The image illustrates a workflow for creating a QA dataset from conference website data.

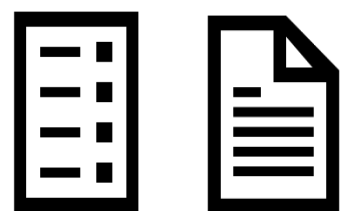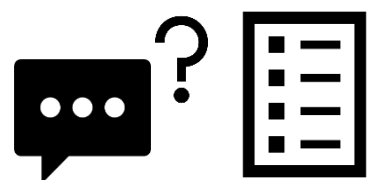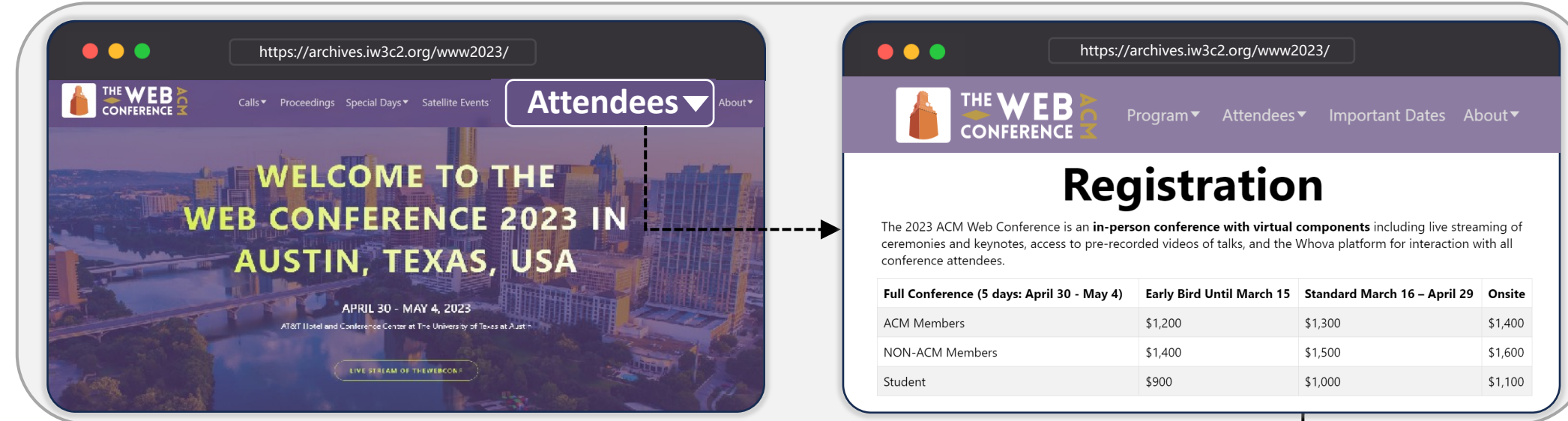**Left panel (workflow steps):**

Semi-structured data
From Conference Website

↓

Change into semi-structured
Conference dataset (json)

↓

Design **QA pairs**, based on the
Conference dataset
(classified into 4 kinds)

Right panel (website screenshots):

https://archives.iw3c2.org/www2023/

THE WEB CONFERENCE ACM

Calls▾ Proceedings Special Days▾ Satellite Events | **Attendees ▾** | About▾

WELCOME TO THE
WEB CONFERENCE 2023 IN
AUSTIN, TEXAS, USA

APRIL 30 - MAY 4, 2023

AT&T Hotel and Conference Center at The University of Texas at Aust...

LIVE STREAM OF THE WWW.CONF

https://archives.iw3c2.org/www2023/

THE WEB CONFERENCE ACM

Program▾ Attendees▾ Important Dates About▾

Registration

The 2023 ACM Web Conference is an **in-person conference with virtual components** including live streaming of ceremonies and keynotes, access to pre-recorded videos of talks, and the Whova platform for interaction with all conference attendees.

Full Conference (5 days: April 30 - May 4)	Early Bird Until March 15	Standard March 16 – April 29	Onsite
ACM Members	$1,200	$1,300	$1,400
NON-ACM Members	$1,400	$1,500	$1,600
Student	$900	$1,000	$1,100

JSON dataset block:

```
{
  "WWW2023": {...
    "Attendees": {
      "Registration": {...
        "Introduce": ["The 2023 ACM Web Conference is an in-person
          conference ... including live streaming of ceremonies and keynotes,"...],
        "Register Fee(All fees in US dollars)": {
          "Full Conference (5 days: April 30 – May 4)": [
            {  "Type": " ACM Members ",
               "Early Bird Until March 15": "$1,200", ... },
            {  "Type": " NON-ACM Members ", ...
               "Onsite": "$1,600"},
            {  "Type": " Student ",
               "Early Bird Until March 15": "$900",
               "Standard March 16 - April 29": "$1000",
               " Onsite ": "$1100"}],          ...
      "Student Travel Award": {
          "Application Deadline" : "March 31, 2023",  ... }          ... }
```

Bottom panel (four QA pair examples):

{"question": "When is the **application deadline** for the student travel grant?",
"answer": "Deadline for student travel grants is **March 31, 2023**.",
"from":"WWW2023>>Attendees>>Student Travel Award"}

Extraction_atomic

{"question": "How much does an **ACM member** have to pay if they register for the **full conference** during the **Early Bird period**?" ,
"answer": "$1200",
"from":"WWW2023>>Attendees>>Registration>>Register Fee(All fees in US dollars)"}

Extraction_complex

{
"question": "Will there be **live-streamed** sessions for those who can't attend in person?",
"answer": "Yes",
"from":"WWW2023>>Attendees>>Registration>>Introduce"
}

Reasoning_atomic

{"question": "How much do 10 **NON-ACM members** and 10 **students** have to pay if they all register **onsite** for the **full conference**?" ,
"answer": "$27000",
"from":"WWW2023>>Attendees>>Registration>>Register Fee(All fees in US dollars)"}

Reasoning_complex