

# Collaborative Caption Editing System

IBM Research – Tokyo  
IBM Japan



# Enhancing the Quality of Caption Editing System

Speaker: **Ali Sobhi**

sobhi@us.ibm.com

Authors: **Ali Sobhi**  
**Reiko Nagatsuma**  
**Takashi Saitoh**

Session ID: WEB-027

Wednesday February 29, 2012  
3:10 PM – 4:10 PM

Grand Hyatt – Elizabeth F

<<original>>

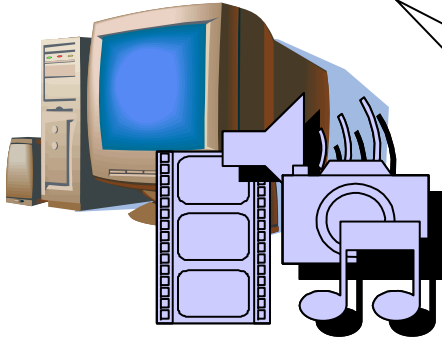
abc...

<<after recognition>>

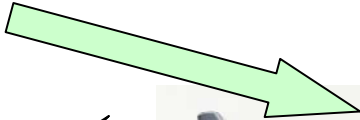
XXXXX...

Too hard to understand if there is no caption

1: Auto Speech Recognition



Language model problem

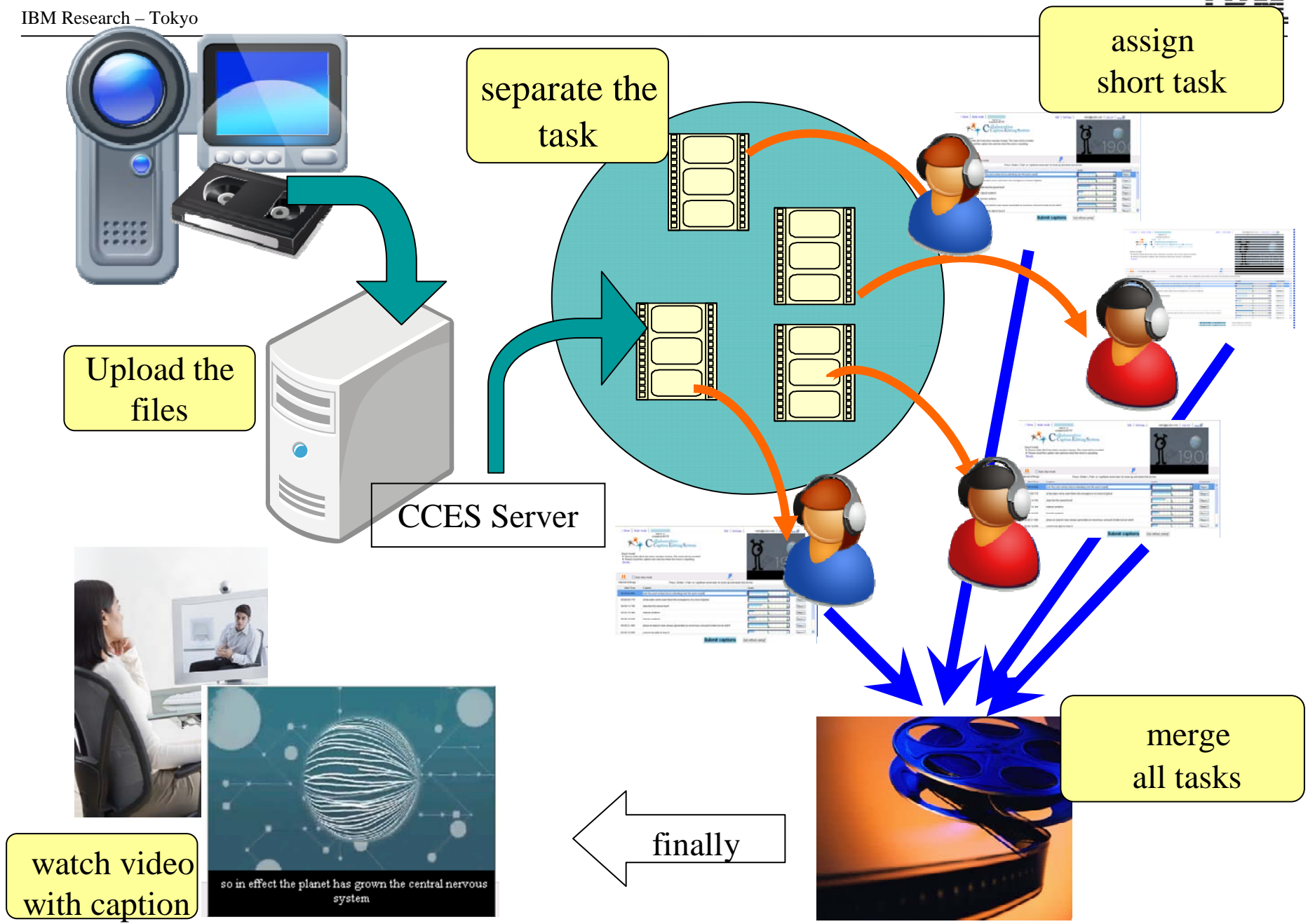


2 : Transcribe all text



Hard work to transcribe for 1 person

- Hard to set up the equipment
- Not too easy to input
- Loss of concentration



Upload the files

separate the task

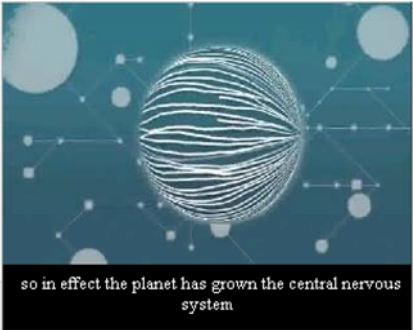
assign short task

CCES Server

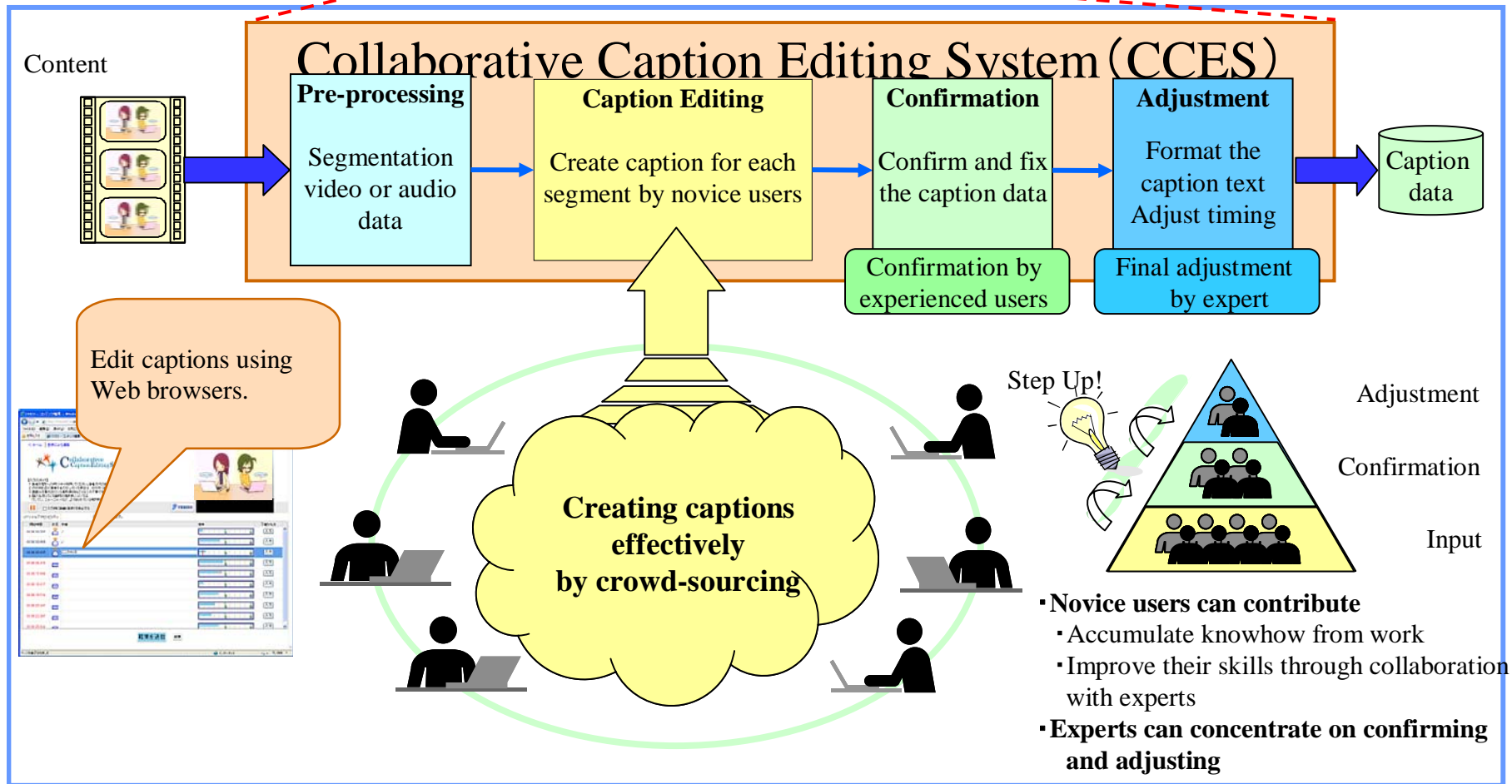
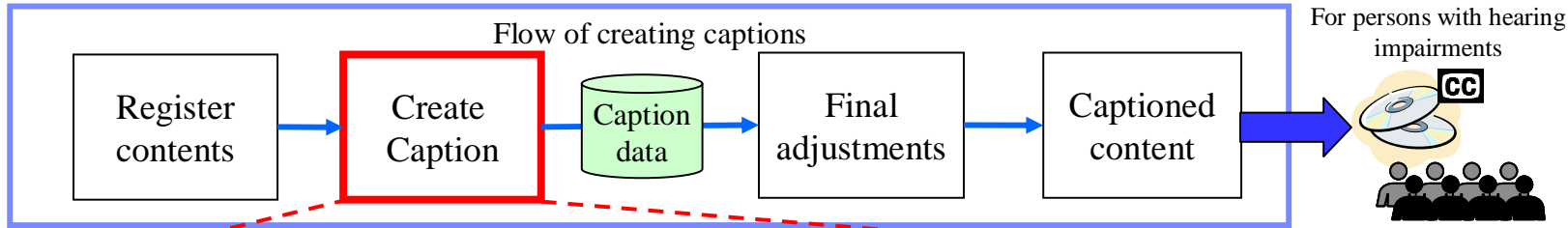
merge all tasks

finally

watch video with caption



# Overview: Collaborative Caption Editing System (CCES)



< Home | Audio mode | [Progress Bar]

Edit | Settings |

reiko@p.ibm.com | Log out | Help

degree of completion(1/12)



[Input Guide]

⚠ Please enter Alt+0 key when speaker change. The mark will be inserted.

⚠ Please insert the caption into next line when the word is separating.

[Details...](#)



Confirm the caption immediately

Auto stop mode



Internet of things

Press <Enter>, <Tab> or <up/down arrow key> to move up and down line by line

| Start Time   | Caption   | Audio | Comment |
|--------------|---|-------|---------|
| 00:00:32.950 | instrument and it's all interconnected so we can have actually have access to it                |       | Report  |
| 00:00:37.650 | so in effect the planet has grown the central nervous system                                    |       | Report  |
| 00:00:42.629 | so is about a billion people using the internet to the moment and that's set to great for me to |       | Report  |
| 00:00:47.550 | do it in the next couple of years and we think that's quite a lot                               |       | Report  |
| 00:00:51.180 | that's the last ten years i've been looking at  |       | Report  |
| 00:00:54.629 | devices being linked together using networks  |       | Report  |

Reporting function

**Submit captions**

Quit without saving?

Repeat the phrase automatically

# *Demo*

# User Pilot: IBM Centennial Volunteer Program

- Target content: Lecture video from Waseda University.
  
- Total participants : 143 (Nov. 12th – Dec. 23th)
  - 5 participants are working for confirmation task
  
- Participants who login more than 3days : 80
  - Average work items/login : 16 items (16 items x ~30 seconds each = ~8 minutes)
  - Average operation time: 45 min
  
- Active time
  - Morning: before 9:00 am
  - Lunch time: noon -1:00 pm
  - Evening: after 7:00 pm -11:00 pm

and holiday, Saturday



# User Pilot: IBM Centennial Volunteer Program

- Average operation time (educational content)
  - Input captions: 163 sec / each content (30 sec)
  - Confirmation: 82 sec / each content (30 sec)
  
- Real operation time (25 min content)
  - Input captions: 151 min
  - Confirmation: 66 min
  - Adjustment: 48 min
  - **Total:** **257 min** (~4.2 hours - an expert spends ~8 hours to caption 15 min. of content)
  
- Comments from users are the following:
  - I am pretty impressed with this...
  - Response of the system is pretty good, and user interface is very useful.
  - It is good for novice PC user because:
    - No need for special equipment (only browser)
    - Repeat one phrase automatically
    - Available to input easily with short segmentation

## ■ Learning effectiveness (sample data of 3 users)

