裕民

Visual Generation



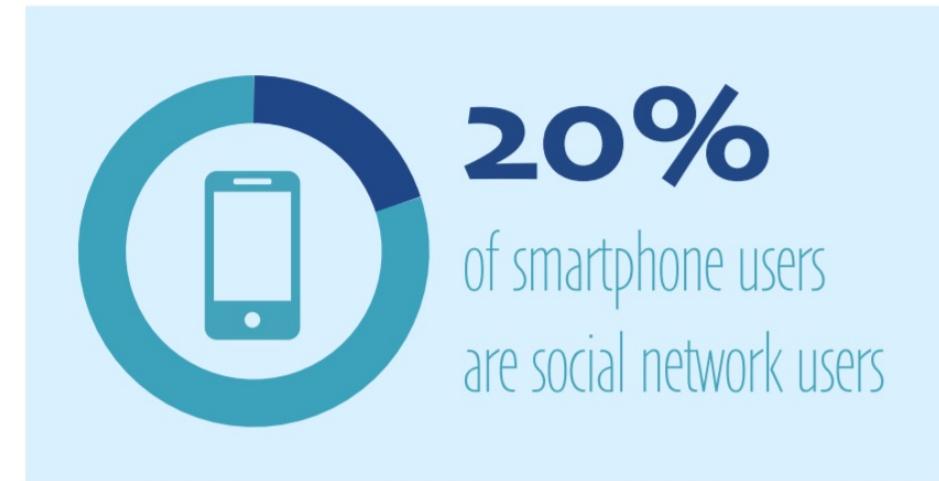
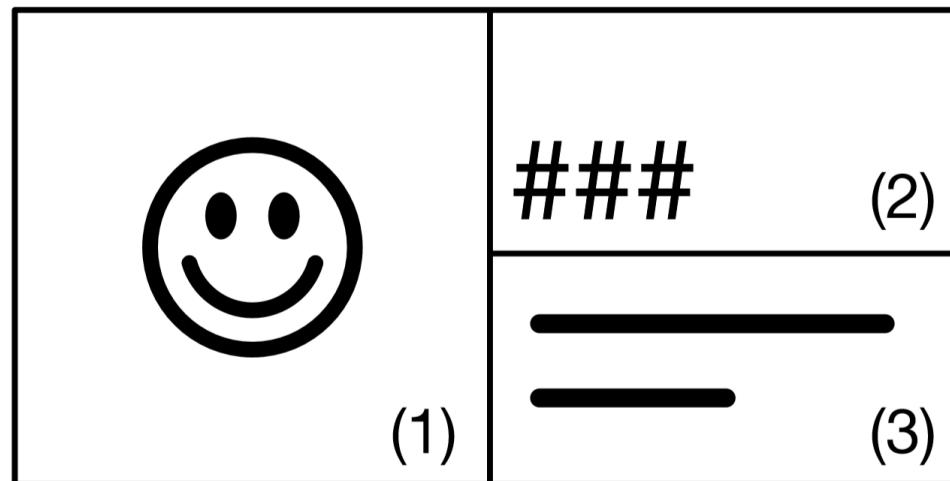
- Step 1: Identify candidates on each design dimension
- Step 2: Enumerate **combinations** to identify valid infographics
- Step 3: **Rank** to find good ones for users to choose or further refine



Layout Candidates

A set of blueprints describing the overall look of resulting infographics

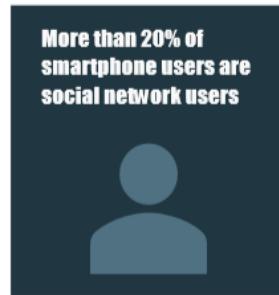
- The aspect ratio of the canvas, how it is divided into regions
- The attribute constraints of visual elements that can be placed in each region
- E.g.,



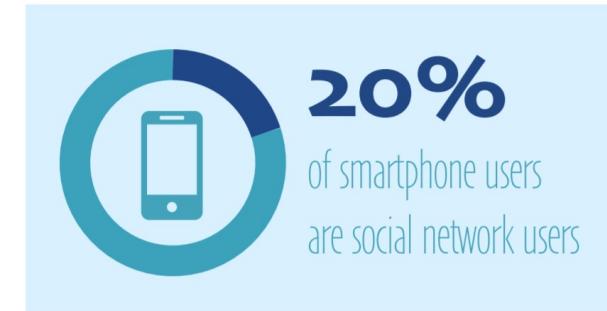


Description Candidates

E.g., “More than 20% of smartphone users are social network users.”



More than
20% of smartphone users
are social network users



Snippets of different information and lengths for different blueprints

E.g., Number: 20%, Whole: “of smartphone users”, “smartphone users”, Part: “are social network users”, Modifier: “more than”, After-number: “of smartphone users are social network users”, etc.



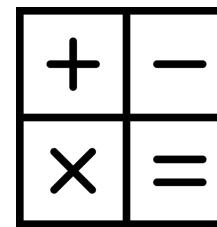
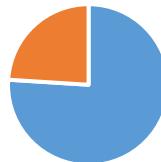


Graphic Candidates



e.g., “More than 76% of students find math difficult.”

- Identify graphic elements from different components
 - Number: 76%
 - Whole: “of students”
 - Part: “find math difficult”



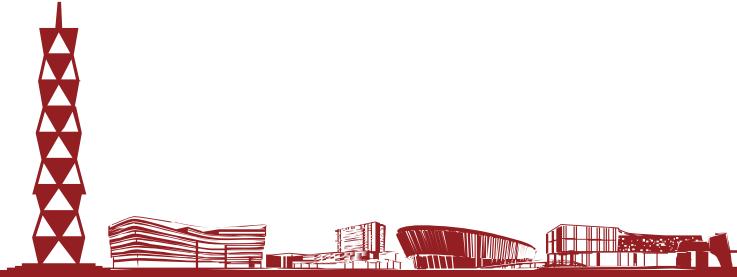
Color Palette Candidates

Considerations

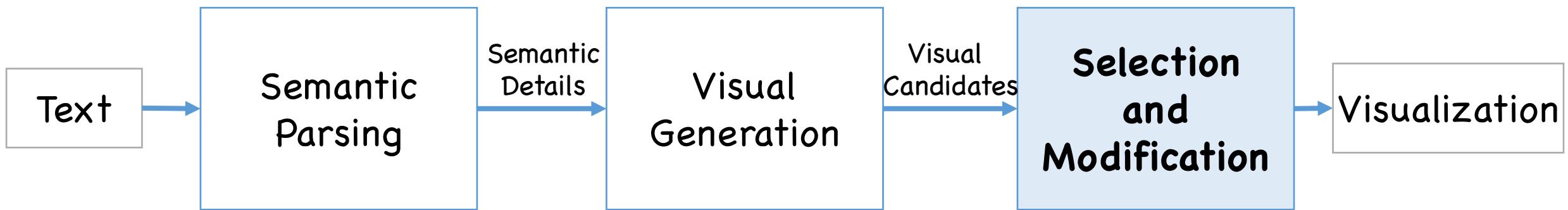
- Harmonic inside a palette
- The semantic of statements

Solution

- Build a library of color palettes with descriptive keywords
- Word2Vec to find best-matching palette



Text-to-Viz Pipeline



Rank the Synthesized Infographics



Semantic

The semantic effectiveness of selected icons and color palette



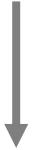
Informative



The completeness of information delivered visually

Visual

The visual appearance of the generated infographic

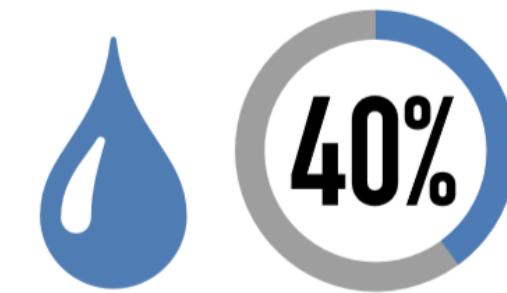
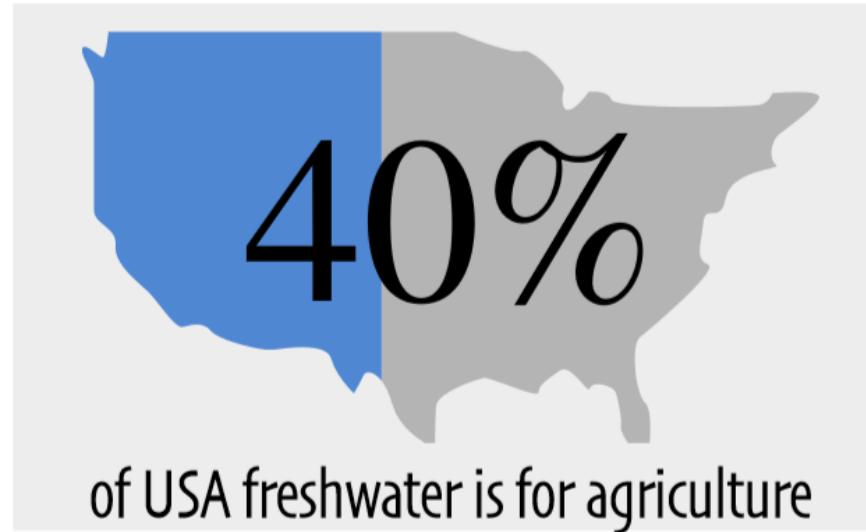


Score = Word2vec scores + Ratio of non-stop words + Ratio of the area



Sample Examples

- “40% of USA freshwater is for agriculture”



Sample Examples

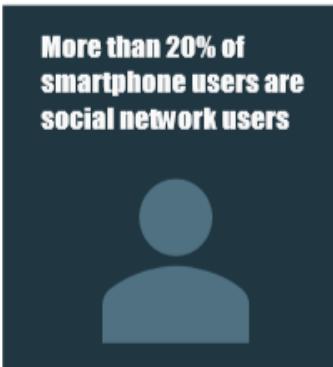
- “More than 20% of smartphone users are social network users”



More than
20% of smartphone users
are social network users



More than **20%** of
smartphone users are
social network users



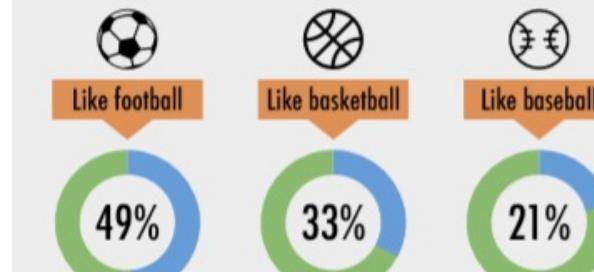
Sample Examples

- “Among all students, 49% like football, 32% like basketball, and 21% like baseball”

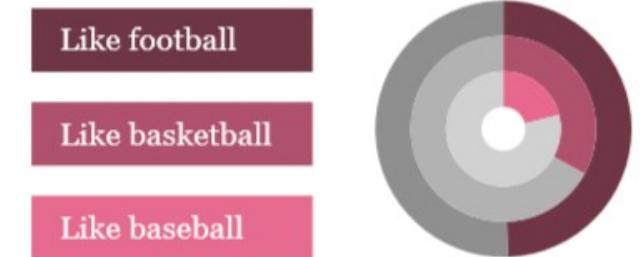
AMONG ALL STUDENTS

Like football	49%	
Like basketball	33%	
Like baseball	21%	

Among all students

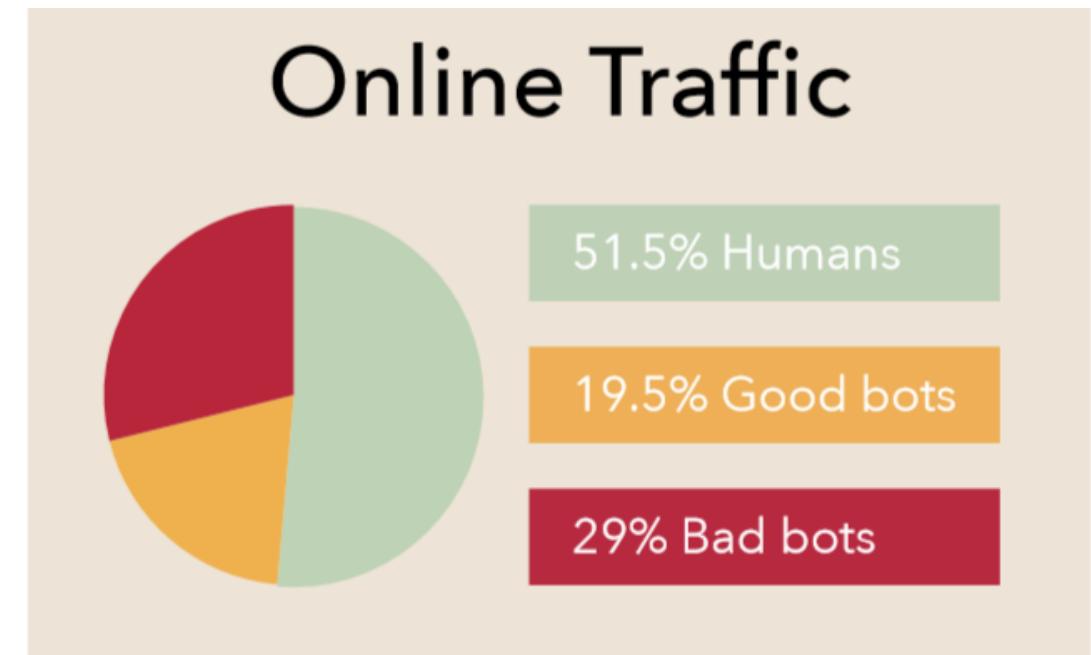
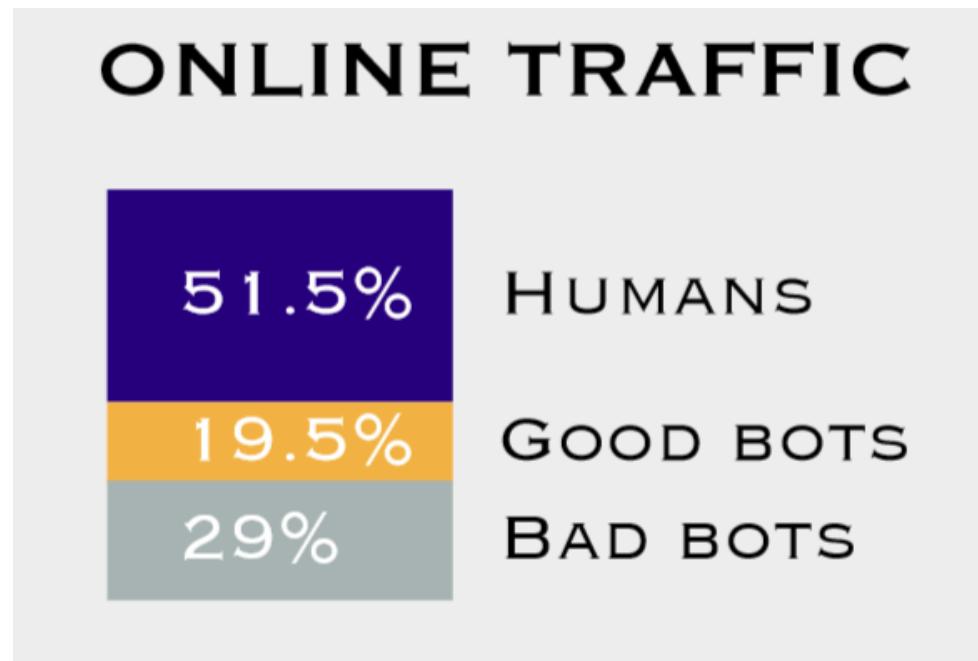


AMONG ALL STUDENTS



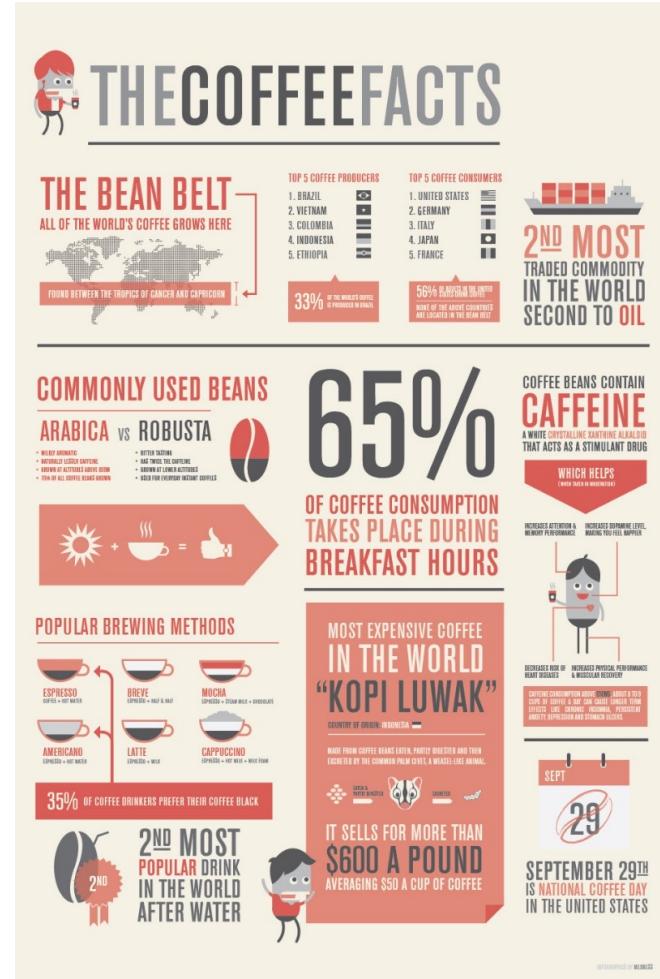
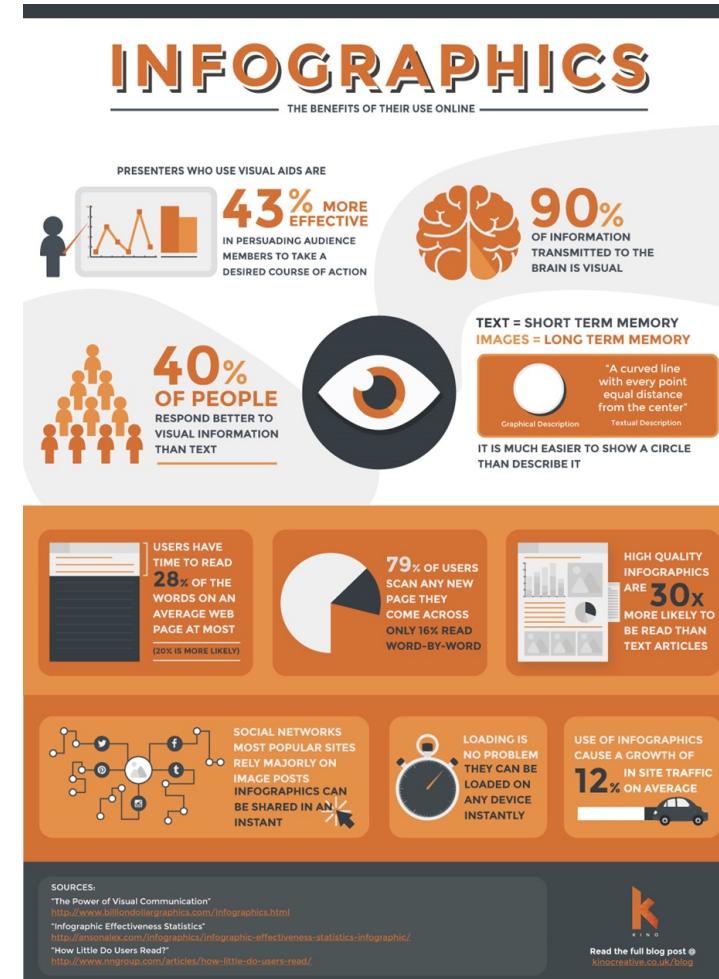
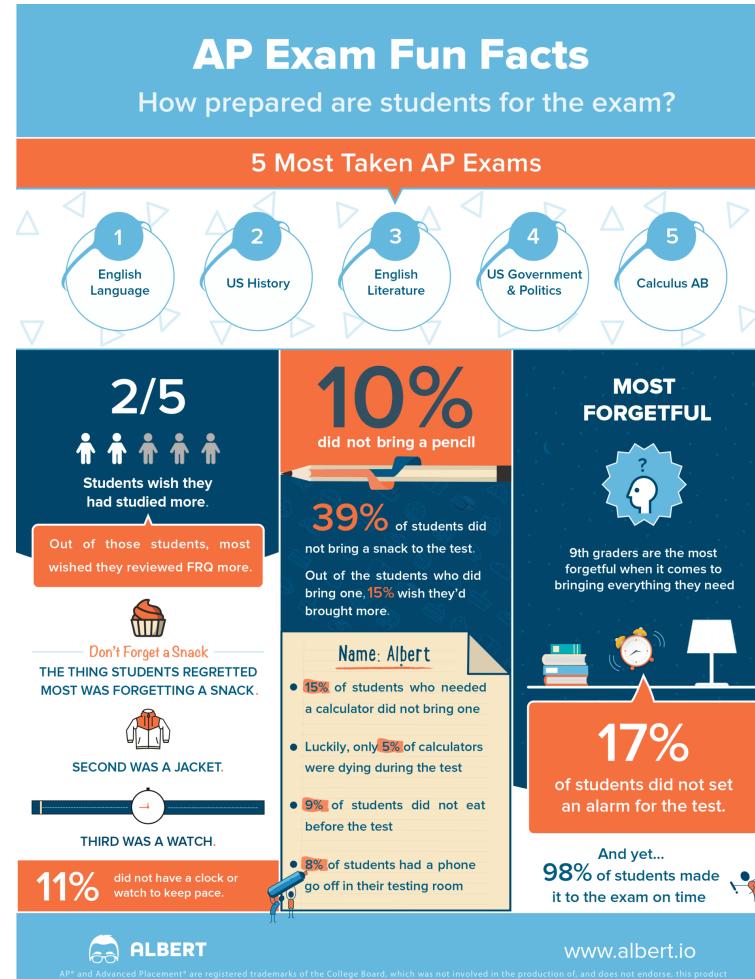
Sample Examples

- “Humans made 51.5% of online traffic, while good bots made 19.5% and bad bots made 29%”



Research Case 3

Automatic Generation of Fact Sheets



Fact Sheets

Data facts

- Numerical or statistical results

Visualizations and Infographics

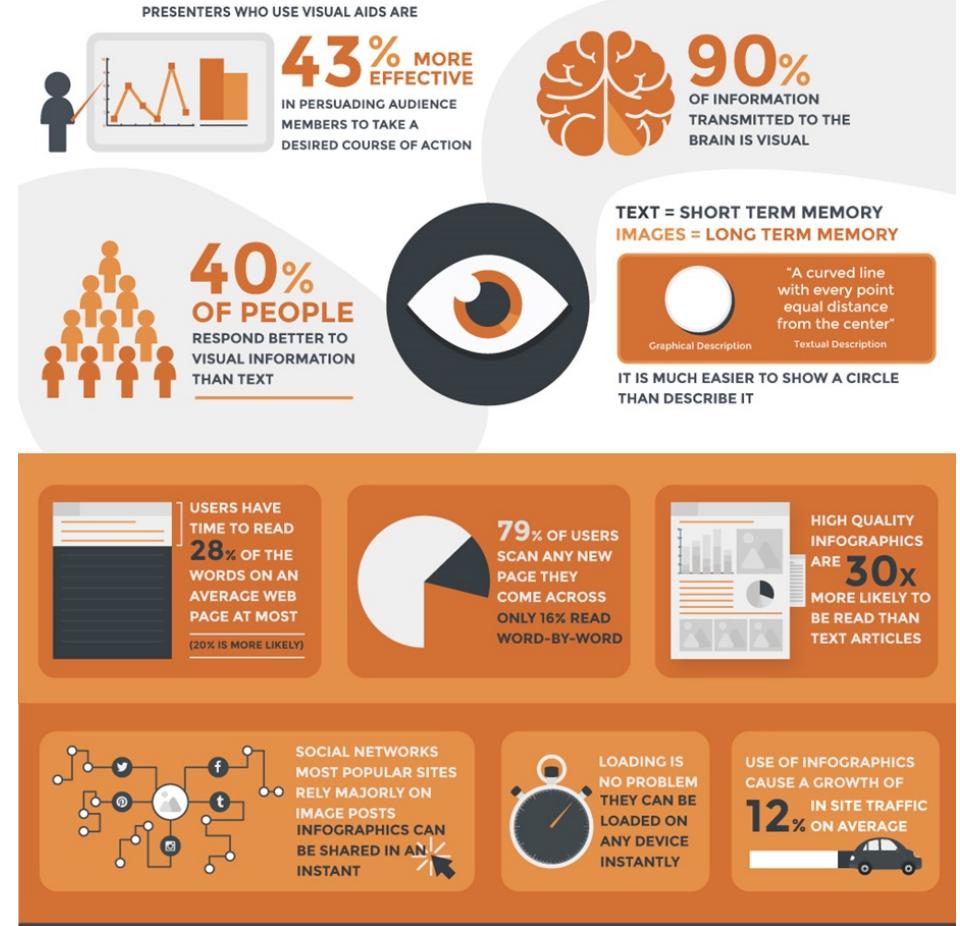
- Illustrate data facts

Composition of information

- Ease the process of absorbing information

INFOGRAPHICS

THE BENEFITS OF THEIR USE ONLINE

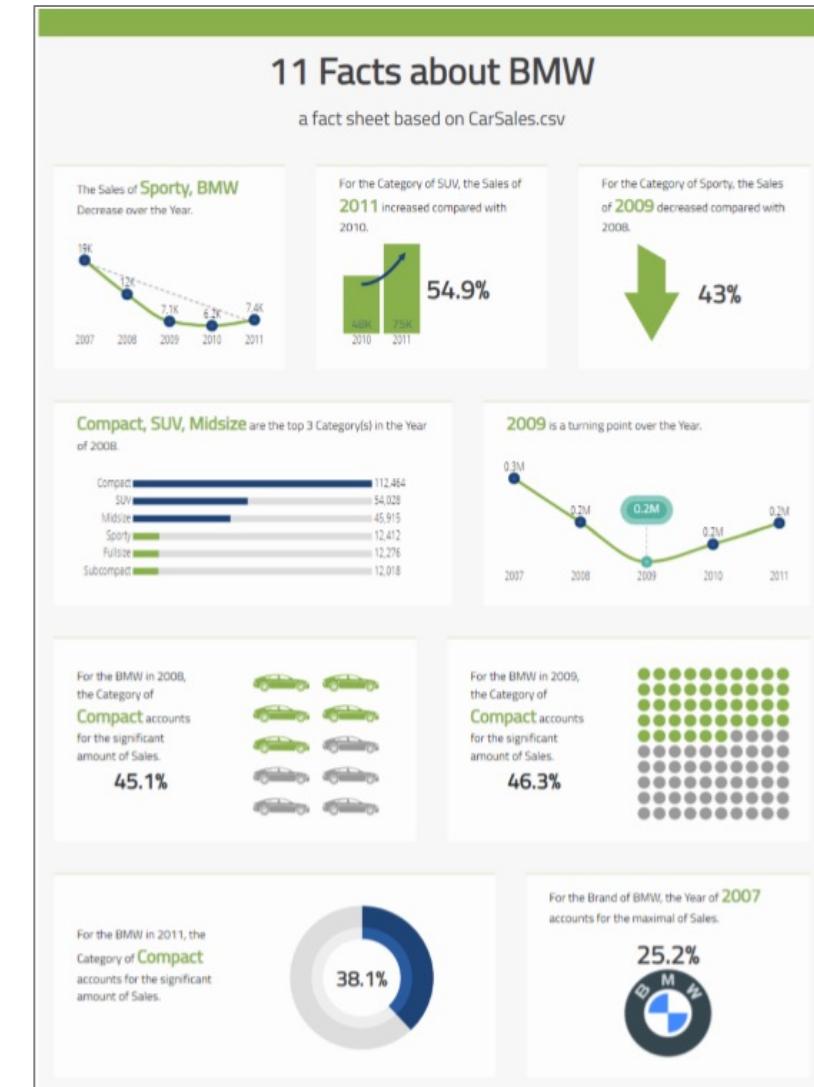




Data Table

A	B	C	D	E	F
1	Year	Brand	Category	Model	Sales
2	1/1/2007	BMW	Compact	BMW 3-Se	142490
3	1/1/2008	BMW	Compact	BMW 3-Se	112464
4	1/1/2009	BMW	Compact	BMW 3-Se	90960
5	1/1/2010	BMW	Compact	BMW 3-Se	100910
6	1/1/2011	BMW	Compact	BMW 3-Se	94371
7	1/1/2007	BMW	Compact	BMW 5-Se	54142
8	1/1/2007	BMW	Fullsize	BMW 7-Se	14773
9	1/1/2008	BMW	Fullsize	BMW 7-Se	12276
10	1/1/2009	BMW	Fullsize	BMW 7-Se	9254
11	1/1/2010	BMW	Fullsize	BMW 7-Se	12253
12	1/1/2011	BMW	Fullsize	BMW 7-Se	11299
13	1/1/2008	BMW	Midsize	BMW 5-Se	45915
14	1/1/2009	BMW	Midsize	BMW 5-Se	40109
15	1/1/2010	BMW	Midsize	BMW 5-Se	39488
16	1/1/2011	BMW	Midsize	BMW 5-Se	51491
17	1/1/2007	BMW	Sporty	BMW 6-Se	9033
18	1/1/2008	BMW	Sporty	BMW 6-Se	6533
19	1/1/2009	BMW	Sporty	BMW 6-Se	3549
20	1/1/2010	BMW	Sporty	BMW 6-Se	2418
21	1/1/2011	BMW	Sporty	BMW 6-Se	3903
22	1/1/2007	BMW	Sporty	BMW Z4	10097
23	1/1/2008	BMW	Sporty	BMW Z4	5879
24	1/1/2009	BMW	Sporty	BMW Z4	3523
25	1/1/2010	BMW	Sporty	BMW Z4	3804

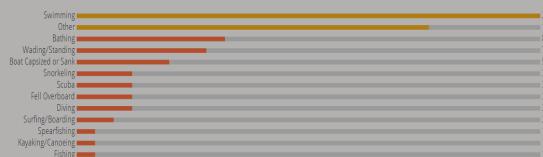
Fact Sheet



9 Facts about Swimming

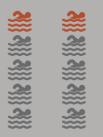
a fact sheet based on SharkAttacks.csv

Swimming, Other are the top 2 Activity(s) in the Gender of Female.



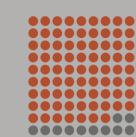
For the Gender of Female, the Activity of **Swimming** accounts for the maximal of Attacks.

24.1%



For the Activity of Swimming, the Gender of **Male** accounts for the maximal of Attacks.

86.7%



For the Fatal in Female, the Activity of **Swimming** accounts for the significant amount of Attacks.

For the Not Fatal in Female, the Activity of **Swimming** accounts for the significant amount of Attacks.

22.7%



For the Activity of Swimming, the Fatal of **Not Fatal** accounts for the maximal of Attacks.

62.5%



For the Fatal in Male, the Activity of **Swimming** accounts for the maximal of Attacks.

33.8%



Fatal, Swimming Attacks by Gender(s).

305

25

Male

Female

For all Activity, the Activity of **Swimming** accounts for the maximal of Attacks.

78.3%



11 Facts about Sporty

a fact sheet based on CarSales.csv

Sporty Sales by Year(s).

0.2M
0.1M
87K
85K
81K

2007
2008
2009
2010
2011

The total number of **Brand(s)** included in the Category of Sporty

3

Sporty Sales by Brand(s).

0.4M
52K
47K

Ford
BMW
Volkswagen

For the Category of **Sporty**, the Brand(s) account for the maximal of Sales.

81.5%

The Sales of Sporty, BMW

Decrease over the Year.

19K
17K
15K
13K
12K

2007
2008
2009
2010
2011

The Sales of **Sporty, Ford**

Decrease over the Year.

0.1M
0.11M
0.12M
0.13M
0.14M

2007
2008
2009
2010
2011

The Sales of **Sporty, Volkswagen**

Decrease over the Year.

13K
13K
12K
11K
10K

2007
2008
2009
2010
2011

For the Category of **Sporty**, the Sales of **2009** decreased compared with **2008**.

-43.0%

2007
2008
2009
2010

Sporty, BMW Sales by Year(s).

14K
12K
7.4K
7.1K
6.2K

2007
2008
2009
2010
2011

For the Category of **Sporty**, the Sales of **2009** decreased compared with **2008**.

-43.0%

2007
2008
2009
2010
2011

Sporty, BMW Sales by Year(s).

14K
12K
7.4K
7.1K
6.2K

2007
2008
2009
2010
2011

For the Sporty in 2008, the Brand of **Ford** accounts for the maximal of Sales.

78.3%

For the Sporty in 2007, the Brand of **Ford** accounts for the maximal of Sales.

80.9%

For the Sporty in 2007, the Brand of **Ford** accounts for the maximal of Sales.

80.9%

11 Facts about BMW

a fact sheet based on CarSales.csv

The Sales of **Sporty, BMW**

Decrease over the Year.

15K
13K
11K
9K
7K

2008
2009
2010
2011

For the Category of SUV, the Sales of **2011** increased compared with **2010**.

54.9%

For the Category of Sporty, the Sales of **2009** decreased compared with **2008**.

-43%

2007
2008
2009
2010
2011

For the BMW in 2008, the Category of **Compact** accounts for the significant amount of Sales.

45.1%

For the BMW in 2009, the Category of **Compact** accounts for the significant amount of Sales.

46.3%

For the BMW in 2011, the Category of **Compact** accounts for the maximal of Sales.

38.1%

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

25.2%

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the Brand of BMW, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the BMW in 2011, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
2010
2011

For the SUV, the Brand of **BMW** accounts for the maximal of Sales.

75K
66K
54K
48K
38K

2007
2008
2009
2010
2011

For the BMW in 2007, the Category of **Compact** accounts for the maximal of Sales.

48.8%

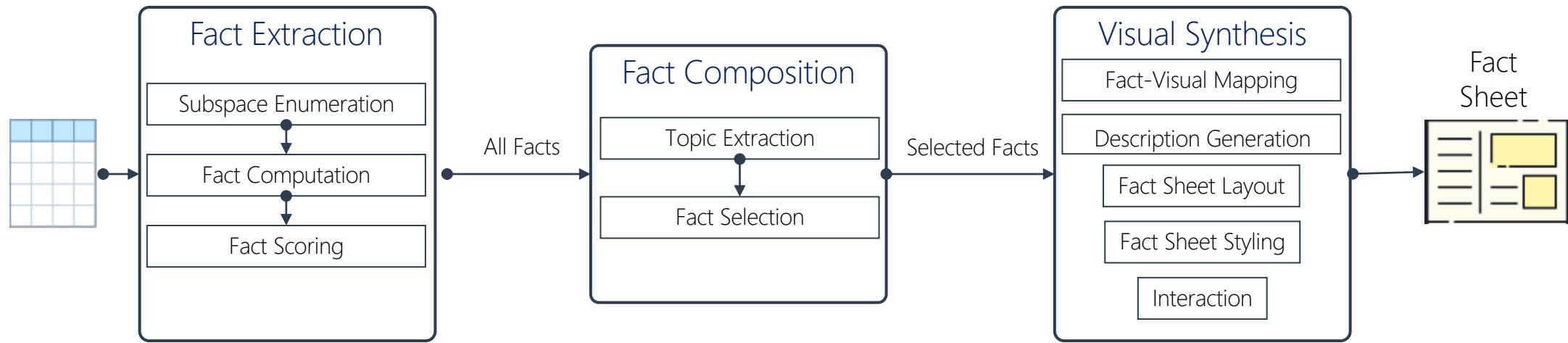
2007
2008
2009
2010
2011

For the Brand of BMW, the Year of **2007** accounts for the maximal of Sales.

25.2%

2007
2008
2009
20

System Pipeline



Data Input

Story Type

Story Elements

DataShot Home About

Data Input

CarSales.csv Browse

Story Type

- 10 Facts about BMW
- 10 Facts about Compact
- 10 Facts about Ford
- 10 Facts about GMC
- 10 Facts about Honda
- 10 Facts about Hyundai
- 10 Facts about Midsized
- 10 Facts about Pickup**
- 10 Facts about Smart

Story Elements

Potential Story Elements

- Proportion (Category): Pickup in All
- Proportion (Category): Pickup in 2007
- Proportion (Category): Pickup in 2008
- Proportion (Category): Pickup in 2009
- Proportion (Category): Pickup in 2010
- Proportion (Category): Pickup in 2011
- Categorization (Year): Pickup
- Proportion (Year): 2007 in Pickup
- Proportion (Year): 2008 in Pickup
- Proportion (Year): 2011 in Pickup
- Proportion (Year): 2010 in Pickup
- Proportion (Year): 2009 in Pickup
- Trend (Year): Pickup
- Extreme (Year): 2009 in Pickup

Selected Story Elements

- Trend (Year): Pickup,GMC
- Difference (Year): 2009 in Pickup,GMC
- Extreme (Year): 2009 in Pickup
- Proportion (Brand): GMC in Pickup
- Proportion (Brand): GMC in 2007,Pickup
- Value (Year): Pickup
- Value (Brand): Pickup

Generate New Story

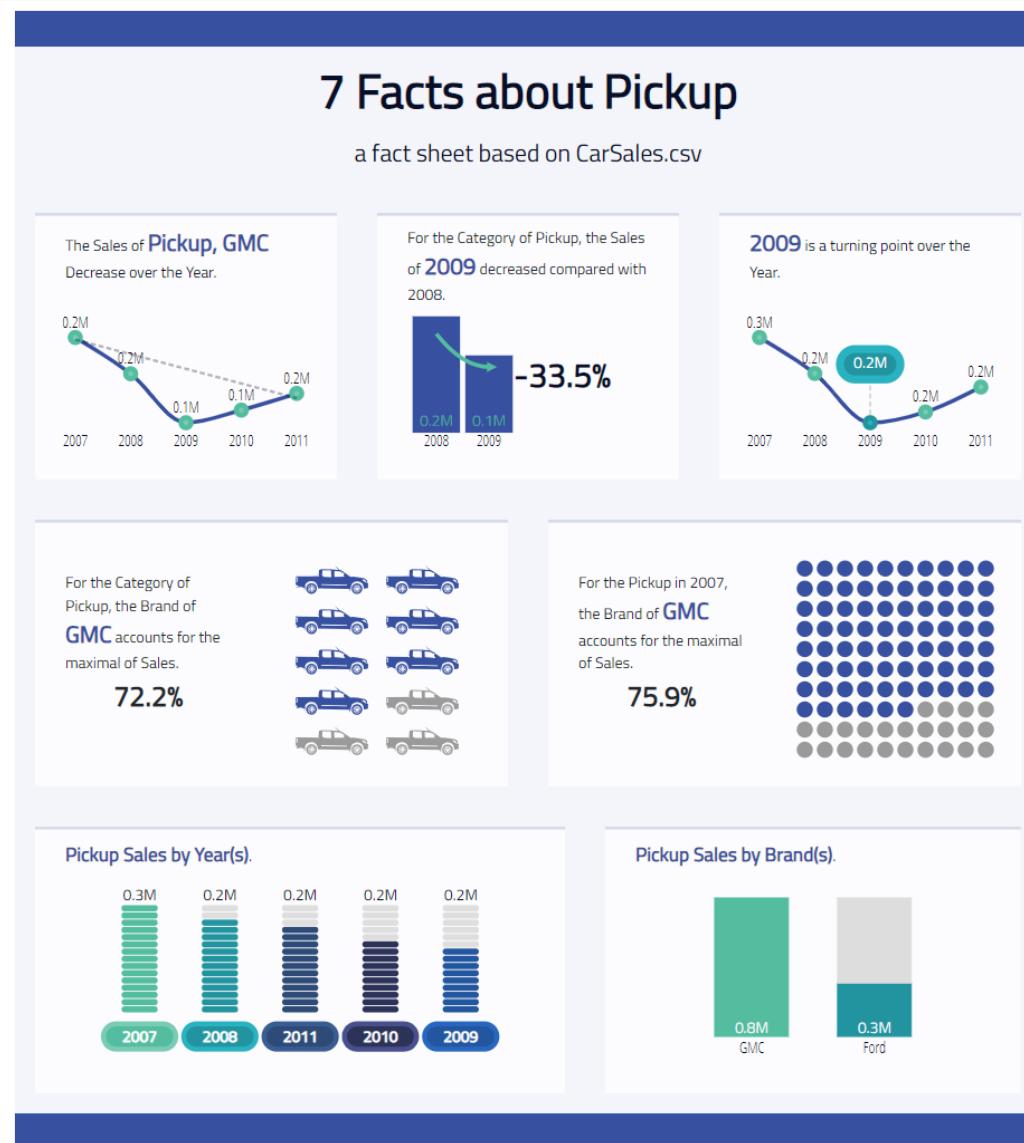
Theme

- Cool Jade
- Mountain Breeze
- Sea Sunset
- Warm Afternoon
- Ultra Violet
- Living Coral
- Technology**

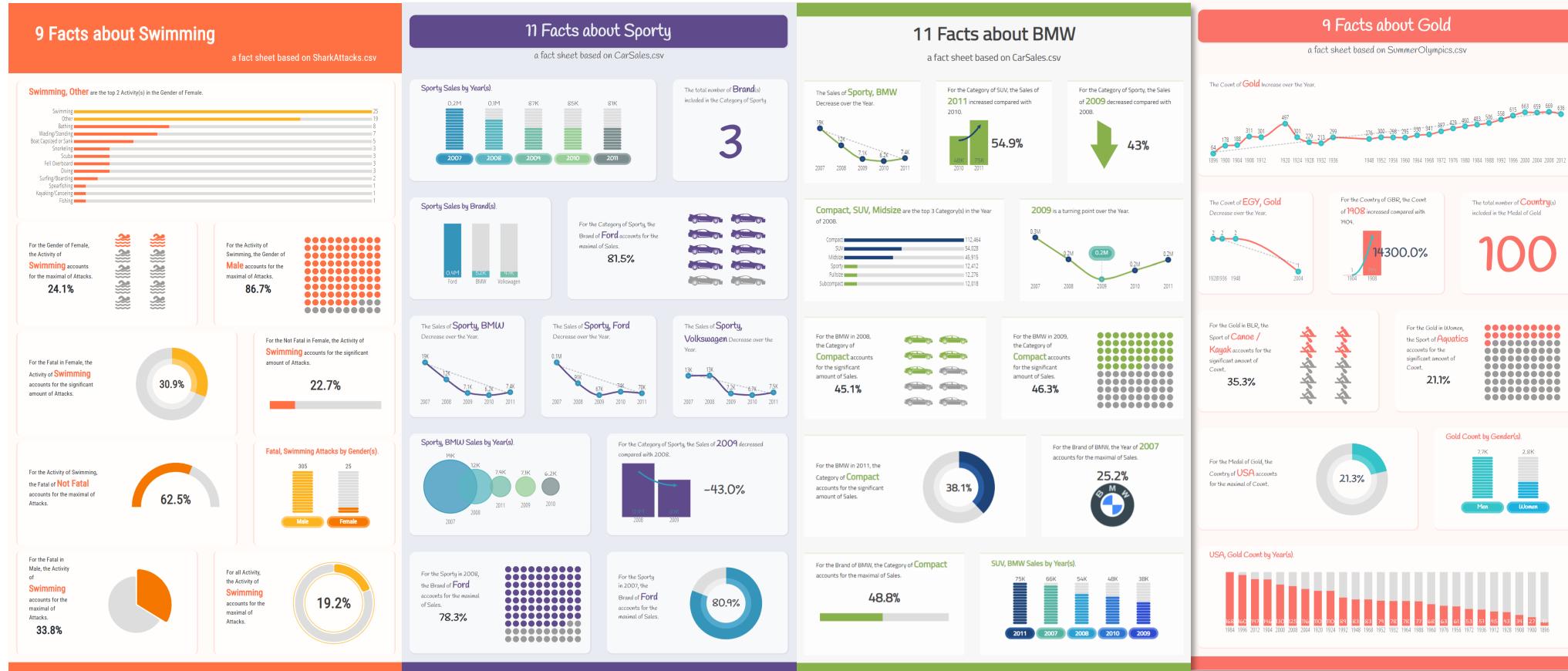
Style

- Paper
- Brick
- Brackets

Export



Example Results





Summary

- Good visualizations are accumulating rapidly
 - Diverse, creative, and professional
 - Clear patterns: alignment, color, font, etc.
- Understand and reuse existing designs
 - Design ideas: How to present information effectively
 - Design details: How to present information aesthetically
- Challenges
 - Understand: Extract semantic models from examples
 - Reuse: Find appropriate examples to reuse

Take-home Message

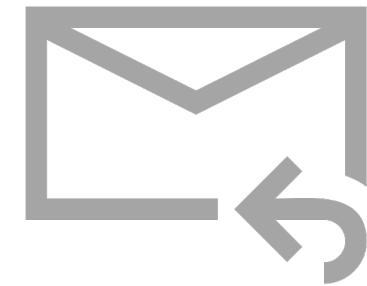
- Visualization data is emerging and stored in different formats including graphics, programs, and hybrid.
- Currently AI4VIS research primarily focused on 7 questions.
- Representation of visualization data is a challenging issue since visualization data is multi-modal, data-specific, susceptible-to-local-detail and construed with domain knowledge.
- It is therefore demanding to develop ML models that are more tailored to visualization data

<https://ai4vis.github.io>



Quan Li

Questions?
Thank you 😊



[liquan@shanghaitech
.edu.cn](mailto:liquan@shanghaitech.edu.cn)