

# Help, It looks Confusing:

GUI Task Automation Through Demonstration and Follow-up Questions

Thanapong Intharah,

Daniyar Turmukhambetov

Gabriel J. Brostow

University College London

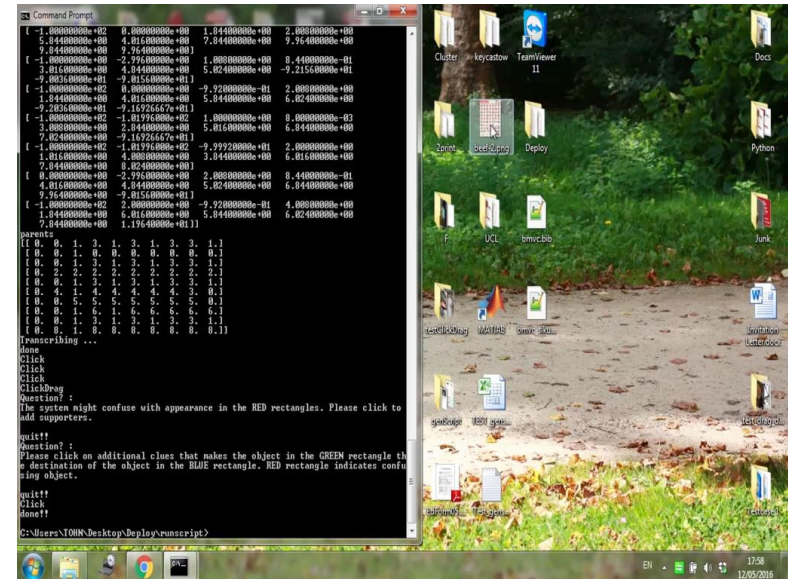
# Target Users:

People Who Want To ...

# Target Users: People Who Want To ...

(Linear) Have a macro/script where one click does many actions.

*Example: switch everything to high contrast mode.*



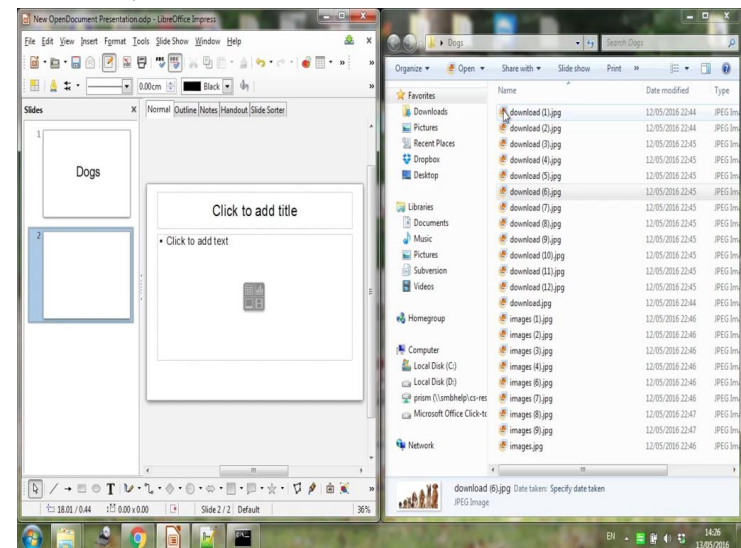
# Target Users: People Who Want To ...

(Linear) Have a macro/script where one click does many actions.

*Example: switch everything to high contrast mode.*

(Loop) Do iterative tasks, automatically.

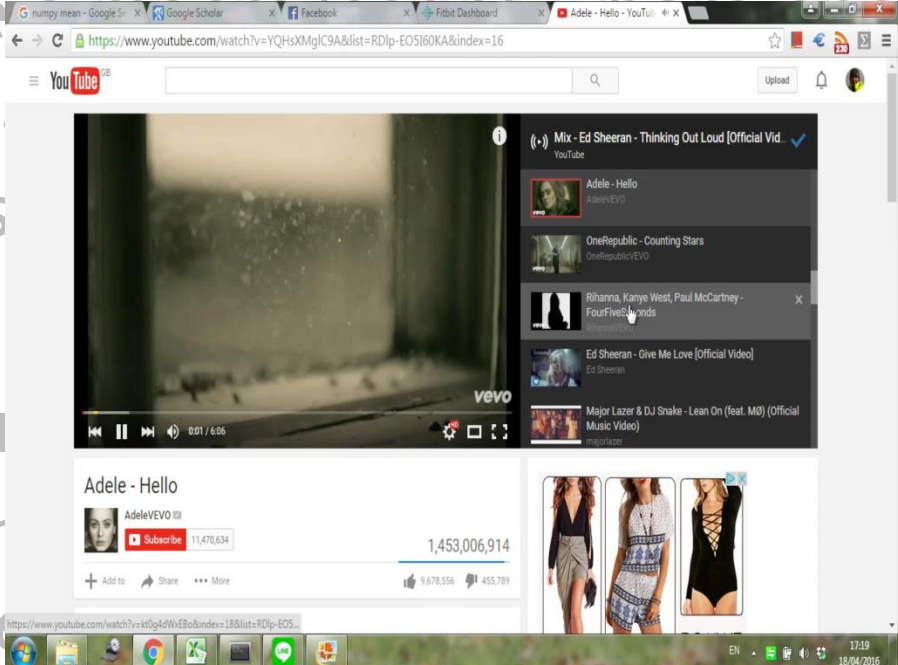
*Example: make each photo in this folder into a separate slide in PowerPoint.*



# Target Users: People Who Want To ...

(Linear) Have many actions.  
*Example: scroll down to click does contrast mode.*

(Loop) Do iterations.  
*Example: repeat the same actions on a separate slide.*

A screenshot of a web browser displaying a YouTube video player. The video is titled 'Adele - Hello' and is from the channel 'AdeleVEVO'. The video player shows a blurred image of Adele. Below the video, there are buttons for 'Add to', 'Share', and 'More'. The video has 1,453,006,914 views and 9,678,556 likes. On the right side of the video player, there is a list of recommended videos, including 'Adele - Hello', 'OneRepublic - Counting Stars', 'Bhanna, Kanye West, Paul McCartney - FourFiveSeconds', 'Ed Sheeran - Give Me Love [Official Video]', and 'Major Lazer & DJ Snake - Lean On (feat. MØ) (Official Music Video)'. The browser's address bar shows the URL 'https://www.youtube.com/watch?v=YQHsXMgIC9A&list=RDlp-EOS360KA&index=16'. The Windows taskbar is visible at the bottom of the screen.
















(Monitoring) React the same way each time ...  
*Example: click the “skip ad” during a youtube video.*

# Before HILC

# Before HILC
















- Nearest neighbors : Programming by Demonstration
  - Sikuli, Sikuli Slides (*Yeh et al 2009*).
  - Koala (*Little et al 2007*), CoScripter (*Leshed et al 2008*).
  - More systems can be found in “Watch what I do” (*Cypher and Halbert 1993*).

# Feature Comparison

















	Domain Independent	No Programming Skill Required	Deal with Non-linear Tasks
Sikuli			
Sikuli Slides			
Koala, CoScriptor			
Sheepdog, Familiar			
HILC			



















# Feature Comparison

	Domain Independent	No Programming Skill Required	Deal with Non-linear Tasks
➡ Sikuli			
Sikuli Slides			
Koala, CoScriptor			
Sheepdog, Familiar			
HILC			
















# Feature Comparison

	Domain Independent	No Programming Skill Required	Deal with Non-linear Tasks
Sikuli			
 Sikuli Slides			
Koala, CoScriptor			
Sheepdog, Familiar			
HILC			

















# Feature Comparison

	Domain Independent	No Programming Skill Required	Deal with Non-linear Tasks
Sikuli			
Sikuli Slides			
 Koala, CoScriptor			
Sheepdog, Familiar			
HILC			

# Feature Comparison

	Domain Independent	No Programming Skill Required	Deal with Non-linear Tasks
Sikuli			
Sikuli Slides			
Koala, CoScriptor			
Sheepdog, Familiar			
HILC			

# Feature Comparison

	Domain Independent	No Programming Skill Required	Deal with Non-linear Tasks
Sikuli			
Sikuli Slides			
Koala, CoScriptor			
Sheepdog, Familiar			
 HILC			

# Before HILC

- Nearest neighbours : Programming by Demonstration
  - Sikuli, Sikuli Slides (*Yeh et al 2009*).
  - Koala (*Little et al 2007*), CoScripter (*Leshed et al 2008*).
  - More systems can be found in “Watch what I do” (*Cypher and Halbert 1993*).
- Differences :
  - HILC is software-agnostic;  
no need for Accessibility API's and works across applications.
  - Inputs users feed to HILC resemble how a human teaches another to complete a task.
  - Able to deal with non-linear task.

# From User's Perspective


Phase A. HILC observes a user<sup>1</sup> demonstrating a task.

Phase B. HILC asks a user<sup>2</sup> to clarify confusing steps, if any,  
and then synthesizes a task script.

Phase C. A user<sup>3</sup> triggers the script to complete the task.

# From User's Perspective

## Phase A. HILC observes a user<sup>1</sup> demonstrating a task.




```
Command Prompt - python genLog.py --p=USER007\002_21
4.87831624 3.91831624 6.89831624 7.89831624 8.91831624
[ 0. -2.56084188 2.88831624 -0.11168376 4.90831624
1.90831624 5.91831624 4.90831624 10.90831624 -91.83168376]
[-100. 2. -0.12168376 4.89831624 1.89831624
6.91831624 3.91831624 6.91831624 7.89831624 12.91831624]]
parents
[[ 0. 0. 1. 8. 9. 8. 9. 2. 3. 3.]
[ 0. 0. 1. 0. 9. 0. 9. 0. 3. 0.]
[ 0. 0. 1. 8. 9. 8. 9. 2. 3. 3.]
[ 0. 2. 1. 2. 9. 2. 9. 2. 3. 2.]
[ 0. 0. 1. 8. 9. 8. 9. 2. 3. 3.]
[ 0. 4. 1. 4. 9. 4. 9. 4. 3. 8.]
[ 0. 0. 5. 5. 5. 5. 5. 5. 8.]
[ 0. 0. 1. 6. 9. 6. 9. 6. 3. 6.]
[ 0. 0. 1. 8. 9. 8. 9. 2. 3. 3.]
[ 0. 8. 1. 8. 9. 8. 9. 8. 3. 8.]]
Transcribing ...
done
Click
RClick
RClick
ClickDrag
Question? :
The system might confuse with appearance in the RED rectangles. Please click to
add supporters.

quit!!
Question? :
Please click on additional clues that makes the object in the GREEN rectangle the
destination of the object in the BLUE rectangle. RED rectangle indicates confusing
object.

forrtl: error (200): program aborting due to control-C event
Image PC Routine Line Source
libifcoremd.dll 000007FEDFD4314 Unknown Unknown Unknown
kernel32.dll 0000000076D44803 Unknown Unknown Unknown
kernel32.dll 0000000076D0652D Unknown Unknown Unknown
ntdll.dll 0000000076F8C521 Unknown Unknown Unknown
Traceback (most recent call last):
  File "rerun.py", line 13, in <module>
    subprocess.call(['python', 'basic_genscript_offset_sniffer.py', '-p='+main_
path, '-s=sniffer'])
  File "C:\Users\TOHN\Anaconda2\lib\subprocess.py", line 522, in call
    return Popen(*popenargs, **kwargs).wait()
  File "C:\Users\TOHN\Anaconda2\lib\subprocess.py", line 1007, in wait
    _subprocess.INFINITE)
KeyboardInterrupt

C:\Users\TOHN\Desktop\Deploy\runcscript>cd ..
C:\Users\TOHN\Desktop\Deploy>python genLog.py --p=USER007\002_21
```



Instructor



Teacher



End-user



# From User's Perspective

## Phase B. HILC asks a user<sup>2</sup> to clarify confusing steps.

```
Command Prompt - python rerun.py --p=C:\Users\TOHN\Desktop\Deploy\USER007\002_Z1
[ 0.0000000e+00 -1.0000000e+00 0.0000000e+00 -1.0000000e+00
 0.0000000e+00 -1.0000000e+00 -1.0000000e+00 -1.0000000e+00
 0.0000000e+00 -1.0000000e+02 ]
[ -1.0000000e+02 0.0000000e+00 -1.0000000e+00 0.0000000e+00
 -1.0000000e+00 0.0000000e+00 -1.0000000e+00 -1.0000000e+00
 -1.0000000e+00 0.0000000e+00 ]
bestPath
[ 0. 1. 0. 1. 0. 1. 2. 3. 0. 1.]
max_score
[[ 4.0000000e-03 -1.9960000e+00 2.0120000e+00 8.4400000e-01
 4.0200000e+00 4.8440000e+00 6.0240000e+00 7.8440000e+00
 1.0307333e+01 -9.0156000e+01 ]
[ -1.0000000e+02 2.0080000e+00 8.0000000e-03 4.0160000e+00
 1.8440000e+00 6.0240000e+00 5.8440000e+00 8.0240000e+00
 9.8440000e+00 1.2650667e+01 ]
[ -1.0000000e+00 -1.0760000e+00 1.0000000e+00 2.8440000e+00
 3.0160000e+00 5.8440000e+00 6.9940000e+00 7.8440000e+00
 8.9640000e+00 -9.0156000e+01 ]
[ -1.0000000e+02 0.0000000e+00 1.8440000e+00 2.0080000e+00
 5.8440000e+00 4.0160000e+00 7.8440000e+00 9.9640000e+00
 9.8440000e+00 9.9640000e+00 ]
[ -1.0000000e+00 -2.9960000e+00 1.0000000e+00 8.4400000e-01
 3.0160000e+00 4.8440000e+00 5.0240000e+00 -9.2156000e+01
 -9.0036000e+01 -9.0156000e+01 ]
[ -1.0000000e+02 0.0000000e+00 -9.9200000e-01 2.0080000e+00
 1.8440000e+00 4.0160000e+00 5.8440000e+00 6.0240000e+00
 -9.2036000e+01 -9.1672667e+01 ]
[ -1.0000000e+02 -1.0199600e+02 1.0000000e+00 8.0000000e-03
 3.0080000e+00 2.8440000e+00 5.0160000e+00 6.8440000e+00
 7.0240000e+00 -9.1672667e+01 ]
[ -1.0000000e+02 -1.0199600e+02 -9.9992000e+01 2.0000000e+00
 1.0160000e+00 4.0080000e+00 3.8440000e+00 6.0160000e+00
 7.8440000e+00 8.0240000e+00 ]
[ 0.0000000e+00 -2.9960000e+00 2.0080000e+00 8.4400000e-01
 4.0160000e+00 4.8440000e+00 5.0240000e+00 6.8440000e+00
 9.9640000e+00 -9.0156000e+01 ]
[ -1.0000000e+02 2.0000000e+00 -9.9200000e-01 4.0080000e+00
 1.8440000e+00 6.0160000e+00 5.8440000e+00 6.0240000e+00
 7.8440000e+00 1.1964000e+01 ]]
parents
[[ 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.1
 0. 0. 1. 0. 0. 0. 0. 0. 0. 0.1
 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.1
 0. 2. 2. 2. 2. 2. 2. 2. 2. 2.1
 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.1
 0. 4. 1. 4. 4. 4. 4. 4. 4. 3.0.1
 0. 0. 5. 5. 5. 5. 5. 5. 5. 0.1
 0. 0. 1. 6. 1. 6. 6. 6. 6. 6.1
 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.1
 0. 8. 1. 8. 0. 8. 8. 8. 8. 8.1 ]]
Transcribing ...
done
```



Instructor



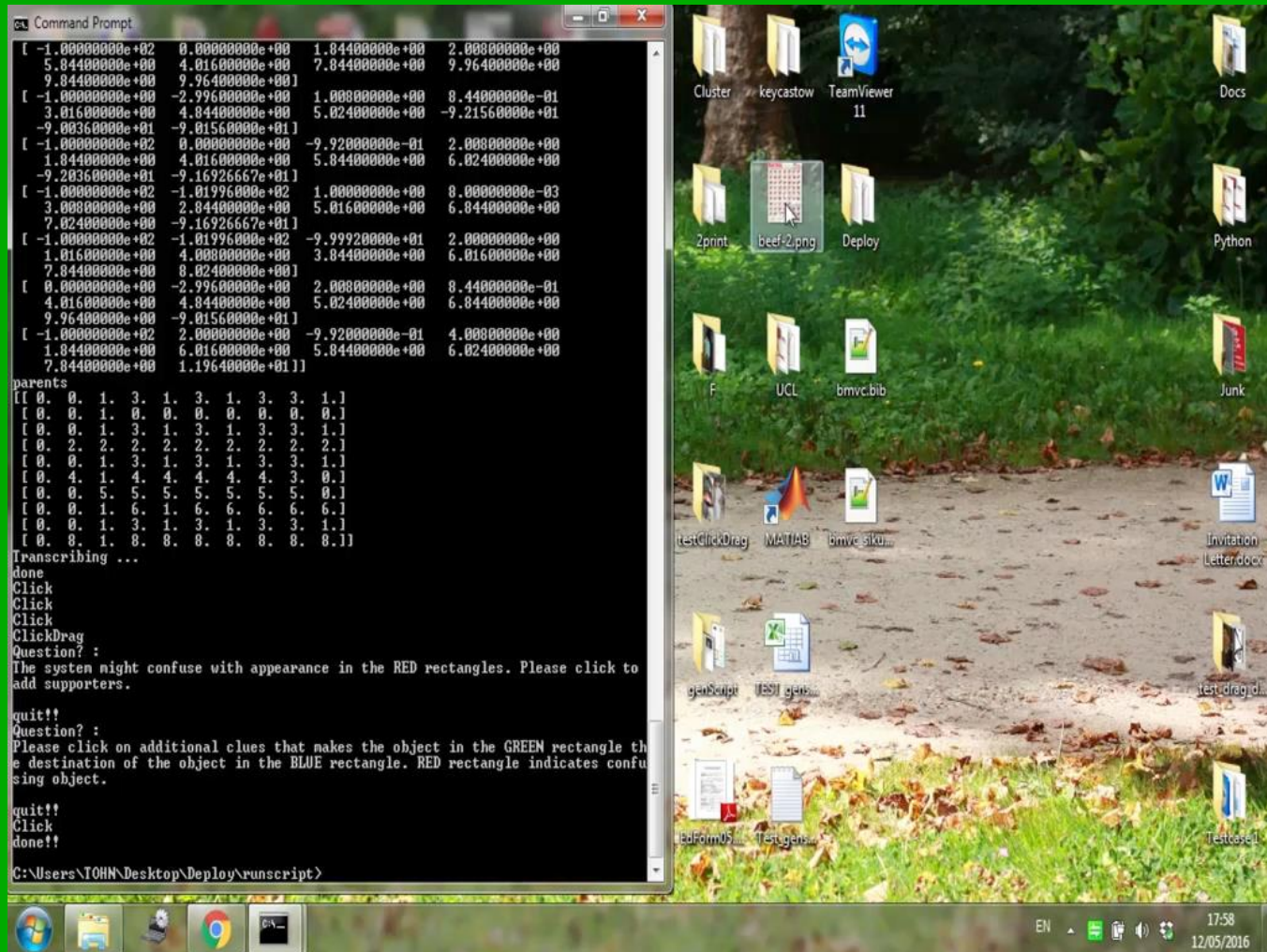
Teacher



End-user

# From User's Perspective

## Phase C. A user<sup>3</sup> triggers the script to complete the task.



Instructor



Teacher



End-user

# From development perspective

Phases

## Task Classification

		Linear	Looping	Monitoring
A: Demonstration		Normal	+Signal	+Signal
B: Teaching		Supporters	Positives False positives Supporters	Visual Cues
C: Running		Execute	Execute	Execute



# Phase A : Demonstration Phase

- Sniffer

- Video



## Log-file

```
INFO:root:6195671,000000.bmp,1405,357,no press,  
INFO:root:6195703,000001.bmp,1439,197,no press,  
INFO:root:6195734,000002.bmp,1463,168,no press,  
INFO:root:6195781,000003.bmp,1480,144,no press,  
INFO:root:6195796,000004.bmp,1482,134,no press,  
INFO:root:6195859,000005.bmp,1482,132,no press,  
INFO:root:6195906,000006.bmp,1482,119,no press,  
INFO:root:6195937,000007.bmp,1482,117,no press,  
INFO:root:6196046,000008.bmp,1483,123,no press,  
INFO:root:6196062,000009.bmp,1503,168,no press,  
INFO:root:6196093,000010.bmp,1569,226,no press,  
INFO:root:6196140,000011.bmp,1593,241,no press,  
INFO:root:6196171,000012.bmp,1595,242,no press,  
INFO:root:6196296,000013.bmp,1595,243,no press,  
INFO:root:6196343,000014.bmp,1606,292,no press,  
INFO:root:6196375,000015.bmp,1620,316,no press,  
INFO:root:6196484,000016.bmp,1620,317,no press,  
INFO:root:6196531,000017.bmp,1620,320,no press,  
INFO:root:6196562,000018.bmp,1619,321,no press,  
INFO:root:6196671,000019.bmp,1618,321,no press,  
INFO:root:6196718,000020.bmp,1608,317,no press,  
INFO:root:6196750,000021.bmp,1606,315,no press,  
INFO:root:6196843,000022.bmp,1606,315,no press,  
INFO:root:6196875,000023.bmp,1603,313,no press,  
INFO:root:6196921,000024.bmp,1602,312,no press,  
INFO:root:6196984,000025.bmp,1601,310,no press,  
INFO:root:6197125,000026.bmp,1601,310,press left,  
INFO:root:6197250,000027.bmp,1601,310,no press,  
INFO:root:6197578,000028.bmp,1596,310,no press,
```

## Sequence of basic actions

RightClick()


Click()

RightClick()

Click()

Click()

Click()

For each () in Instances List :

Drag()to()



# Phase B : Teaching phase

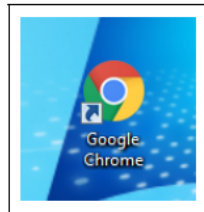
HILC gets follow-up input from user (a.k.a. Teacher):

- more positive examples
- remove false positive examples
- point out the visual cues
- add supporters

# Phase B : Teaching phase

HILC allows users to add more positive examples, remove false positive examples, point out the visual and add supporters

## Supporters



(a)



(b)

	A	B	C	D	E
1					
2					
3					
4					
5					
6					

(c)

First Name	<input type="text"/>
Last Name	<input type="text"/>
<hr/>	
Street Address Line 1	<input type="text"/>
Street Address Line 2	<input type="text"/>
City	<input type="text"/>

(d)

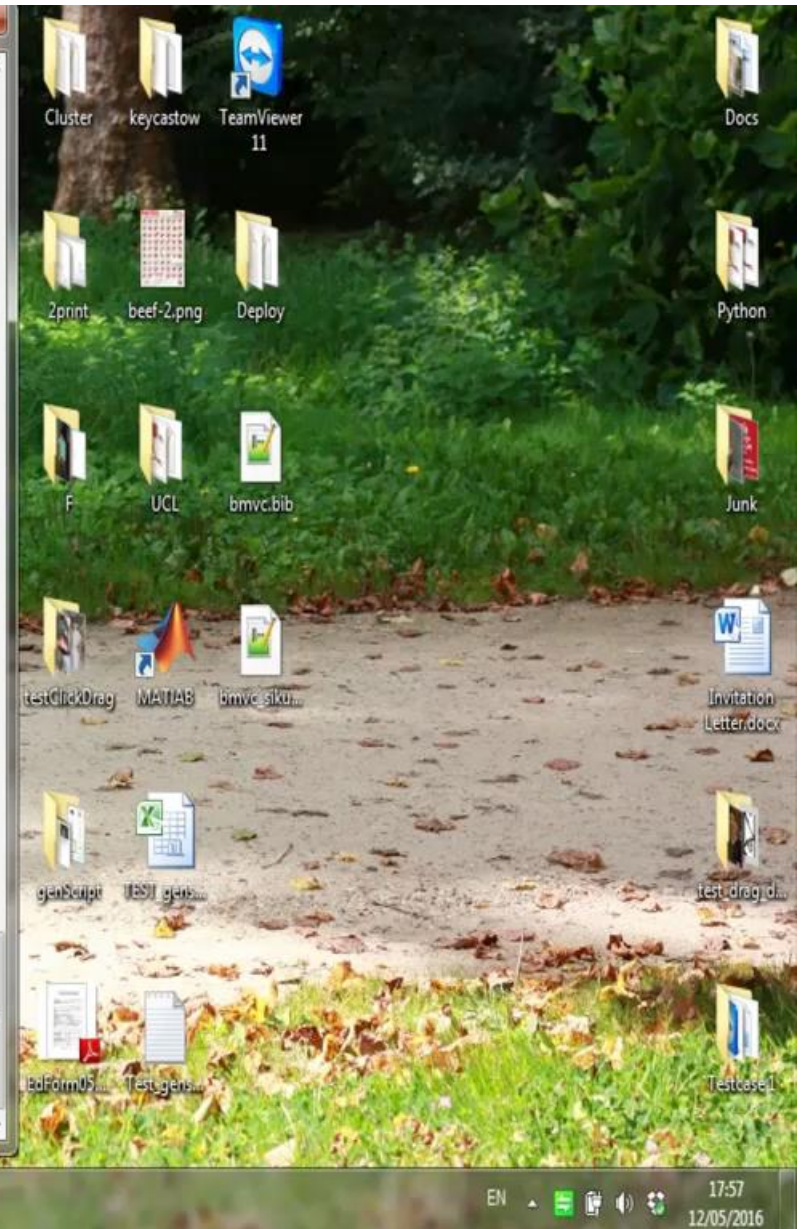
# Teaching phase - How did we do it?

When and What to ask?

# When and What to ask? (Linear tasks)

```
Command Prompt - python rerun.py --p=C:\Users\TOHN\Desktop\Deploy\USER007\002_2\

[ -1.00000000e+02  0.00000000e+00 -1.00000000e+00  0.00000000e+00
-1.00000000e+00  0.00000000e+00 -1.00000000e+00 -1.00000000e+00
-1.00000000e+00  0.00000000e+00 ]
bestPath
[ 0. 1. 0. 1. 0. 1. 2. 3. 0. 1.]
max_score
[[ 4.00000000e-03 -1.99600000e+00 2.01200000e+00 8.44000000e-01
 4.02000000e+00 4.84400000e+00 6.02400000e+00 7.84400000e+00
 1.03073333e+01 -9.01560000e+01 ]
[ -1.00000000e+02 2.00000000e+00 0.00000000e-03 4.01600000e+00
 1.84400000e+00 6.02400000e+00 5.84400000e+00 8.02400000e+00
 9.84400000e+00 1.26506667e+01 ]
[ -1.00000000e+00 -1.07600000e+00 1.00000000e+00 2.84400000e+00
 3.01600000e+00 5.84400000e+00 6.99400000e+00 7.84400000e+00
 8.96400000e+00 -9.01560000e+01 ]
[ -1.00000000e+02 0.00000000e+00 1.84400000e+00 2.00000000e+00
 5.84400000e+00 4.01600000e+00 7.84400000e+00 9.96400000e+00
 9.84400000e+00 9.96400000e+00 ]
[ -1.00000000e+00 -2.99600000e+00 1.00000000e+00 8.44000000e-01
 3.01600000e+00 4.84400000e+00 5.02400000e+00 -9.21560000e+01
 -9.00360000e+01 -9.01560000e+01 ]
[ -1.00000000e+02 0.00000000e+00 -9.92000000e-01 2.00000000e+00
 1.84400000e+00 4.01600000e+00 5.84400000e+00 6.02400000e+00
 -9.20360000e+01 -9.16926667e+01 ]
[ -1.00000000e+02 -1.01996000e+02 1.00000000e+00 8.00000000e-03
 3.00000000e+00 2.84400000e+00 5.01600000e+00 6.84400000e+00
 7.02400000e+00 -9.16926667e+01 ]
[ -1.00000000e+02 -1.01996000e+02 -9.99920000e+01 2.00000000e+00
 1.01600000e+00 4.00000000e+00 3.84400000e+00 6.01600000e+00
 7.84400000e+00 8.02400000e+00 ]
[ 0.00000000e+00 -2.99600000e+00 2.00000000e+00 8.44000000e-01
 4.01600000e+00 4.84400000e+00 5.02400000e+00 6.84400000e+00
 9.96400000e+00 -9.01560000e+01 ]
[ -1.00000000e+02 2.00000000e+00 -9.92000000e-01 4.00000000e+00
 1.84400000e+00 6.01600000e+00 5.84400000e+00 6.02400000e+00
 7.84400000e+00 1.19640000e+01 ]
parents
[[ 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.]
 [ 0. 0. 1. 0. 0. 0. 0. 0. 0. 0.]
 [ 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.]
 [ 0. 2. 2. 2. 2. 2. 2. 2. 2. 2.]
 [ 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.]
 [ 0. 4. 1. 4. 4. 4. 4. 4. 3. 0.]
 [ 0. 0. 5. 5. 5. 5. 5. 5. 5. 0.]
 [ 0. 0. 1. 6. 1. 6. 6. 6. 6. 6.]
 [ 0. 0. 1. 3. 1. 3. 1. 3. 3. 1.]
 [ 0. 0. 1. 8. 0. 8. 8. 8. 8. 8.]]
Transcribing ...
done
Click
Click
Click
```





# When and What to ask? (Looping task)

The screenshot shows a Windows desktop environment. In the background, a file explorer window is open, displaying a directory structure with files and folders. In the foreground, a command prompt window is open, showing the output of a Python script. The taskbar at the bottom displays various application icons and the system clock.

**File Explorer Window:**

- Address bar: < sniffer > Testset > USE >
- Search: Search USE
- Organize > Open > Include in library > Share with > New folder >
- Libraries: Downloads, Pictures, Recent Places, Dropbox, Desktop
- Documents: closeads, excel2bibtex, filename2excel, skipads

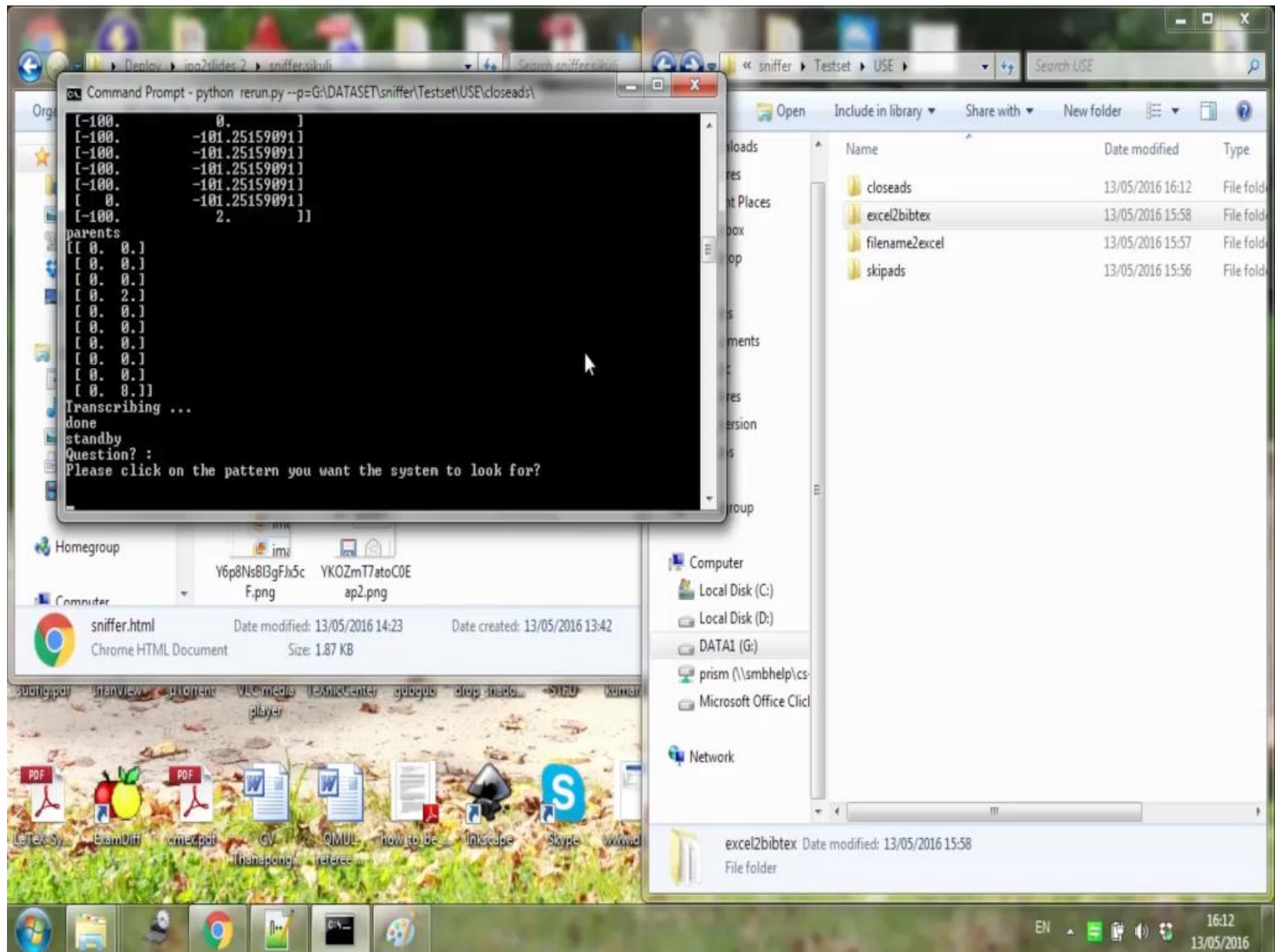
**Command Prompt Window:**

```
python run.py --p=G:\DATASET\sniffer\Testset\USE\excel2bibtex\
parents
[[ 0. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  0. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  0. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  0. 2. 1. 2. 1. 2. 1. 2. 1. 2. 1. 2. 1. 2. 1. 2.
  1. 2. 1. 2. 1. 2. 1. 2. 1. 2. 1. 2. 1. 2. 1. 2.
  0. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  0. 4. 1. 4. 1. 4. 1. 4. 1. 4. 1. 4. 1. 4. 1. 4.
  1. 4. 1. 4. 1. 4. 1. 4. 1. 4. 1. 4. 1. 4. 1. 4.
  0. 0. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5.
  5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5. 5.
  0. 0. 1. 6. 1. 6. 1. 6. 1. 6. 1. 6. 1. 6. 1. 6.
  1. 6. 1. 6. 1. 6. 1. 6. 1. 6. 1. 6. 1. 6. 1. 6.
  0. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0. 1. 0.
  0. 8. 1. 8. 1. 8. 1. 8. 1. 8. 1. 8. 1. 8. 1. 8.
  1. 8. 1. 8. 1. 8. 1. 8. 1. 8. 1. 8. 1. 8. 1. 8.]]
Transcribing ...
done
find locations of each instance in the list
```

**Taskbar:**

- System tray: EN, 16:40, 13/05/2016

# When and What to ask? (Monitoring task)





# Phase C : Running phase



The screenshot displays two overlapping windows on a Windows 7 desktop. The background window is LibreOffice Impress, titled 'New OpenDocument Presentation.odp - LibreOffice Impress'. It shows a presentation with two slides. Slide 1 is titled 'Dogs' and Slide 2 is titled 'Click to add title'. The foreground window is a Windows File Explorer, titled 'Dogs', showing a directory of files. The files are listed in a table with columns for Name, Date modified, and Type. The files are organized into Favorites, Libraries, Homegroup, Computer, and Network. The 'download (6).jpg' file is selected, and its preview is shown at the bottom of the window.

Name	Date modified	Type
download (1).jpg	12/05/2016 22:44	JPEG Im
download (2).jpg	12/05/2016 22:44	JPEG Im
download (3).jpg	12/05/2016 22:45	JPEG Im
download (4).jpg	12/05/2016 22:45	JPEG Im
download (5).jpg	12/05/2016 22:45	JPEG Im
download (6).jpg	12/05/2016 22:45	JPEG Im
download (7).jpg	12/05/2016 22:45	JPEG Im
download (8).jpg	12/05/2016 22:45	JPEG Im
download (9).jpg	12/05/2016 22:45	JPEG Im
download (10).jpg	12/05/2016 22:45	JPEG Im
download (11).jpg	12/05/2016 22:45	JPEG Im
download (12).jpg	12/05/2016 22:45	JPEG Im
download.jpg	12/05/2016 22:44	JPEG Im
images (1).jpg	12/05/2016 22:46	JPEG Im
images (2).jpg	12/05/2016 22:46	JPEG Im
images (3).jpg	12/05/2016 22:46	JPEG Im
images (4).jpg	12/05/2016 22:46	JPEG Im
images (6).jpg	12/05/2016 22:46	JPEG Im
images (7).jpg	12/05/2016 22:46	JPEG Im
images (8).jpg	12/05/2016 22:47	JPEG Im
images (9).jpg	12/05/2016 22:47	JPEG Im
images.jpg	12/05/2016 22:46	JPEG Im

download (6).jpg Date taken: Specify date taken  
JPEG Image

# Scenarios and experiments

Scenario	Basic Actions + Typing					Transcription		Reproduction		Demonstration Time VS Refining Time (average)	
	Click	Click Drag	Double Click	Right Click	Typing						
						Sikuli Slides	Our	Sikuli Slides	Our	Sikuli Slides	Our
1 Mute Audio playback (Linear)	2	0	0	0	0	✓	✓	✓*	✓	10s/49s	10s/27s
2 Turn on High-Contrast-Mode (Linear)	6	1	0	0	0	✓*	✓	✓**	✓	27s/10m	27s/170s
3 Remote access with Team Viewer (Linear)	11	0	0	0	4	✓*	✓	✗	✓	40s/∞	40s/4m
3.2 Remote access with Team Viewer (Linear)	13	0	0	4	0	✓*	✓	✗	✓	37s/∞	37s/7m
4. Skip YouTube ads (Monitoring)	1	0	0	0	0	✗	✓	✗	✓	N/A	10s/5.5m
5. Close YouTube ads (Monitoring)	1	0	0	0	0	✗	✓	✗	✓	N/A	12s/6.9m
6. Create slides out of jpgs folder (Looping)	2x	1x	0	0	0	✗	✓	✗	✓	N/A	35s/10m
7. Create spreadsheet of filenames (Looping)	4x	2x	0	0	4x	✗	✓	✗	✓	N/A	60s/6.6m
8. Create BibTex from spreadsheet (Looping)	9x	0	0	0	8x	✗	✓	✗	✓	N/A	86s/12.5m
9. Move MSWord files to a folder (Looping-Video)	4	1x	0	2	0	✗	✓	✗	✓	N/A	25s/ <b>22m</b>

**Table 1.** User study on our system compared to Sikuli Slides. Scenario 3.2 is an alternative way to perform Scenario 3, without pressing shortcut key combinations that Sikuli Slides is known to be missing. Nevertheless, we eventually realized that Sikuli Slides isn't detecting the right click actions either. (✓ = successful, ✓\* = partially successful, ✓\*\* = can be successful with guidance from the operator, ✗ = can not succeed at the task at all). x represents the number of repeated loops needed to complete the task. Please note that 90% of the refining time for Task 9 is offline - devoted to the time spent on processing video to produce the log-file.

# Limitations

- HILC is not aware of off-screen objects.  
because it only looks at a screenshot image.
- Scrolling is currently not supported.
- HILC is not aware of system state, hence it only relies on a fixed time delay between basic actions.
- HILC relies on having a sniffer, but we have preliminary results from just videos of the GUI.

# Take home messages

- **Computer Vision.**
- **3 Phases.**
- **Follow-up questions.**

**Thank you**

# Q & A



For more information and code please visit <http://visual.cs.ucl.ac.uk/pubs/HILC/>



**Extra slides**



WIKIPEDIA  
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

#### Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

#### Tools

- What links here
- Related changes
- Upload file
- Special pages
- Permanent link
- Page information
- Wikidata item
- Cite this page

Article **Talk**

Read **Edit** View history

Search Wikipedia

# Document Object Model

From Wikipedia, the free encyclopedia



This article includes a [list of references](#), but **its sources remain unclear** because it has **insufficient inline citations**. Please help to [improve](#) this article by [introducing](#) more precise citations. *(August 2010)* ([Learn how and when to remove this template message](#))

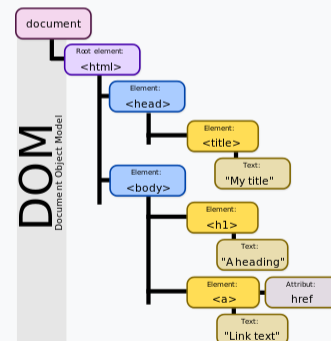
The **Document Object Model (DOM)** is a [cross-platform](#) and [language-independent](#) [application programming interface](#) that treats an [HTML](#), [XHTML](#), or [XML](#) document as a [tree structure](#) wherein each [node](#) is an [object](#) representing a part of the document. The objects can be manipulated programmatically and any visible changes occurring as a result may then be reflected in the display of the document.<sup>[2]</sup>

Principal standardization of DOM was handled by the [W3C](#), which last developed a recommendation in 2004. [WHATWG](#) took over development of the standard, publishing it as a [living document](#). The W3C now publishes stable snapshots of the WHATWG standard.

#### Contents [hide]

- History
- Standards
- Applications
  - Web browsers
  - JavaScript
- Implementations
  - Layout engines

## Document Object Model

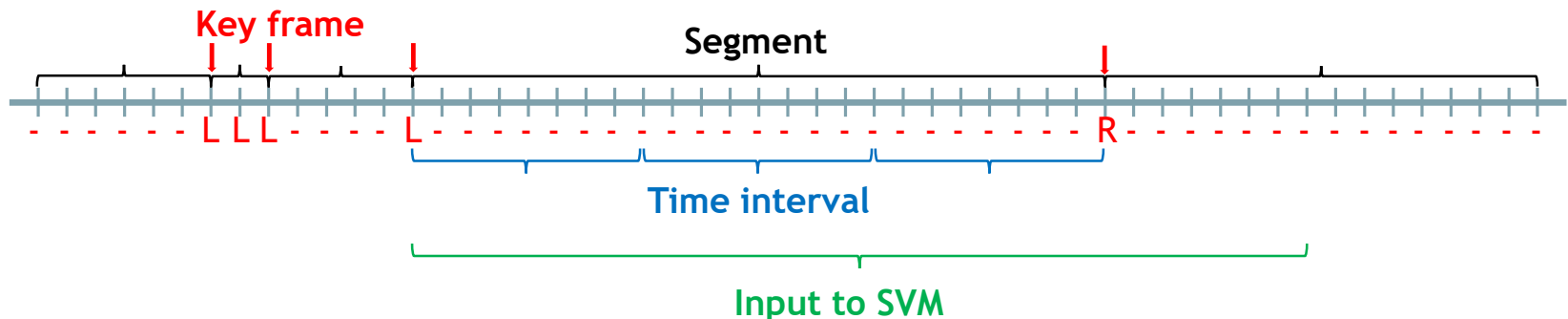


Example of DOM hierarchy in an HTML document

**First published** October 1, 1998; 18 years ago  
**Latest version** DOM4<sup>[1]</sup>

# Demonstration Phase - How did we do it?

- The log file unifies inputs from both video and sniffer
- Transcribe input log-file
  - Viterbi Algorithm
    - Unary potential is computed using SVM + Random Forest
    - Pairwise potential encourages segments from the same basic action to follow the order



Basic actions = {Click, Double Click, Click Drag, Right Click}