Readable Image for the Visually Impaired

EEL-806: Computer Vision
Remote Guest Lecture

SunilKumar.Kopparapu@TCS.COM

TCS Innovation Labs - Mumbai

Tata Consultancy Services Ltd, Yantra Park,
Thane (West), Maharastra 400601,
INDIA.

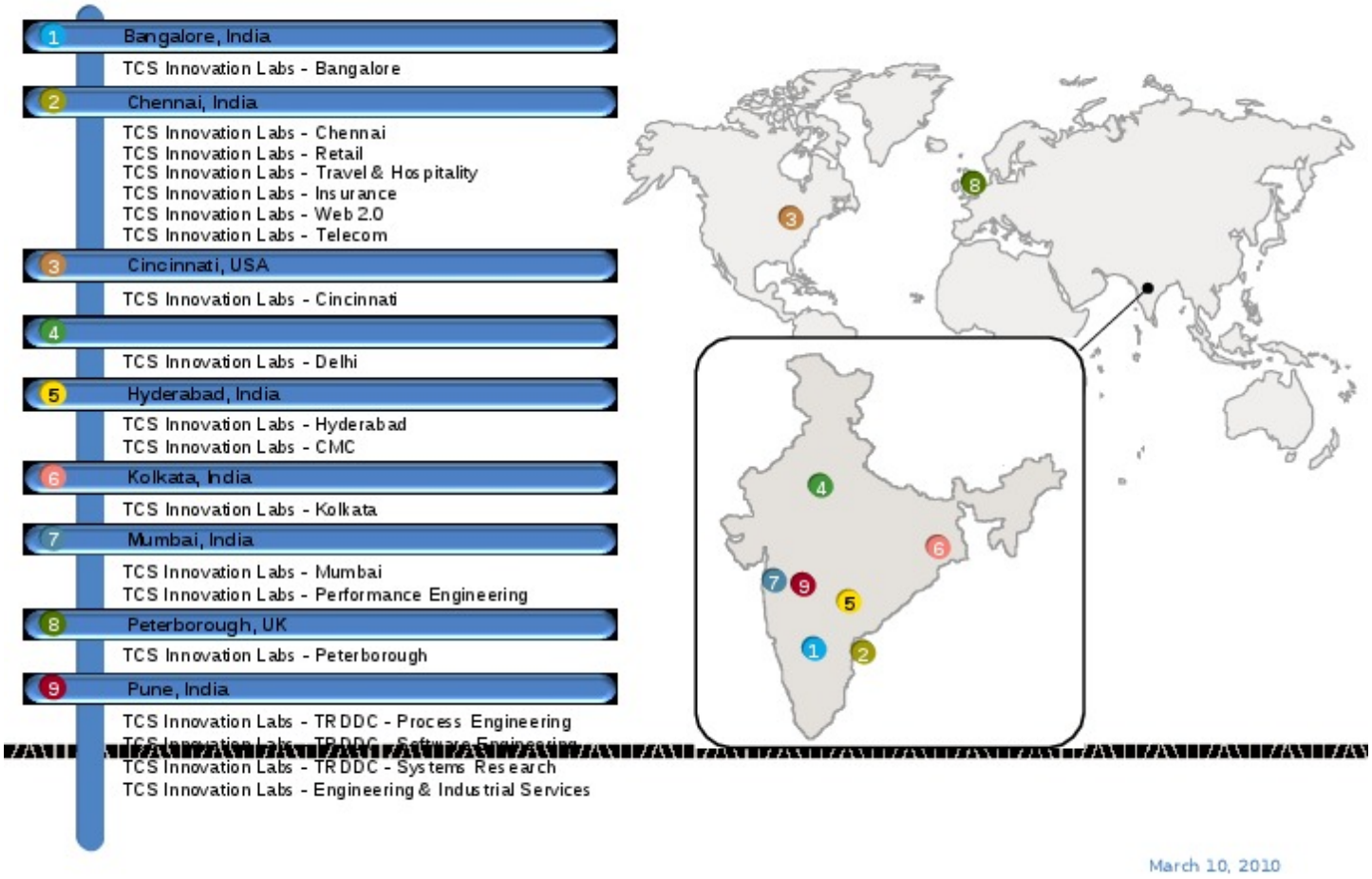
September 26, 2013

Outline

Outline

Outline

TCS Innovation Labs



TCS Innovation Labs - Research Areas

- Systems
  - Cyber Physical Systems
  - Human Centric Systems
  - ICT Systems

TCS Innovation Labs - Research Areas

- Systems
  - Cyber Physical Systems
  - Human Centric Systems
  - ICT Systems
- Scientific Applications
  - Decision Sciences and Algorithms
  - Computational Finance and Risk
  - **Speech and Natural Language**
  - Data Analytics and Information Fusion
  - Web Intelligence and Text Mining
  - Computational Materials Engineering
  - Computational Life Sciences
  - **Mobility and Social Innovation**
  - Multimedia, Graphics and Robotics

TCS Innovation Labs - Research Areas

- Systems
  - Cyber Physical Systems
  - Human Centric Systems
  - ICT Systems
- Scientific Applications
  - Decision Sciences and Algorithms
  - Computational Finance and Risk
  - **Speech and Natural Language**
  - Data Analytics and Information Fusion
  - Web Intelligence and Text Mining
  - Computational Materials Engineering
  - Computational Life Sciences
  - **Mobility and Social Innovation**
  - Multimedia, Graphics and Robotics
- Software
  - Software Testing Research
  - Enterprise Integration
  - Human Aspects of Software Engineering
  - Integrated Computational Material Engineering
  - Requirement Engineering
  - Software Evolution
  - Model-driven Enterprise

TCS Innovation Labs - Research Areas

- Systems
  - Cyber Physical Systems
  - Human Centric Systems
  - ICT Systems
- Scientific Applications
  - Decision Sciences and Algorithms
  - Computational Finance and Risk
  - **Speech and Natural Language**
  - Data Analytics and Information Fusion
  - Web Intelligence and Text Mining
  - Computational Materials Engineering
  - Computational Life Sciences
  - **Mobility and Social Innovation**
  - Multimedia, Graphics and Robotics
- Software
  - Software Testing Research
  - Enterprise Integration
  - Human Aspects of Software Engineering
  - Integrated Computational Material Engineering
  - Requirement Engineering
  - Software Evolution
  - Model-driven Enterprise

Digress

Invention and Innovation Related?

Known adage **Necessity is the mother of invention**

- Invention is
  - the formulation of new ideas
  - for Products or Process or ...

Invention and Innovation Related?

Known adage **Necessity is the mother of invention**

- Invention is
  - the formulation of new ideas
  - for Products or Process or ...
- Innovation is
  - practical application of new inventions
  - into marketable products or services
  - (so that) there is a value add (Economic, Social ...)

Invention and Innovation Related?

Known adage **Necessity is the mother of invention**

- Invention is
  - the formulation of new ideas
  - for Products or Process or ...
- Innovation is
  - practical application of new inventions
  - into marketable products or services
  - (so that) there is a value add (Economic, Social ...)

In Summary

What is an Invention?

Invention

- A new device or a process
  - Should pass originality test to fetch a patent and
  - be sufficiently different from previous inventions
- Minor improvements do not qualify for patents

What is an Invention?

Invention

- A new device or a process
  - Should pass originality test to fetch a patent and
  - be sufficiently different from previous inventions
- Minor improvements do not qualify for patents

Reality Check

- Only a very small percentage of patented inventions have some economic or social value

What is an Invention?

Invention

- A new device or a process
  - Should pass originality test to fetch a patent and
  - be sufficiently different from previous inventions
- Minor improvements do not qualify for patents

Reality Check

- Only a very small percentage of patented inventions have some economic or social value

So what is Innovation?

What is Innovation?

"Innovation is far more about prospecting, mining, refining and adding value than it is about pure invention" William Buxton

But

- Generally the obsession is with inventing something totally unique, rather than extracting value from creative understanding of what is already known.

What is Innovation?

"Innovation is far more about prospecting, mining, refining and adding value than it is about pure invention" William Buxton

But

- Generally the obsession is with inventing something totally unique, rather than extracting value from creative understanding of what is already known.

Fortunately,

- we are TCS **Innovation** Labs.
- Idea is to add value - Social, Economic, ...

What is Innovation?

"Innovation is far more about prospecting, mining, refining and adding value than it is about pure invention" William Buxton

But

- Generally the obsession is with inventing something totally unique, rather than extracting value from creative understanding of what is already known.

Fortunately,

- we are TCS **Innovation** Labs.
- Idea is to add value - Social, Economic, ...

Invention is necessary but not sufficient

What is Innovation?

"Innovation is far more about prospecting, mining, refining and adding value than it is about pure invention" William Buxton

But

- Generally the obsession is with inventing something totally unique, rather than extracting value from creative understanding of what is already known.

Fortunately,

- we are TCS **Innovation** Labs.
- Idea is to add value - Social, Economic, ...

Invention is necessary but not sufficient

TCS Innovation Labs - Mumbai

TCS Innovation Labs - Mumbai

- Speech and Natural Language Processing
  - Speech (Indian Languages; Masses),
  - Script, Image
  - Natural Language Processing
    - Information Retrieval
    - Question Answering Systems

TCS Innovation Labs - Mumbai

- Speech and Natural Language Processing
  - Speech (Indian Languages; Masses),
  - Script, Image
  - Natural Language Processing
    - Information Retrieval
    - Question Answering Systems
- Mobility and Social Innovations
  - mKRISHI (agro advisory system for farmers)
  - mHEALTH (framework for Primary Health Care)
  - GappaGhosti (audio based social networking)
  - Wireless Sensor Networks
  - Human Participatory Sensing

TCS Innovation Labs - Mumbai

- Speech and Natural Language Processing
  - Speech (Indian Languages; Masses),
  - Script, Image
  - Natural Language Processing
    - Information Retrieval
    - Question Answering Systems
- Mobility and Social Innovations
  - mKRISHI (agro advisory system for farmers)
  - mHEALTH (framework for Primary Health Care)
  - GappaGhosti (audio based social networking)
  - Wireless Sensor Networks
  - Human Participatory Sensing

Overview of things that we do. ▶ Video

A Small Deviation!

1. Part I: TCS Innovation Labs
   - TCS Innovation Labs
   - Innovation? Invention?
2. Transition
   - Audio $\longrightarrow$ Image
   - Where is it used?
3. Part II: Accessible Image
   - Image $\longrightarrow$ Audio

Audio ⟶ Image

- Listen to this ▶ Play

Audio ⟶ Image

- Listen to this ▶ Play
- Does it sound familiar?

Audio ⟶ Image

- Listen to this ▶ Play
- Does it sound familiar?
- How about this? ▶ Play

Audio ⟶ Image

- Listen to this ▸ Play
- Does it sound familiar?
- How about this? ▸ Play
- Let us see if this helps ▸ Hidden

Audio ⟶ Image

- Listen to this ▸ Play
- Does it sound familiar?
- How about this? ▸ Play
- Let us see if this helps ▸ Hidden
- Seeing is believing!

Audio ⟶ Image

- Listen to this ▸ Play
- Does it sound familiar?
- How about this? ▸ Play
- Let us see if this helps ▸ Hidden
- Seeing is believing!

Where can this be used?

A System & Method For Visual Message Comm

- Indian Patent Application [1116/MUM/2013; Mar 25, 2013]

ABSTRACT

A SYSTEM AND METHOD FOR VISUAL MESSAGE COMMUNICATION

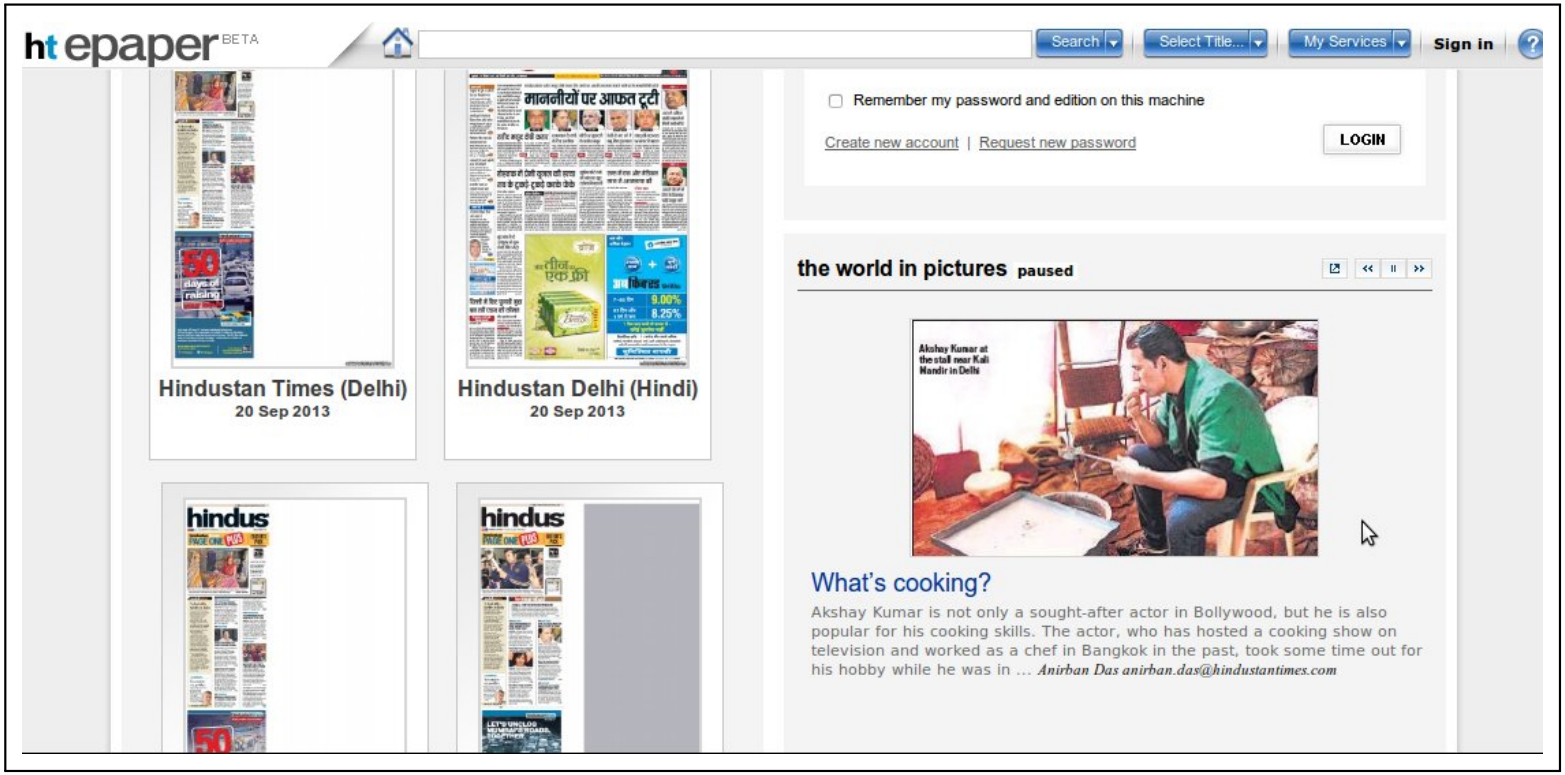
A broadcaster (102) for visual information broadcast, comprising a receiving module (212) configured to receive a first input. The first input comprises a string of characters selected from a text, or an image, or combination thereof. The first input is transformed into a first image having at least two dimensions. The broadcaster further comprises an encoder module (214) configured to convert the first image into a one dimensional signal waveform. Further a modulator module (216) is configured to modulate the one dimensional signal waveform for transmission into a modulated one dimensional signal waveform. The broadcaster further comprises a radio transmitter, wherein the radio transmitter transmits the modulated one dimensional signal waveform over a radio channel, wherein the radio channel is configured to transmit audio signals.

Part II

1. Part I: TCS Innovation Labs
   - TCS Innovation Labs
   - Innovation? Invention?
2. Transition
   - Audio $\longrightarrow$ Image
   - Where is it used?
3. Part II: Accessible Image
   - Image $\longrightarrow$ Audio

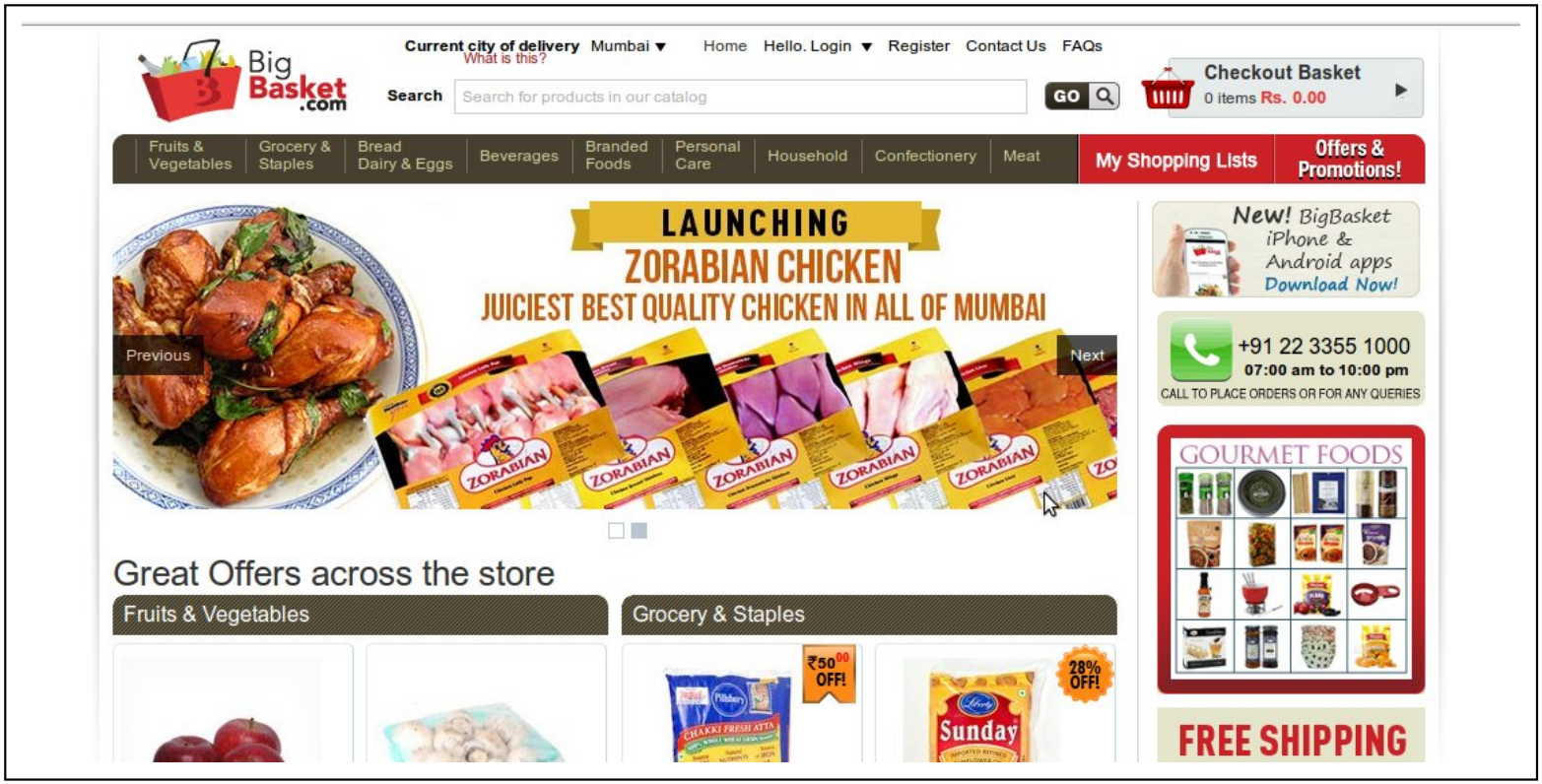
We use pictures all the time

- Example: HT ePaper



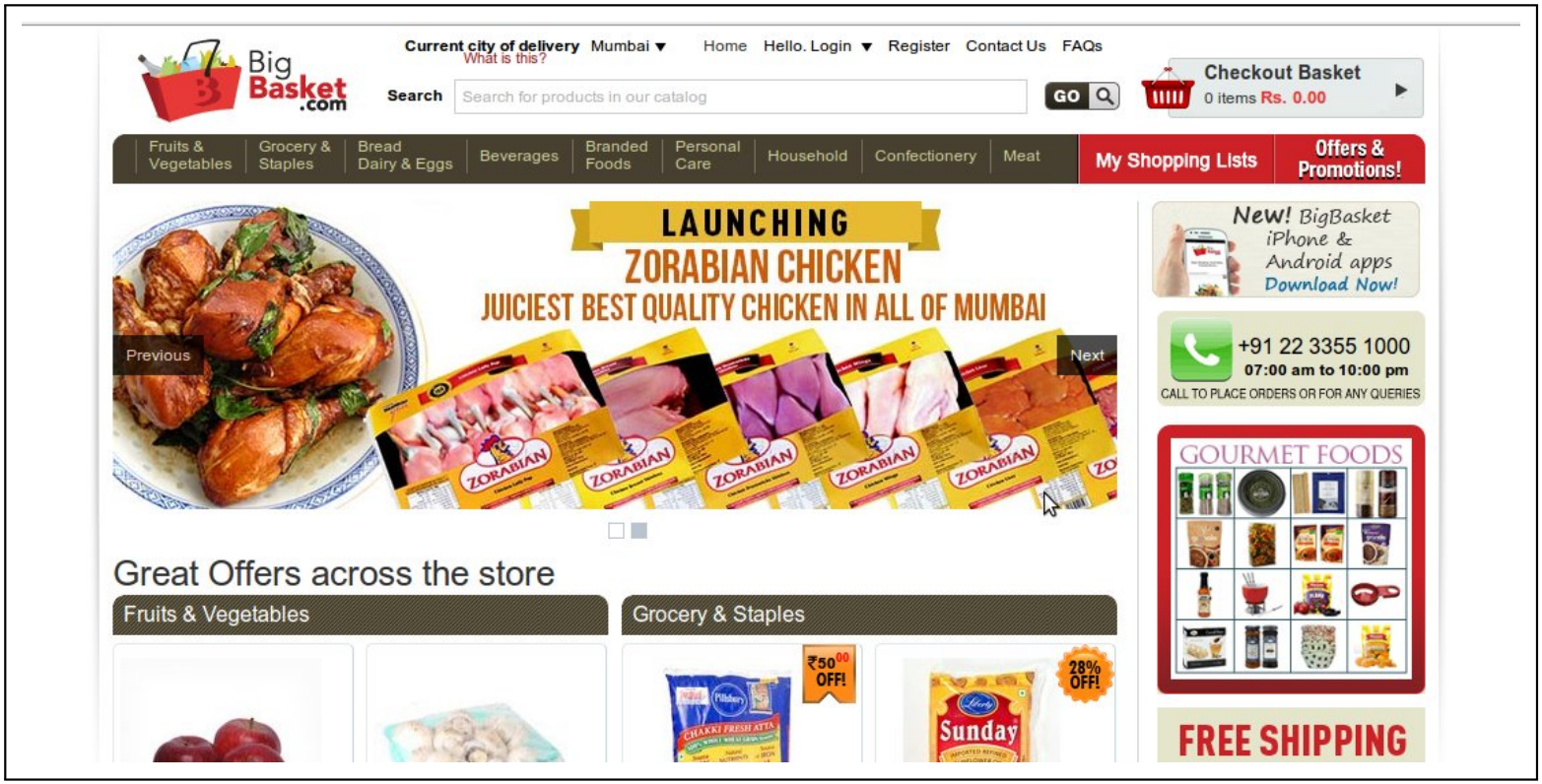
We use pictures all the time

- Example: Big Basket



We use pictures all the time

- Example: Big Basket



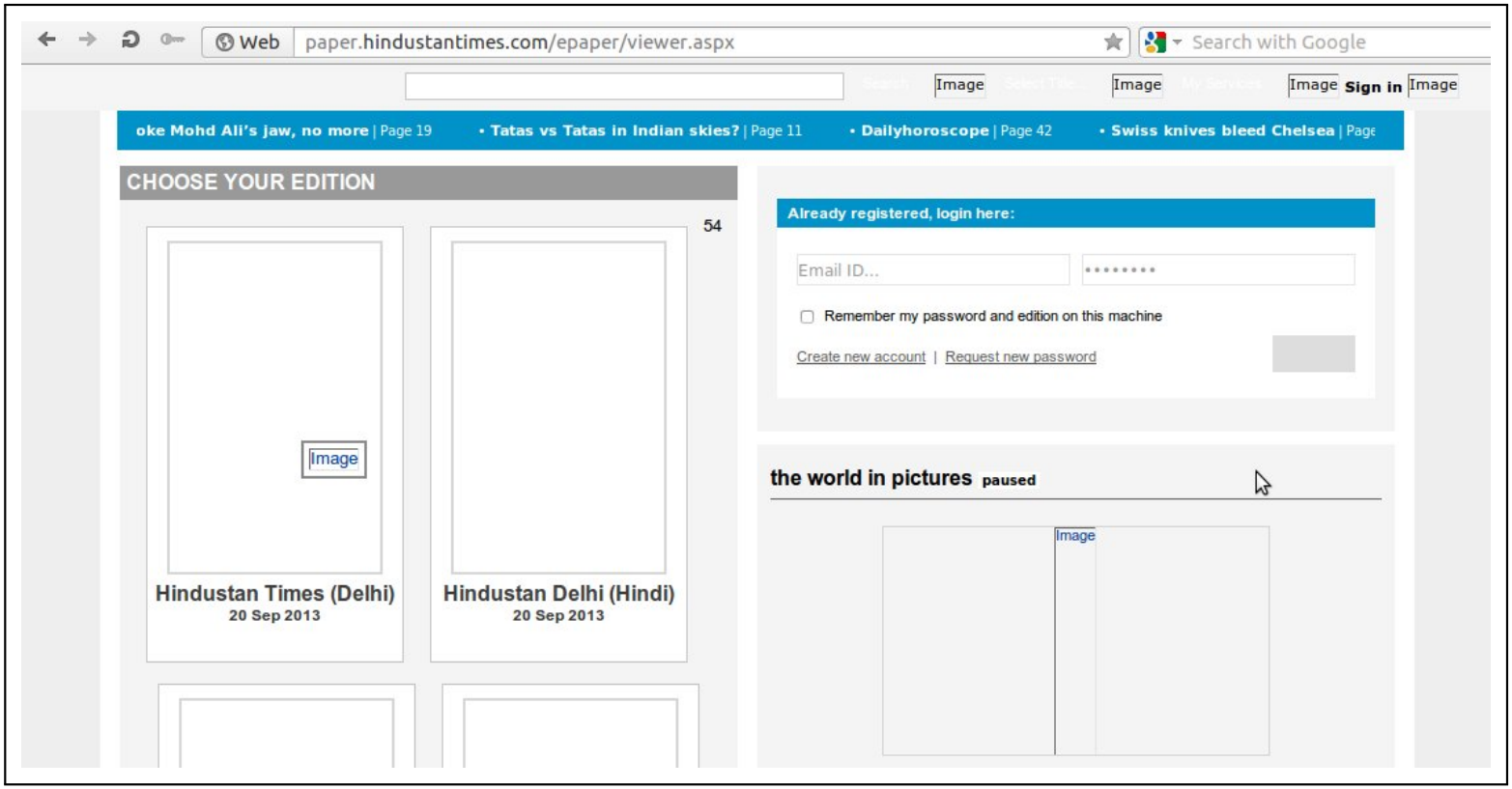
- What if your browser can not render images?

What if Pictures Can't be rendered?

- When your browser can render text and not images

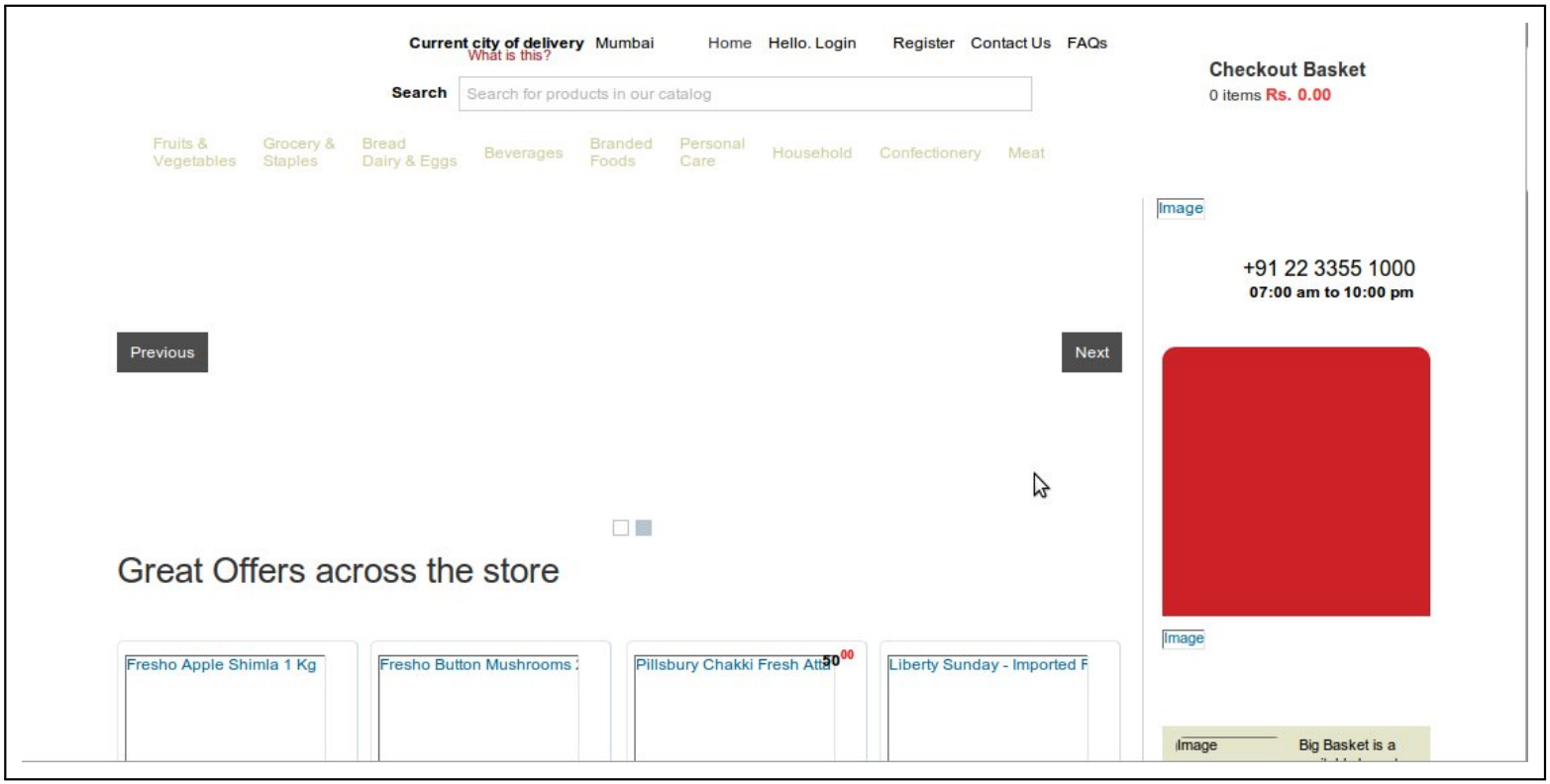
What if Pictures Can't be rendered?

- When your browser can render text and not images
- Example: HT ePaper



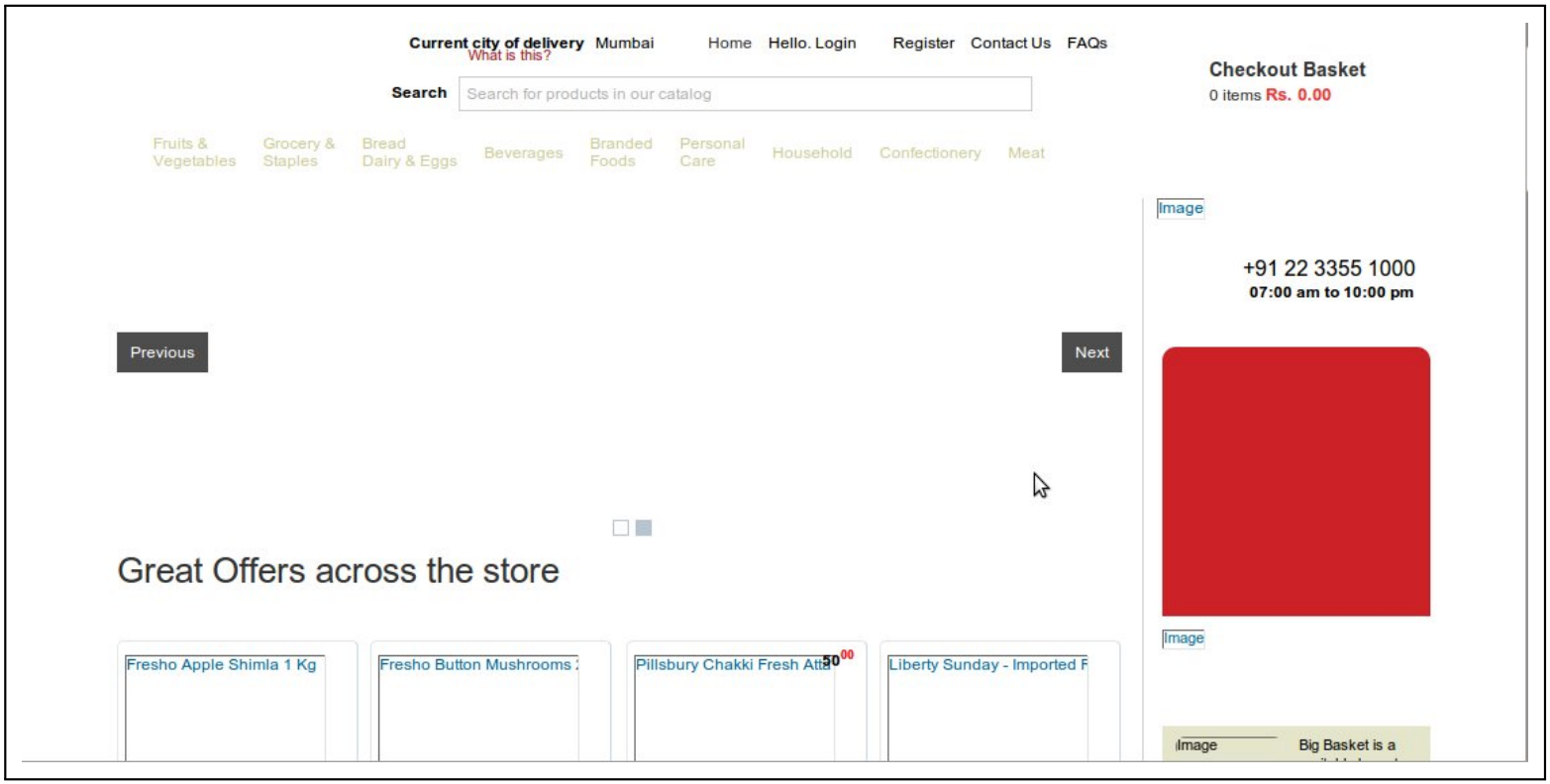
What if Pictures Can't be rendered?

- When your browser can render text and not images
- Example: Big Basket



What if Pictures Can't be rendered?

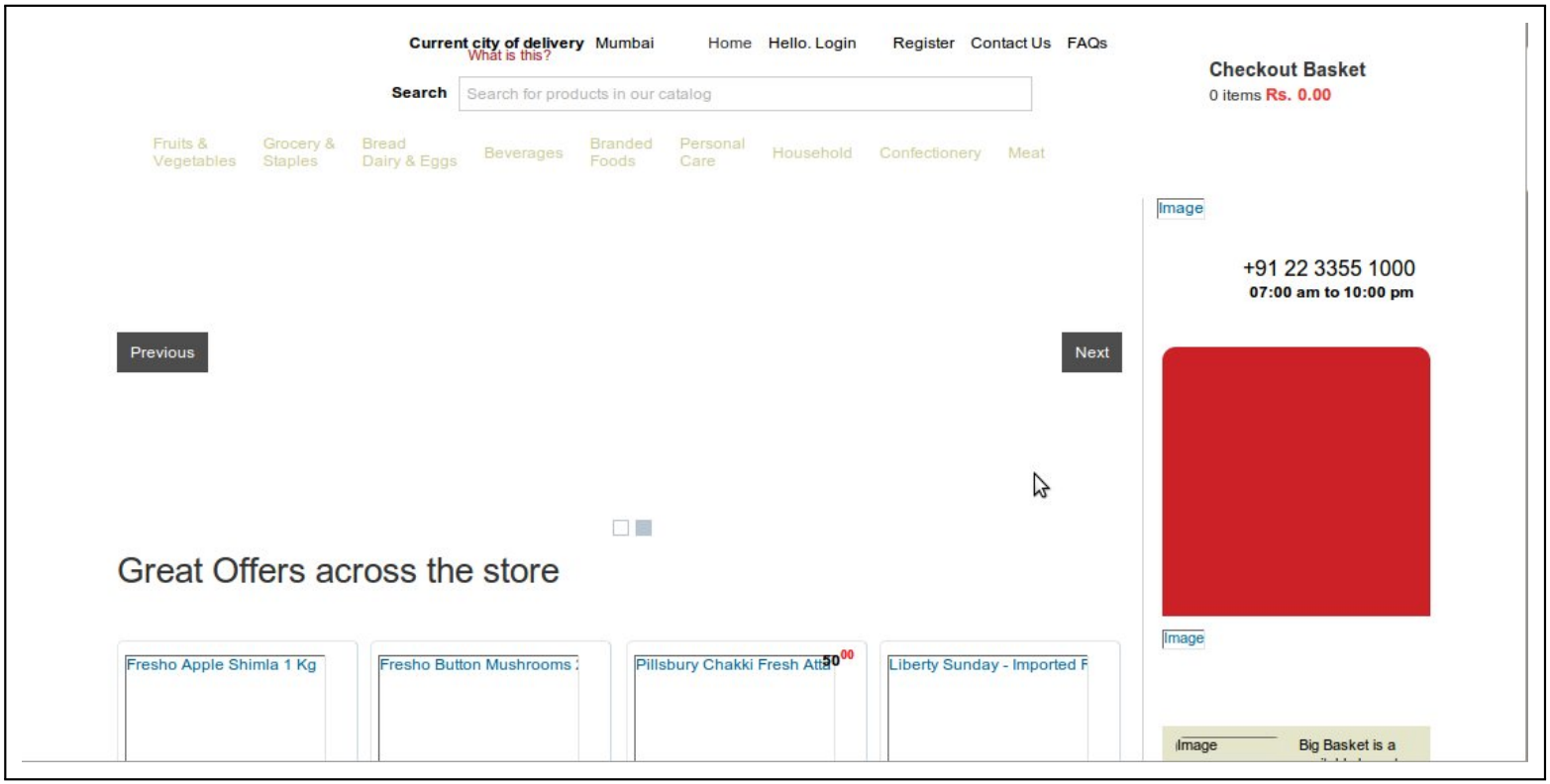
- When your browser can render text and not images
- Example: Big Basket



- All the images are lost!

What if Pictures Can't be rendered?

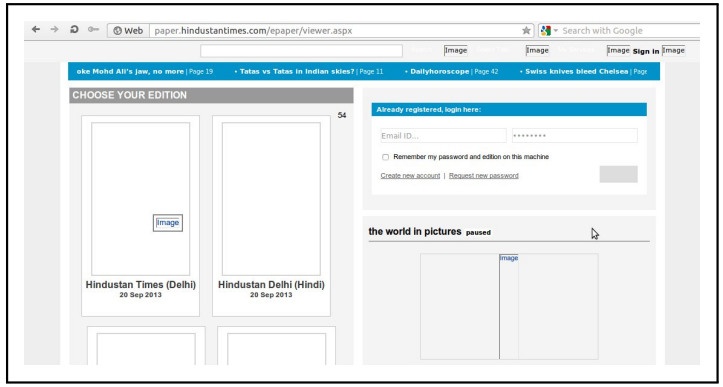
- When your browser can render text and not images
- Example: Big Basket

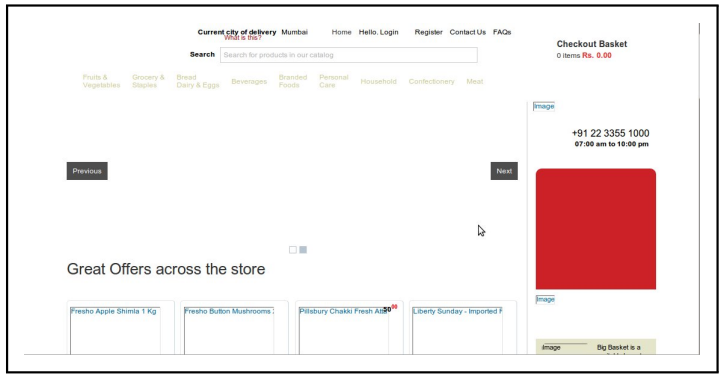


- All the images are lost!

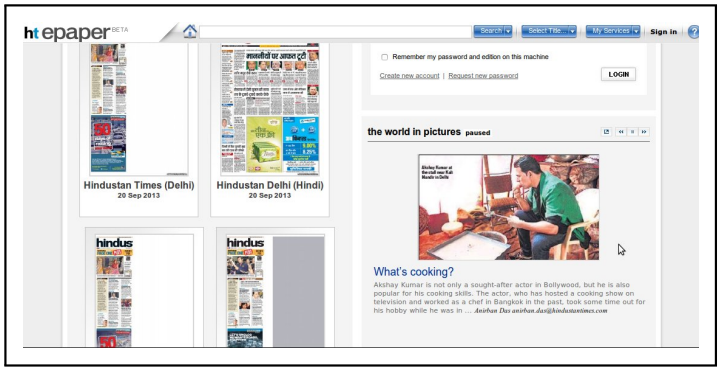
When is this a problem?

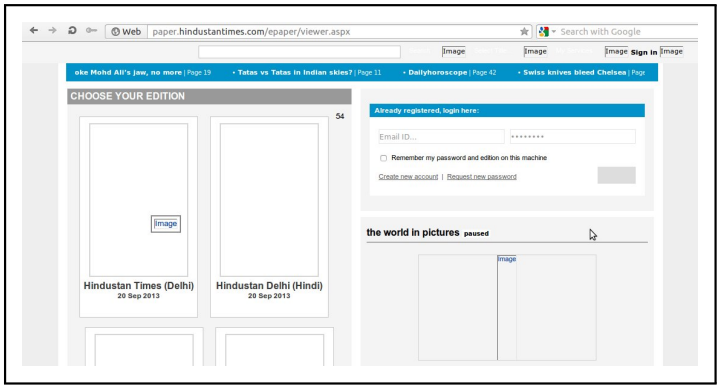
Only when one can visualize images

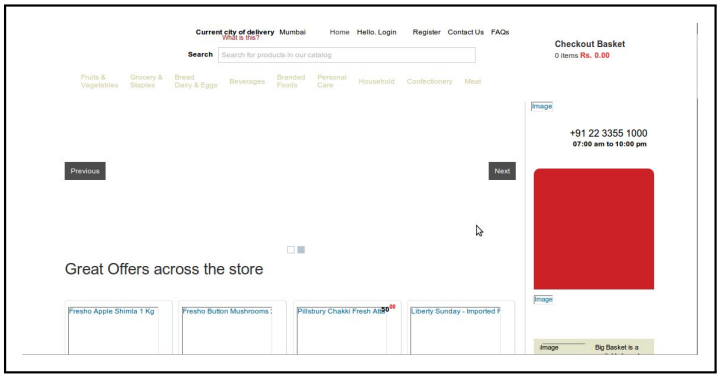




Only when one can visualize images

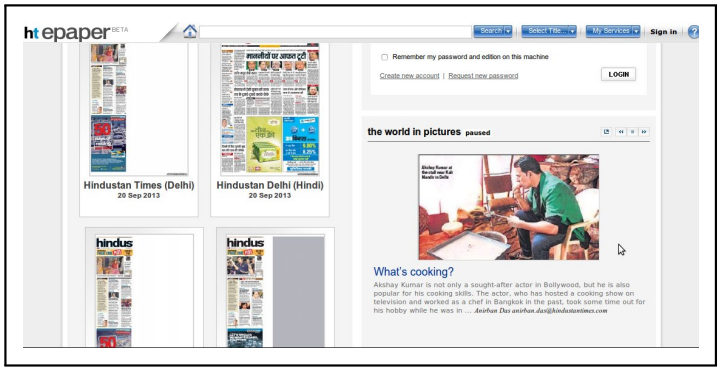


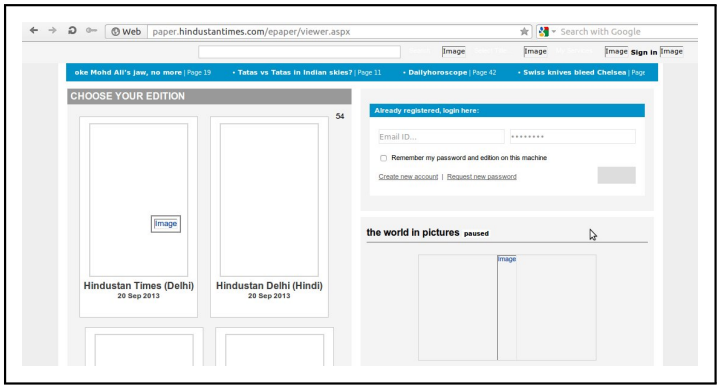


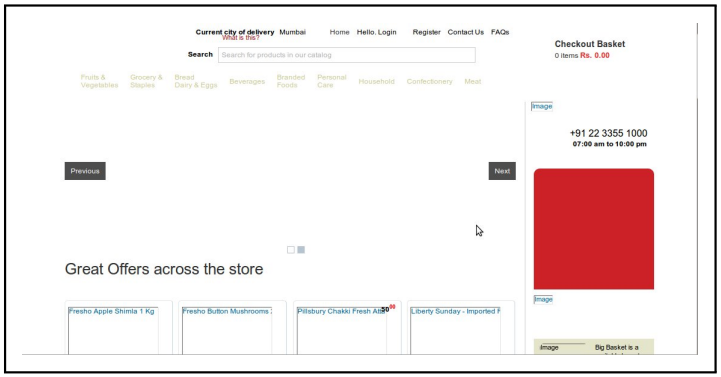


How is this connected to the visually impaired?

Only when one can visualize images







How is this connected to the visually impaired?

Through Screen Readers

Screen Readers

- Piece of Software that identify and interpret what is being displayed on the screen
  - more accurately, information sent to standard output, whether a video monitor is present or not.

Screen Readers

- Piece of Software that identify and interpret what is being displayed on the screen
  - more accurately, information sent to standard output, whether a video monitor is present or not.
- This interpretation is then presented to the user with
  - text-to-speech,
  - sound icons, or
  - a Braille output device.

Screen Readers

- Piece of Software that identify and interpret what is being displayed on the screen
  - more accurately, information sent to standard output, whether a video monitor is present or not.
- This interpretation is then presented to the user with
  - text-to-speech,
  - sound icons, or
  - a Braille output device.
- In general, screen readers read out text
- Visually impaired have access to the text (listening to the read text)

Screen Readers

- Piece of Software that identify and interpret what is being displayed on the screen
  - more accurately, information sent to standard output, whether a video monitor is present or not.
- This interpretation is then presented to the user with
  - text-to-speech,
  - sound icons, or
  - a Braille output device.
- In general, screen readers read out text
- Visually impaired have access to the text (listening to the read text)
- How about graphics?
  - Only if the graphics is equivalently represented by text
  - Or, *Graphic* $\equiv$ *Text Description*

Graphics = *f*(*Text*)

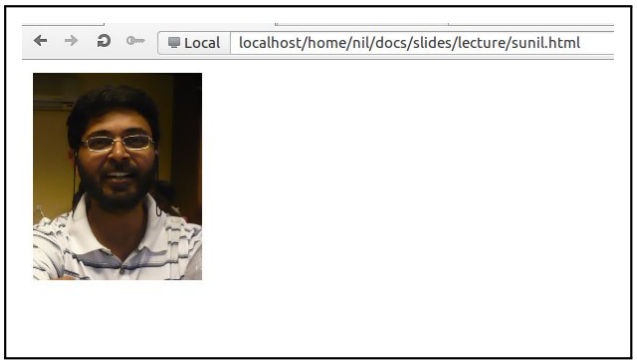
- Document ≡ hyper text markup language (HTML) page
- W3C (World Wide Web Consortium) recommends
  - `alt` attribute, which is designed to be an alternative (text description) for graphics or images on web pages (250 characters) and
  - `longdesc` attribute which is a mechanism to give greater details of the graphic or image (not alternative!).

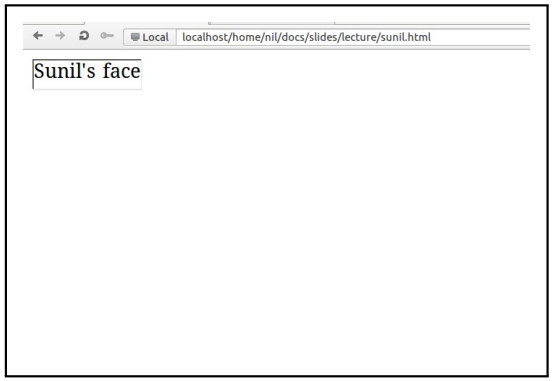
Graphics = *f*(*Text*)

- Document ≡ hyper text markup language (HTML) page
- W3C (World Wide Web Consortium) recommends
  - `alt` attribute, which is designed to be an alternative (text description) for graphics or images on web pages (250 characters) and
  - `longdesc` attribute which is a mechanism to give greater details of the graphic or image (not alternative!).
- Can be used for graphs, plots ▸ How

Graphics = f(Text)

- Document ≡ hyper text markup language (HTML) page
- W3C (World Wide Web Consortium) recommends
  - `alt` attribute, which is designed to be an alternative (text description) for graphics or images on web pages (250 characters) and
  - `longdesc` attribute which is a mechanism to give greater details of the graphic or image (not alternative!).
- Can be used for graphs, plots ▸ How
- How about images? .... In a way
  - `<img src="sunil.gif" alt="Sunil's face">`





`longdesc`?

Describing an Image

-verbosity=1

A flower.

Describing an Image

-verbosity=2

A Yellow flower.

Describing an Image

-verbosity=3

A Yellow flower in midst of green leaves.

Describing an Image

-verbosity=4

A Yellow flower in midst of two smaller yellow flowers in midst of green leaves.

Describing an Image

-verbosity=5

A Yellow flower in midst of two smaller yellow flowers in midst of green leaves. The bigger flower having 9 petals.

Describing an Image

-verbosity=6

A Yellow flower in midst of two smaller yellow flowers in midst of green leaves.
The bigger flower having 9 petals.
Of which 6 petals, to the left, overlapping each other.

Describing an Image

-verbosity=7

A Yellow flower in midst of two smaller yellow flowers in midst of green leaves. The bigger flower having 9 petals. Of which 6 petals, to the left, overlapping each other.
.....

Describing an Image

A Yellow flower in midst of two smaller yellow flowers in midst of green leaves. The bigger flower having 9 petals. Of which 6 petals, to the left, overlapping each other.

....

< Description can go on ..>

Describing an Image

A Yellow flower in midst of two smaller yellow flowers in midst of green leaves. The bigger flower having 9 petals. Of which 6 petals, to the left, overlapping each other.

....

< Description can go on ..>

- Description Speaks about the image. Makes it accessible.

Describing an Image

A Yellow flower in midst of two smaller yellow flowers in midst of green leaves. The bigger flower having 9 petals. Of which 6 petals, to the left, overlapping each other.

....

< Description can go on ..>

- Description Speaks about the image. Makes it accessible.
- Question: How do we describe an image? To make it accessible?

Question Motivates Speaking Image

1. "Readable Image for the Visually Impaired", Universal Access in Human-Computer Interaction. Applications and Services Lecture Notes in Computer Science Volume 6768, 2011, pp 136-145. (HCI International)
2. A Method and System for Construction and Rendering of Annotations associated with an Electronic Image.
   - Indian Patent: 3481/MUM/2010; Dec 22, 2010;
   - US Patent 20120166175; Jun 28, 2012

What is all this about?

- The Problem
  - There is an image on a web page
  - What is the best way to describe this image?

What is all this about?

- The Problem
  - There is an image on a web page
  - What is the best way to describe this image?
- Makes sense, when
  - the browser is unable to render the image or
  - for a visually impaired person can not see it

What is all this about?

- The Problem
  - There is an image on a web page
  - What is the best way to describe this image?
- Makes sense, when
  - the browser is unable to render the image or
  - for a visually impaired person can not see it
- A probable solution

What is all this about?

- The Problem
  - There is an image on a web page
  - What is the best way to describe this image?
- Makes sense, when
  - the browser is unable to render the image or
  - for a visually impaired person can not see it
- A probable solution
  - Describe the image in sufficient (text) detail
  - The text when read out would speak about the image

Earlier Related Work

- Accessibility of graphics specifically in technical documentation.
  - read out block diagrams and the likes of it which are usually accompanied in a technical document, in some sense it uses the descriptions associated with the generation of the block diagrams.

Earlier Related Work

- Accessibility of graphics specifically in technical documentation.
  - read out block diagrams and the likes of it which are usually accompanied in a technical document, in some sense it uses the descriptions associated with the generation of the block diagrams.
- Embed audio data within an image and construct a new image and make the audio data freely recoverable from the embedded image.

Earlier Related Work

- Accessibility of graphics specifically in technical documentation.
  - read out block diagrams and the likes of it which are usually accompanied in a technical document, in some sense it uses the descriptions associated with the generation of the block diagrams.
- Embed audio data within an image and construct a new image and make the audio data freely recoverable from the embedded image.
- Making a priori markings in an image which would be mapped to a corresponding audio embedded into the image file.

Earlier Related Work

- Accessibility of graphics specifically in technical documentation.
  - read out block diagrams and the likes of it which are usually accompanied in a technical document, in some sense it uses the descriptions associated with the generation of the block diagrams.
- Embed audio data within an image and construct a new image and make the audio data freely recoverable from the embedded image.
- Making a priori markings in an image which would be mapped to a corresponding audio embedded into the image file.
- a method that enables visually impaired users to navigate websites and hear high quality audio of narration and description of each website

In Summary

- Associate audio (description) with image

In Summary

- Associate audio (description) with image
- At best, have markers in image so that only the corresponding audio is active.

In Summary

- Associate audio (description) with image
- At best, have markers in image so that only the corresponding audio is active.
- Is this sufficient?

In Summary

- Associate audio (description) with image
- At best, have markers in image so that only the corresponding audio is active.
- Is this sufficient?
  - Description is sequential
  - Is there one complete description for an image?
  - Do all of us describe the image in the same way?
  - Isn't the creator of the image and the viewer of the image different?

In Summary

- Associate audio (description) with image
- At best, have markers in image so that only the corresponding audio is active.
- Is this sufficient?
  - Description is sequential
  - Is there one complete description for an image?
  - Do all of us describe the image in the same way?
  - Isn't the creator of the image and the viewer of the image different?

- Lets look at some of these ....

Describing an Image Revisited

Describing an Image Revisited

How do we describe this image to a visually impaired?

Describing an Image Revisited

How do we describe this image to a visually impaired?
Know, W3C mandates the use of `alt` and `longdesc` in HTML

Using `alt` and `longdesc`

1. `alt` = Two People conversing (not very informative)

Q: Unique?

Using `alt` and `longdesc`

1. `longdesc` = Two people, one in a black coat and a red tie with a black eye wear, balding, fair, ··· sitting on the left ··· and the other in a white shirt and left leg over the right leg, dark complexioned, with black hair and beard ··· sitting on the left of the person in coat ··· sitting on a brownish wooden bench set against a dark brown checkered wall bearing a caption WINE SALES written in Roman all capital Font in light brown with bushes of flowers colored red and white to their right and red and violet to their left with all the flowers in front and some more flowers handing from the top colored yellow and violet with green leaves surrounding them, <more description> conversing.

Unique?

Using `alt` and `longdesc`

Unique?

1. `longdesc` = Two people, sitting on a brownish wooden bench set against a dark brown checkered wall bearing a caption WINE SALES written in Roman all capital Font in light brown with bushes of flowers colored red and white to their right and red and violet to their left with all the flowers in front and some more flowers handing from the top colored yellow and violet with green leaves, one person is in a black coat and a red tie with a black eye wear, balding, fair, $\cdots$ sitting on the left $\cdots$ while the other in a white shirt and left leg over the right leg, dark complexioned, with black hair and beard $\cdots$ sitting on

Describing an Image

`alt` or `longdesc` can be used by a visually impaired person to see the image using a TTS.

Describing an Image

`alt` or `longdesc` can be used by a visually impaired person to see the image using a TTS.

- `alt`: cryptic and conveys an overall description of the image (a typical caption associated with an image),
- `longdesc` is more informative and captures finer details embedded in the image.

Describing an Image

`alt` or `longdesc` can be used by a visually impaired person to see the image using a TTS.

- `alt`: cryptic and conveys an overall description of the image (a typical caption associated with an image),
- `longdesc` is more informative and captures finer details embedded in the image.

So a good longish description can possibly make the image accessible. Are there any problems?

Describing an Image

`alt` or `longdesc` can be used by a visually impaired person to see the image using a TTS.

- `alt`: cryptic and conveys an overall description of the image (a typical caption associated with an image),
- `longdesc` is more informative and captures finer details embedded in the image.

So a good longish description can possibly make the image accessible. Are there any problems?

- The description a user infers when looking at this image depends on
  - his personal interest,
  - the context in which the user is looking at the image
  - cultural background ▸ More and

Describing an Image

`alt` or `longdesc` can be used by a visually impaired person to see the image using a TTS.

- `alt`: cryptic and conveys an overall description of the image (a typical caption associated with an image),
- `longdesc` is more informative and captures finer details embedded in the image.

So a good longish description can possibly make the image accessible. Are there any problems?

- The description a user infers when looking at this image depends on
  - his personal interest,
  - the context in which the user is looking at the image
  - cultural background ▸ More and
- However the description is sequential (the view of the describer!)

Describing an Image

`alt` or `longdesc` can be used by a visually impaired person to see the image using a TTS.

- `alt`: cryptic and conveys an overall description of the image (a typical caption associated with an image),
- `longdesc` is more informative and captures finer details embedded in the image.

So a good longish description can possibly make the image accessible. Are there any problems?

- The description a user infers when looking at this image depends on
  - his personal interest,
  - the context in which the user is looking at the image
  - cultural background ▸ More and
- However the description is sequential (the view of the describer!)

So?

The Scenario

- Make an image accessible
  - a longish description of an image enables accessibility
- However,
  - the description would be sequential and
  - in the manner articulated by the publisher of the image
- A viewer is
  - forced to visualize the image as articulated by the publisher
- What if the viewer wishes to see the image differently?
- What can be done?
  - Allow the publisher to describe the way he likes
  - Allow the viewer to see the image the way he likes
  - Remove sequence aspect of description
- How?
  - Allow multiple level description of the image
  - Allow asynchronous access to description

Multilevel `longdesc`

Multi-level description of the image

- Let image I to be made up of $M \times N$ pixels,
- $I(k, l)$ represents a pixel in the $(k, l)^{th}$ position
- Let there be K levels of descriptions,
  - `level` K is the coarsest description (say "Two people conversing")
  - `level` 1 be the finest description (every pixel described).
  - `level` n has a coarser description than `level` $n + 1$.

`level` K: Two people conversing
:::
`level` ..: :::
`level` 3: A person
`level` 3: Head
`level` 2: Hair
`level` 1: Black

▸ Multi-level Description

Building Multilevel `longdesc`

A Computer Vision Task!

Computer Vision

Make computers understand images and video.

What kind of scene?

Where are the cars?

How far is the building?

...

Building Multilevel `longdesc`

A task for you ;-) ▸ EEL806

[Teaching home]

EEL806: Computer Vision

If you are doing the course, please join the Piazza forum.

Instructors: Sumeet Agarwal and Hiranmay Ghosh
4 credits (3-0-2)
Pre-requisites: EEL205 & EC120
Overlaps with: CSL840
I Semester 2013–14
M Th 17–18:20, II-247

How can machines or computers be equipped with the ability to process and understand the visual information in the world around us? This course will look at algorithms and models for vision. Our approach will be to think of real-world vision as a learning and inference problem from noisy data. In particular, we will focus on employing a statistical machine learning framework for the task of building classifiers for recognising various visual phenomena.

Evaluation components

- Minors: 20% (best of 2) [paper I]

Building Multilevel `longdesc`

Semi supervised

- The finest level details corresponding to each pixel can be captured automatically
  - just the colour of the pixel; map RGB value to a colour

Building Multilevel `longdesc`

Semi supervised

- The finest level details corresponding to each pixel can be captured automatically
  - just the colour of the pixel; map RGB value to a colour
- The next level, description can be achieved by a basic image segmentation
  - pixels that have similar properties grouped together and assigned a name

Building Multilevel `longdesc`

Semi supervised

- The finest level details corresponding to each pixel can be captured automatically
  - just the colour of the pixel; map RGB value to a colour
- The next level, description can be achieved by a basic image segmentation
  - pixels that have similar properties grouped together and assigned a name
- Example (brick wall behind the two people),
  - a pixel in this area would have a `level` 1 description as brown color while
  - the `level` 2 description would be brick wall.

Building Multilevel `longdesc`

Semi supervised

- The finest level details corresponding to each pixel can be captured automatically
  - just the colour of the pixel; map RGB value to a colour
- The next level, description can be achieved by a basic image segmentation
  - pixels that have similar properties grouped together and assigned a name
- Example (brick wall behind the two people),
  - a pixel in this area would have a `level` 1 description as brown color while
  - the `level` 2 description would be brick wall.

Clearly a high-level vision task of image analysis, image interpretation

Building Multilevel `longdesc`

Semi supervised

- The finest level details corresponding to each pixel can be captured automatically
  - just the colour of the pixel; map RGB value to a colour
- The next level, description can be achieved by a basic image segmentation
  - pixels that have similar properties grouped together and assigned a name
- Example (brick wall behind the two people),
  - a pixel in this area would have a `level` 1 description as brown color while
  - the `level` 2 description would be brick wall.

Clearly a high-level vision task of image analysis, image interpretation

What do we have?

What do we have?

- We now have a ▸ Multi-level Description.

```
<description>
        <image>
                <name> converse.jpg </name>
                <size> M x N </size>
        </image>

        <pixel, x, y>
                <level 1>
                        <des> Black </des>
                </level 1>
                <level 2>
                        <des> Hair </des>
                </level 2>
                <level 3>
                        <des> Head </des>
                </level 3>
                <level 4>
                        <des> A person </des>
                        <des> Sitting to the right </des>
                </level 4>
                ...
                ...
                ...
                <level K>
                        <des> Two people having conversation </des>
                </level K>
        </pixel, x, y>
</description>
```

How do we use it?

What do we have?

- We now have a ▸ Multi-level Description.

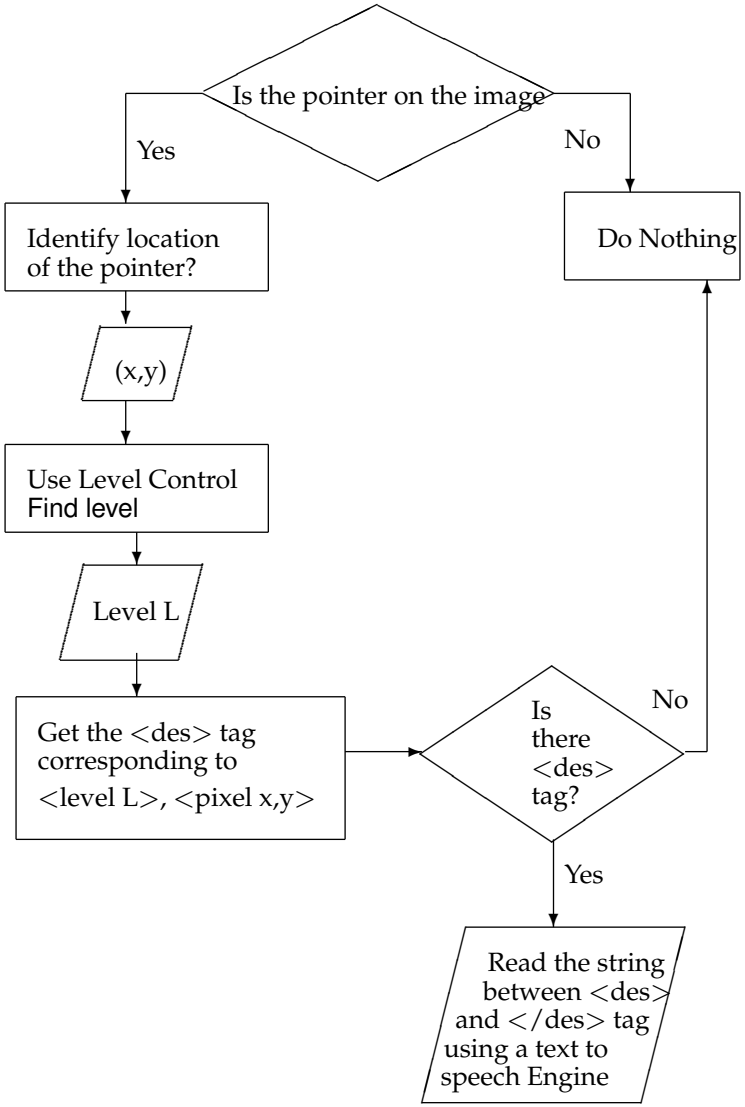
```
<description>
        <image>
                <name> converse.jpg </name>
                <size> M x N </size>
        </image>

        <pixel, x, y>
                <level 1>
                        <des> Black </des>
                </level 1>
                <level 2>
                        <des> Hair </des>
                </level 2>
                <level 3>
                        <des> Head </des>
                </level 3>
                <level 4>
                        <des> A person </des>
                        <des> Sitting to the right </des>
                </level 4>
                ...
                ...
                ...
                <level K>
                        <des> Two people having conversation </des>
                </level K>
        </pixel, x, y>
</description>
```

- For every pixel (x, y) and for every `level` 1 to `level` K there is a description.

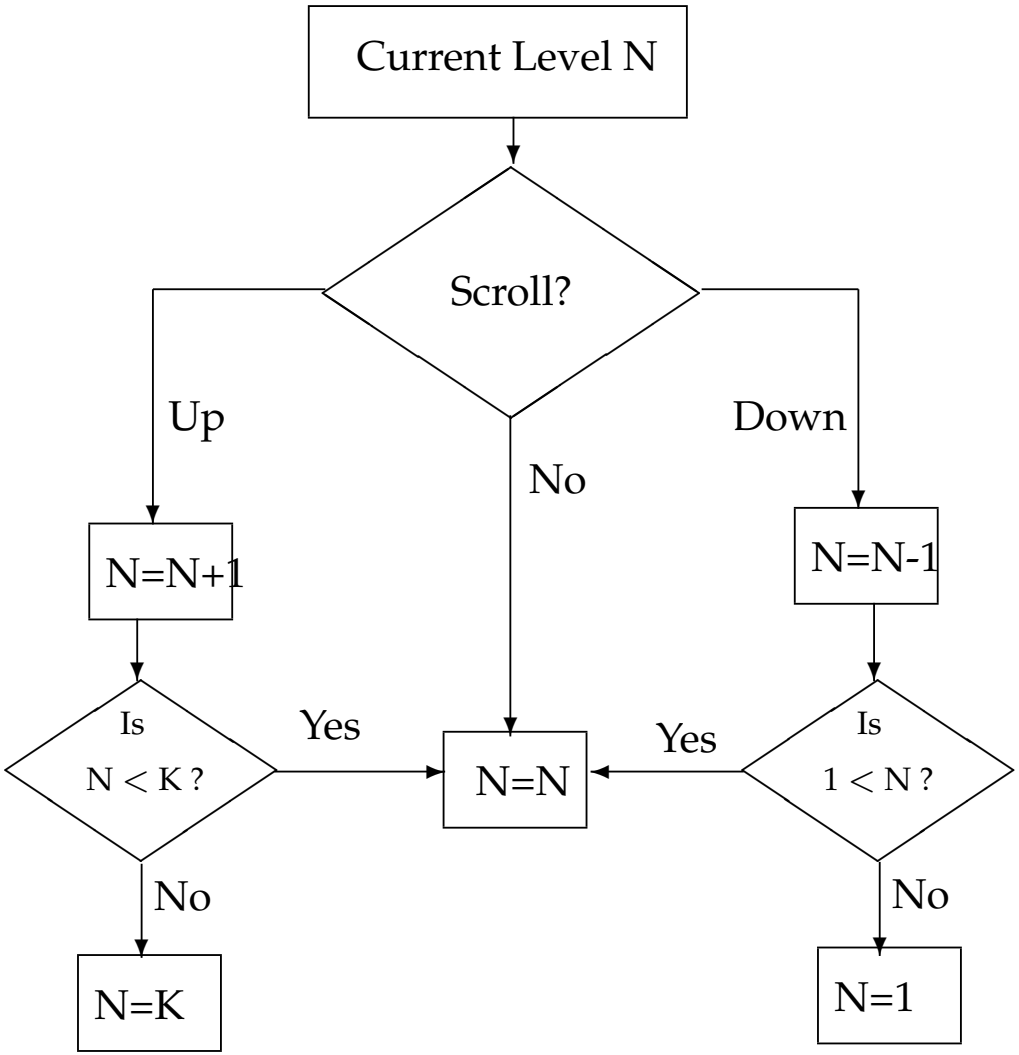
How do we use it?

Accessing Multilevel `longdesc`



Assumption: Mouse with a scroll button.

Accessing Multilevel `longdesc`



Making Image Accessible!

A Method and System for Construction and Rendering of Annotations associated with an Electronic Image.

- Indian Patent: 3481/MUM/2010; Dec 22, 2010;
- US Patent 20120166175; Jun 28, 2012

---

1. a method to annotate images at different levels,
2. a method to access information about the image at different levels,
3. a method to capture annotation in a description file at different levels and
4. a method that makes accessible an image in an non-sequential fashion.

Thank You

- Queries?
- Comments
- Suggestions?

Dr Sunil Kopparapu

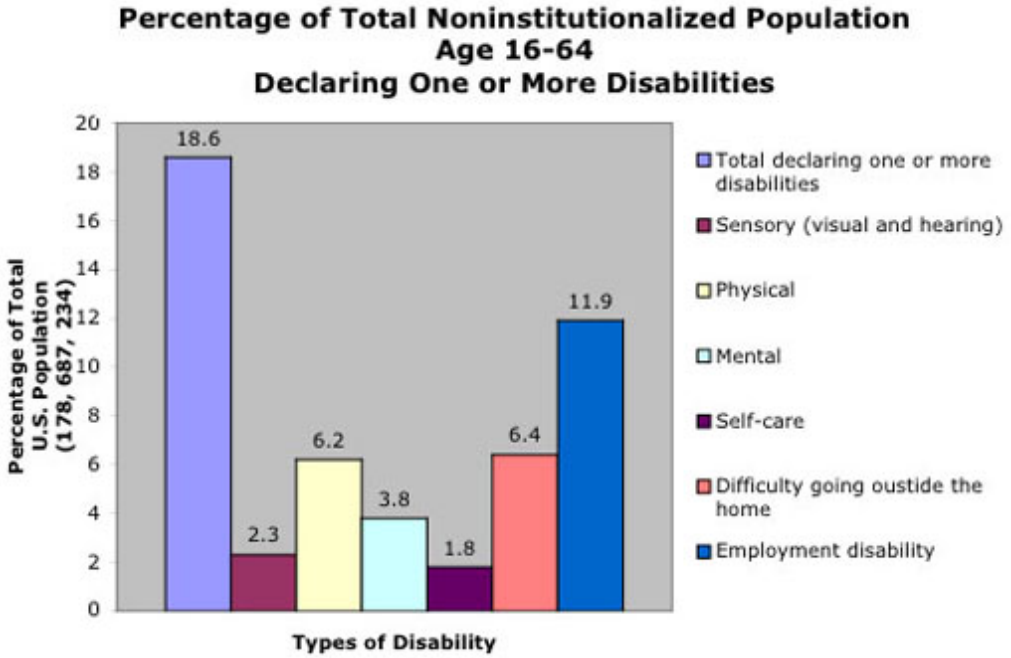
SunilKumar.Kopparapu@TCS.Com
TCS Innovation Lab - Mumbai
Tata Consultancy Services Limited
Yantra Park, Thane (West), India.

END

Additional Material

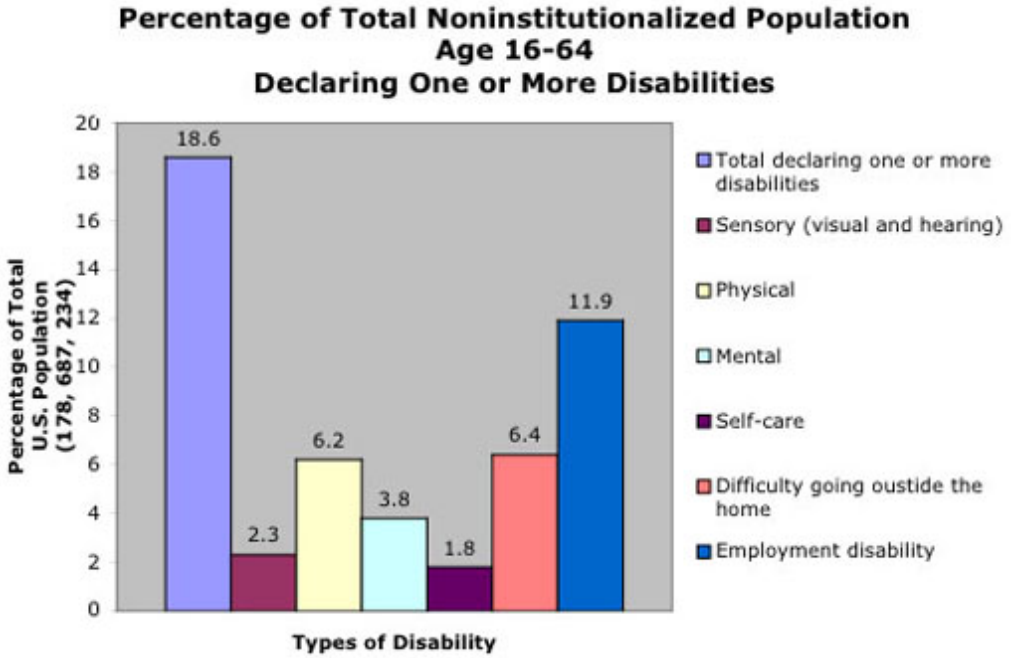
Stop and Go Back

Graphs and Plots



- `alt` = Graph of percentage of total U.S. non-institutionalized population age 16-64 declaring one or more disabilities

Graphs and Plots



- `longdesc =` The percentages of total U.S. non-institutionalized population age 16-64 declaring one or more disabilities are: Total declaring one or more disabilities: 18.6%, Sensory (visual and hearing): 2.3%, Physical: 6.2%, Mental: 3.8%, Self-care: 1.8%, Difficulty going outside the home: 6.4%, Employment disability: 11.9%.

Go Back

How do we see Images?

Cultural variation in eye movements during scene perception. Proc. of the National Academy of Sciences of the USA 102(35), 1262912633 (2005),

- When viewing naturalistic scenes, Westerners attend more to focal objects, whereas East Asians attend more to contextual information
- How? Observing eye movements of American and Chinese participants while they viewed photographs with a focal object on a complex background.
- The sequence in which the image is visualized by a person depends on the cultural background of the person viewing the image.

Go Back