# Google Chrome Extension: NewsPrettifier

Using Go, HTML, CSS, Javascript, and GCE API

第20組

# OUTLINE

1. About the extension
    a. What it is
    b. Purpose
    c. Features
2. Implementation
    a. Languages
    b. Frameworks/Libraries
    c. Execution pipeline
    d. Overview
    e. Server/Database/SQL
    f. Parsing
    g. Front-end Design
3. Demonstration

# News Prettifier

# ABOUT THE EXTENSION

# WHAT IT IS

- A Google Chrome Extension which “prettifies” certain news websites
  - Remove irrelevant information
    - Advertisements
    - Related articles
    - Links
  - Add customization to the prettified pages

**Related Article:** Climate change concerns sour Germany's love for cars

For the first time ever, this week the Greens polled in first place, usurping the CDU, according to the Forsa Trendbarometer conducted by CNN affiliate, RTL.

Merkel is nothing if not a pragmatic dealmaker. But she has tried this political combination before and failed. In the past, Merkel has said she only intends to complete her term and ruled out running for any political office after 2021. Is her successor ready to take on the challenge?

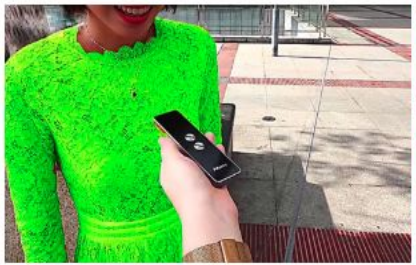
On Monday, when asked if she was prepared to become chancellor in the event of a coalition collapse, Kramp-Karrenbauer took it in her stride.

She said: "There are very good reasons to not recklessly end a government, considering the situation in Germany and Europe. We at the CDU are obliged to provide reliability and stability."

"You can be assured that the CDU is ready for any possibility that may or may not happen."

PAID CONTENT Smartfeed

Mac User Warning: The Only



Forget "Language Barriers" With

Now Playing Merkel admits need for...

Tyra Banks talks new 'SI' cover with Don Lemon

Rex Tillerson proves something very important...

Jordan's King tells Kushner Palestinians must...

Judith Kerr, author of "The Tiger Who Came to Tea...

Advertisement

By using this site, you agree to our updated Privacy Policy and our Terms of Use.

# FEATURES

- Remove irrelevant content from the news page
  - Links
  - Related news
  - Advertisement
  - Buttons
  - Anything except article content
- Customization
  - Choose which theme you want to use
  - Choose font size

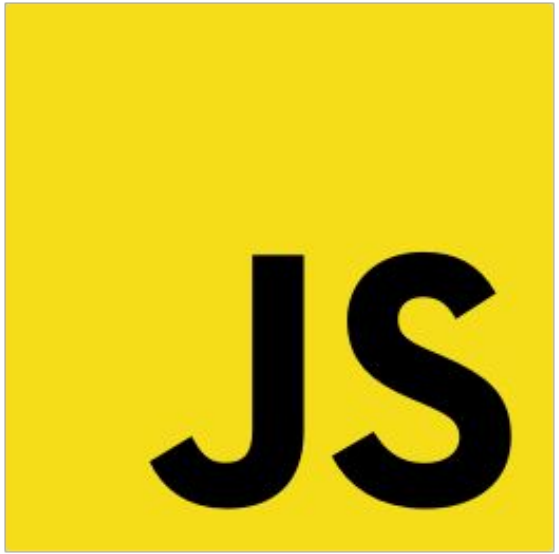
# FEATURES

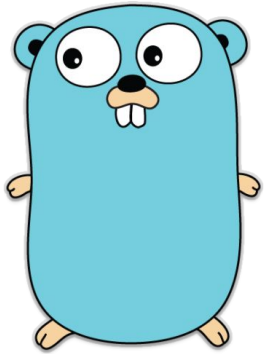
- Account
  - Users can register an account
  - Keep the customization preferences
  - Save article history
    - Can go back to pages the user viewed previously

# FEATURES: SUPPORTED WEBSITES

- ABC News
- Fox News
- Sky News
- The New York Times
- CNN
- China Post
- Taipei Times

# IMPLEMENTATION



# LANGUAGES

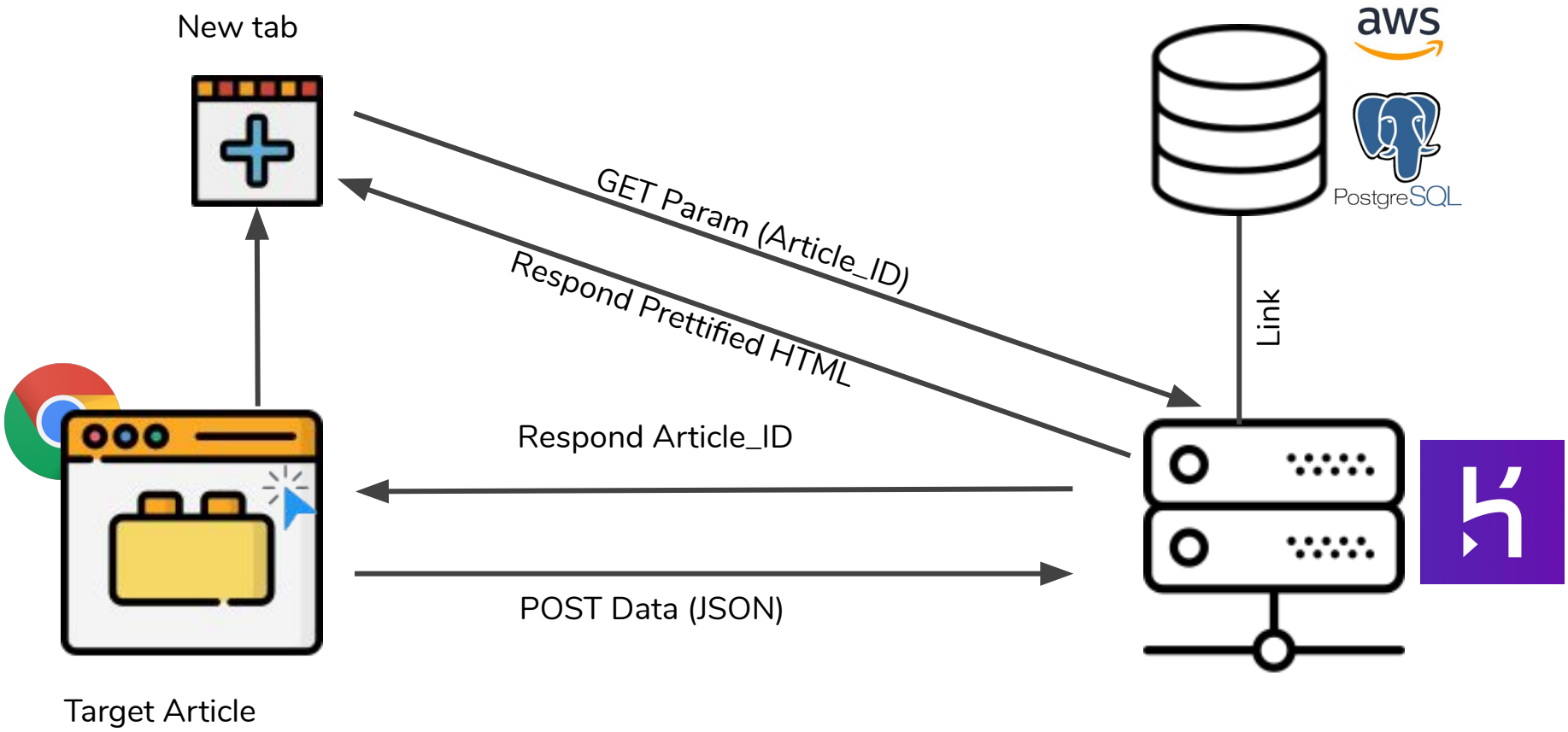
- Go
  - Backend
  - Rendering Template
- SQL
- HTML/CSS
- Javascript
  - Scraping
  - Chrome extension uses Javascript

# FRAMEWORKS/LIBRARIES

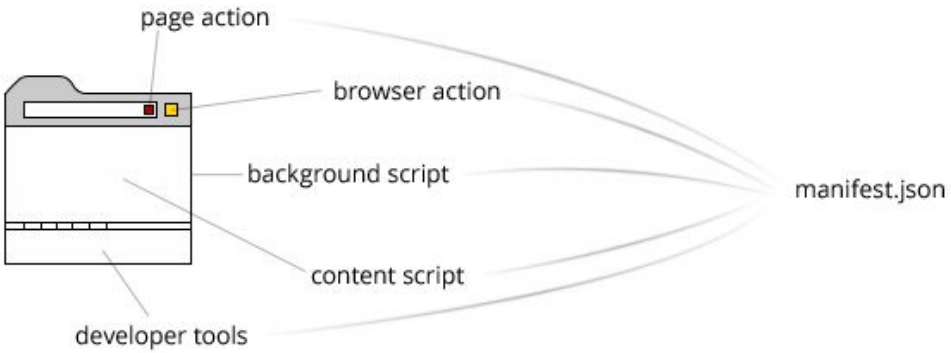
- Javascript
  - Ajax
    - HTTP Requests
  - jQuery
    - Scraping
  - Chrome Extension API
    - Used for GUI and Login system

# EXECUTION PIPELINE

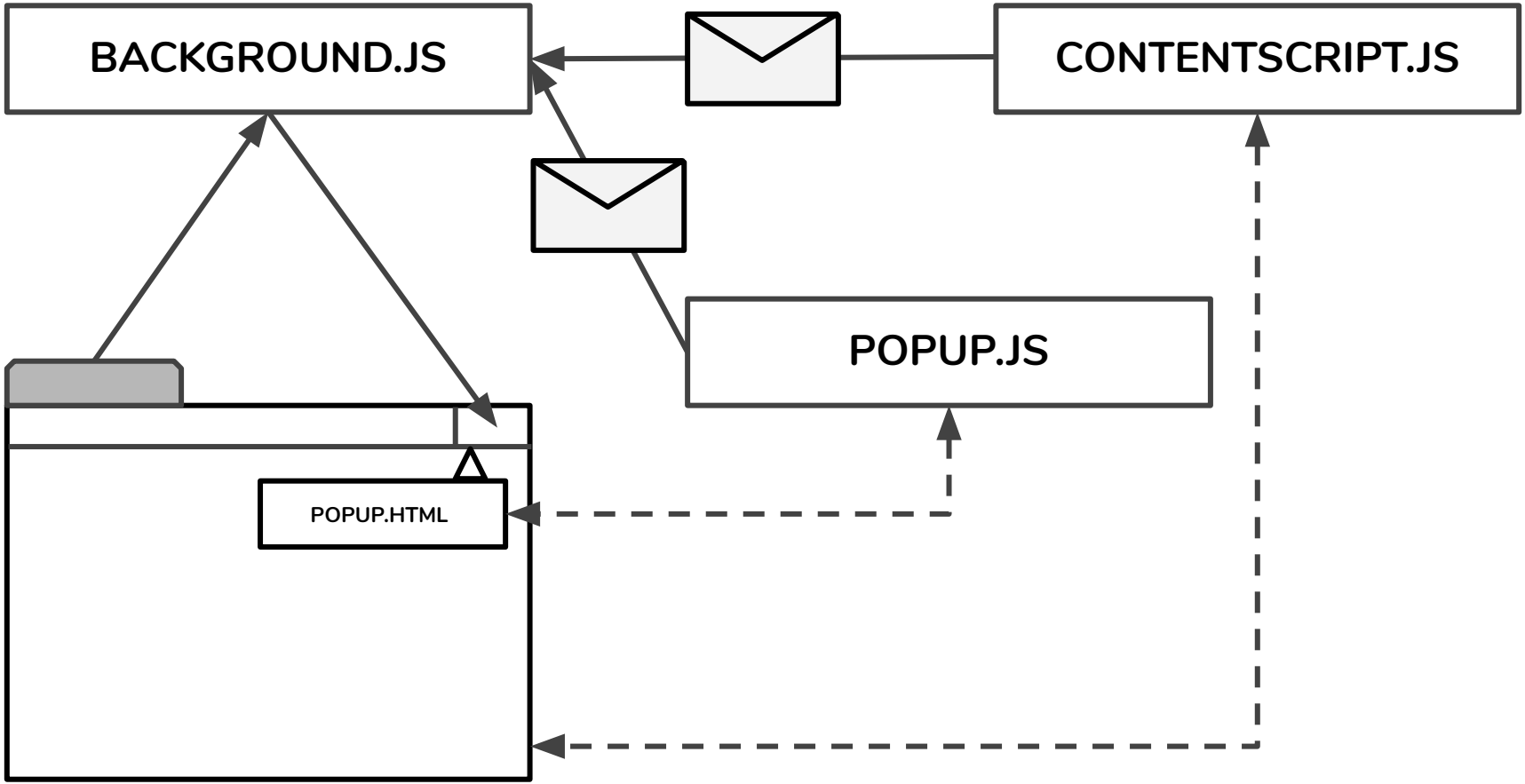
| CONTENT SCRIPT | BACKGROUND | SERVER | POPUP.JS | NEW TAB |
|---|---|---|---|---|
| Content script runs on every newly opened tab. The script will parse the news websites according to their structure, and send the parsing result (a JSON file) to Background. | Background receives the parsing result (a JSON file) from Content script, and will issue a POST request to the server (hosted on Heroku). The server will provide Background with a unique Article ID for the requested article. | Server has a route on '/article' which listens the POST request, and dispatch 'createArticle' function to that creates and saves the article in database. After that, responds a unique Article_ID back to client. | If the user presses the "Prettify" button (if it is available), popup.js will send a message to the Background, ordering it to open a new tab with the Article ID that it got from the server. | Background opens a new tab with the prettyfied news article.<br><br>Server responds the rendered prettified article to browser. |


New tab
aws
PostgreSQL
GET Param (Article_ID)
Respond Prettified HTML
Link
Respond Article_ID
POST Data (JSON)
Target Article

# CHROME EXTENSION OVERVIEW




- Controlled by Manifest.json
  - Permissions
  - Page/Browser Action
- Content Script
  - Acts on the current page, no fetch needed
- Background Script
  - For communication
- Pop Up HTML/JS
  - Shows when page action is enabled


BACKGROUND.JS
CONTENTSCRIPT.JS
POPUP.JS
POPUP.HTML

# MESSAGE PASSING

- Message passing is done through Chrome API
  - chrome.extension
- Two types of connections
  - Send a one time message (and receive optional response)
  - Establish a connection between two extension parts
    - Allows continuous message exchange
- We use messages to communicate between content.js, background.js, and popup.js

# MESSAGE PASSING: SINGLE MESSAGE

```
chrome.runtime.sendMessage({greeting: "hello"}, function(response) {
  console.log(response.farewell);
});
```

```
chrome.runtime.onMessage.addListener(
  function(request, sender, sendResponse) {
    console.log(sender.tab ?
                "from a content script:" + sender.tab.url :
                "from the extension");
    if (request.greeting == "hello")
      sendResponse({farewell: "goodbye"});
  });
```

# GO SERVER

- gorilla/mux
  - URL router and dispatcher
- gorilla/securecookie
  - User Cookie Values
- net/http
  - HTTP Implementation
    - GET, POST, UPDATE, DELETE, REDIRECT etc.
- lib/pq
  - Postgres Drive for Go
- text/template
  - Rendering Template

# DATABASE TABLE

| account | | |
|---|---|---|
| username | VARCHAR | PK |
| email | VARCHAR | NOT NULL |
| password | VARCHAR | NOT NULL |
| color | INT | DEFAULT 0 |
| size | INT | DEFAULT 0 |

| article | | |
|---|---|---|
| article_id | UUID | PK |
| title | VARCHAR | NOT NULL |
| author | VARCHAR | NOT NULL |
| content | VARCHAR | NOT NULL |
| origin | VARCHAR | NOT NULL |
| username | VARCHAR | DEFAULT “” |
| time | TIMESTAMP | CURRENT_TIMESTAMP |

# SQL FOR ACCOUNT

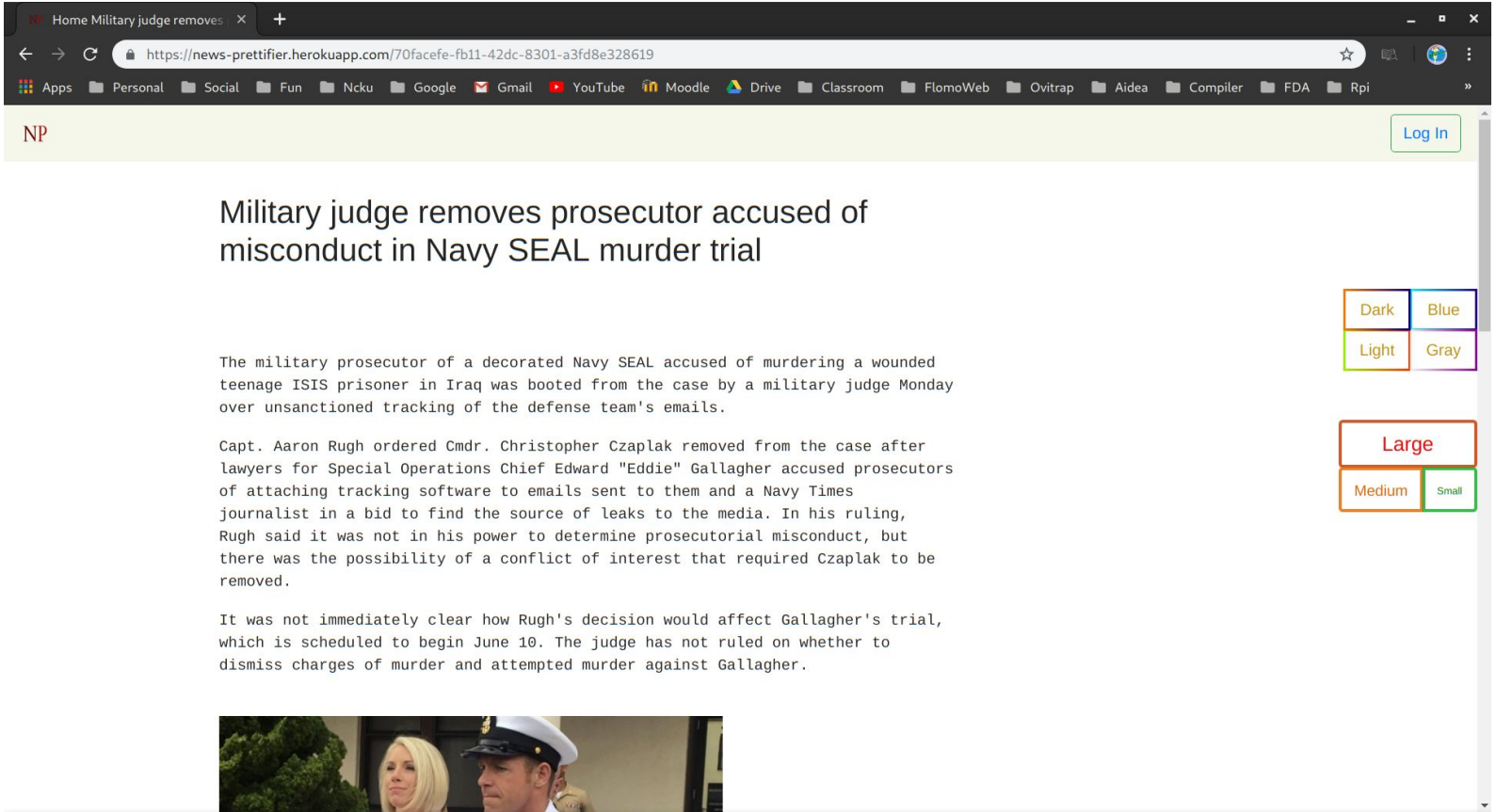
- getAccount
  - SELECT password, email, color, size FROM account where username = {};
- updateAccountSettings
  - UPDATE account SET color={}, size={} WHERE username = {};
- createAccount
  - INSERT INTO account (username, email, password) VALUES ({}, {}, {});

# SQL FOR ARTICLE

- createArticle
  - INSERT INTO article (title, author, content, origin) VALUES ({},{},{},{});
- getArticle
  - SELECT title, author, content, origin FROM article WHERE article_id={};
- updateArticleUser
  - UPDATE article SET username=$1 WHERE article_id={};
- getArticles
  - SELECT article_id, title FROM article WHERE username={};

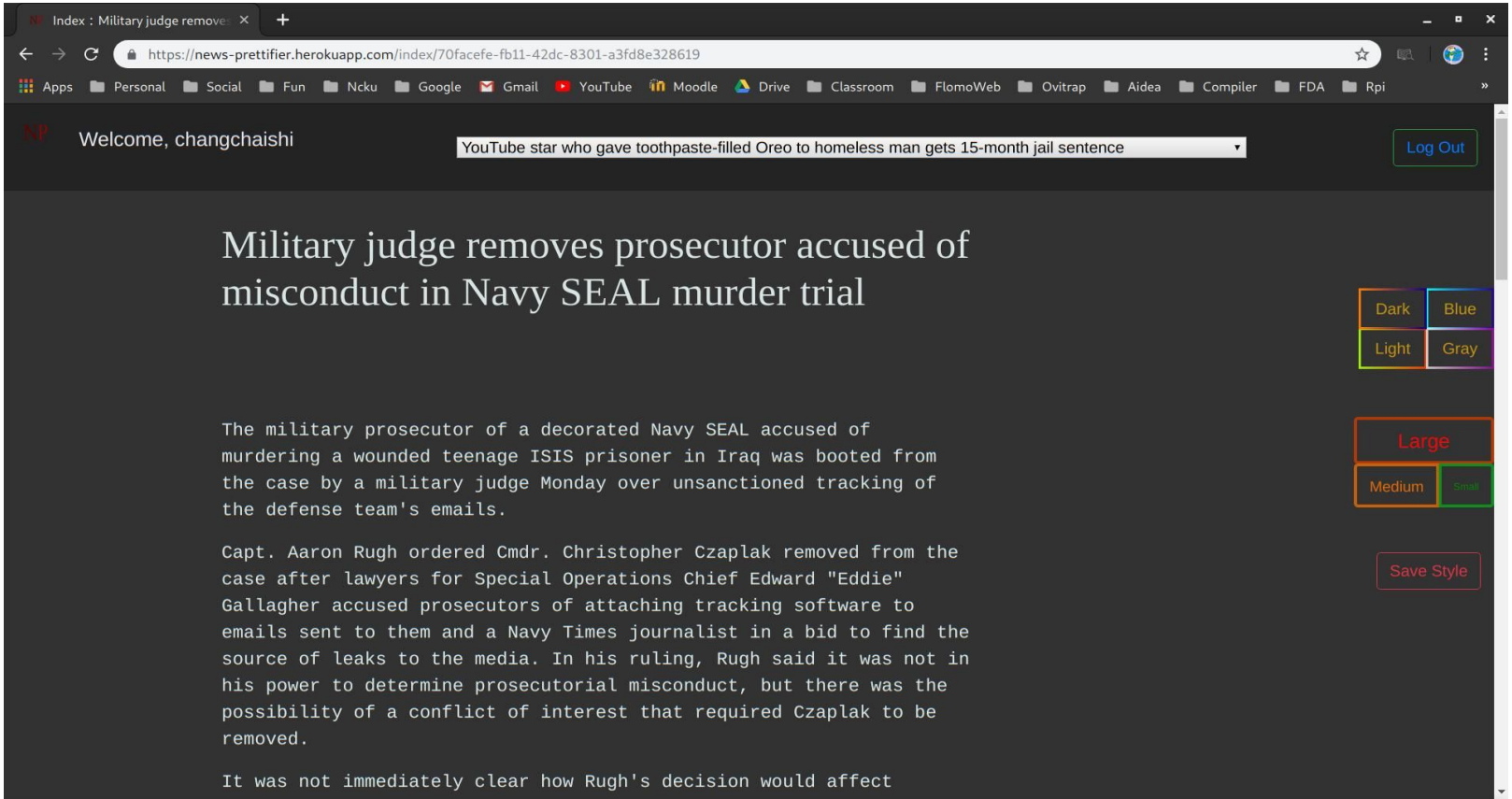
# HOME PAGE (User Logged out)

- URL: GET /{ uuid }



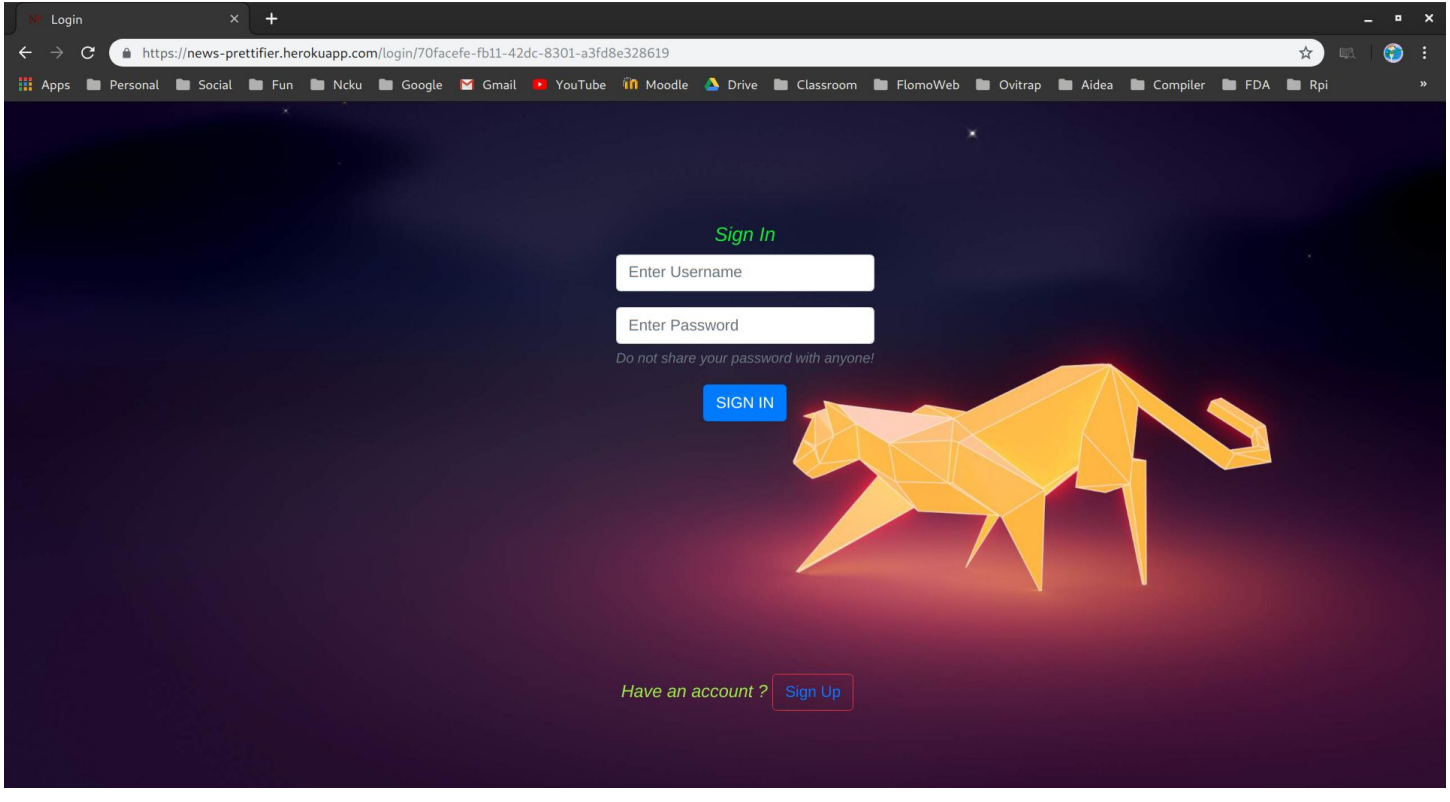
# INDEX PAGE (User Logged in)

- URL: GET /index/{ uuid }



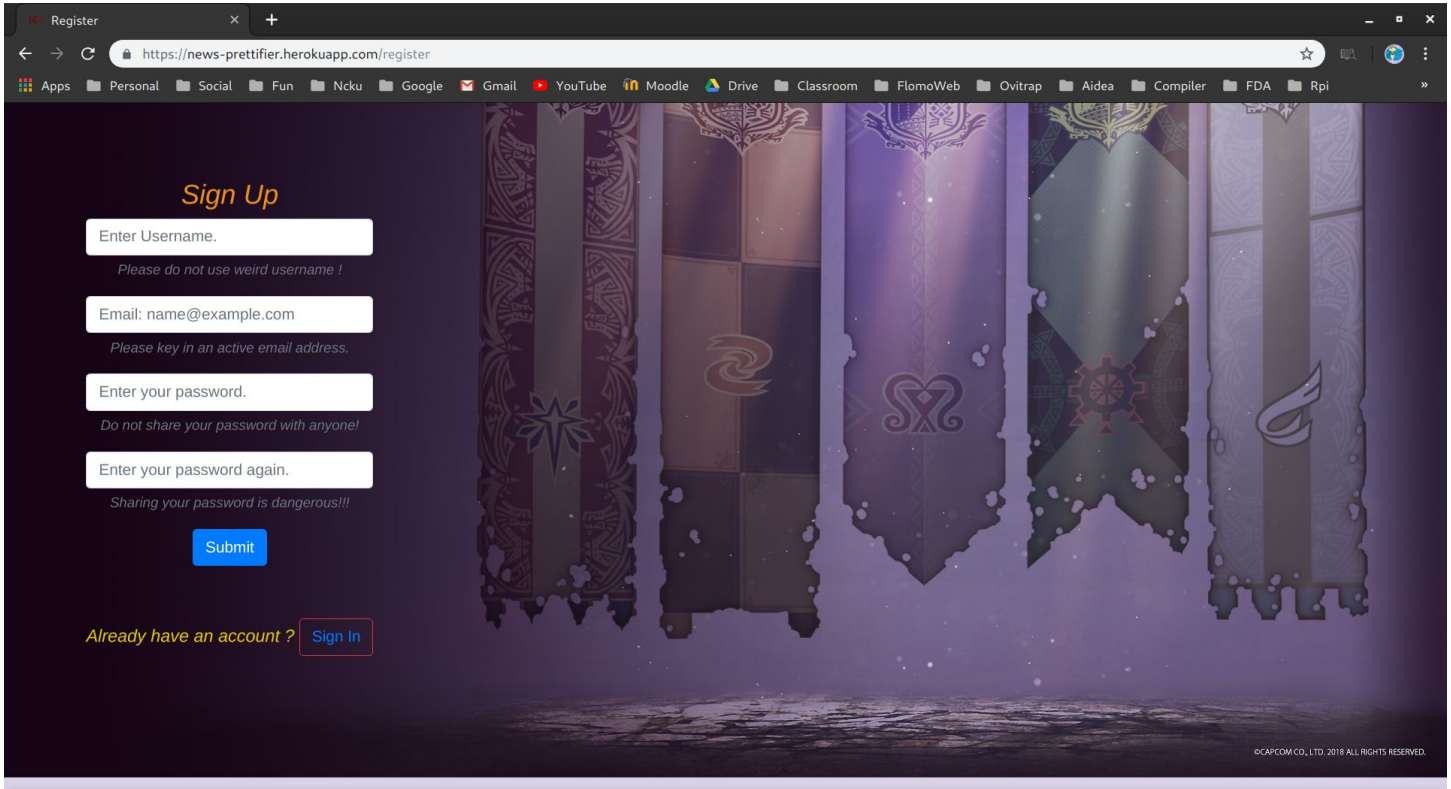
# LOGIN PAGE

- URL: GET /login/{ uuid }



# REGISTER PAGE

- URL: GET /register



# NEWS ARTICLE PARSING

- Scrape the 7 websites: ABC News, Fox News, Sky News, New York Times, CNN, China Post, Taipei Times
- jQuery

# NEWS ARTICLE PARSING: ATTRIBUTES

- URL
- Hostname
- Title
- Author(s)
- Content (article text and images)

```
var url = window.location.href;
var hostname = window.location.hostname;
var title = "";
var author = "";
var content = "";
```

Men are suing women who ac

https://edition.cnn.com/2019/06/05/europe/metoo-defamation-trials-sandra-muller-france-intl/index.html

CNN World + International Edition +

Muller, who had recently been honored as a "silence breaker" and person of the year by Time magazine, said at the time that she was surprised by Brion's change of strategy.

Writing on her Facebook page, Muller said: "The #balancetonporc movement has inspired victims to speak out, and has shed light on a real social problem that is still such a taboo."

t up."

div.el__embedded.el__embedded--fullwidth 649.17 × 422.85

French journalist Sandra Muller, who started the #balancetonporc campaign.

Vowing to "fight this until the end," Muller said that she hoped that the "trial will be an opportunity to hold a real debate on the ways to combat sexual

Elements Console Sources 54 5

```
<div class="el__embedded el__embedded--standard">…</div>
<div class="zn-body__paragraph">…</div>
<div class="zn-body__read-more">…</div>
<div class="zn-body__read-all">
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">And that's exactly what he's done.</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">"With this summons, however, they want to force me to shut up."</div>
  <div class="el__embedded el__embedded--fullwidth">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">The threat of defamation only compounds that fear.</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="el__embedded el__embedded--standard">…</div>
  <div class="zn-body__paragraph">…</div>
  <div class="zn-body__paragraph">…</div>
```

... div div div div div img.media__image.media__image--responsive

Styles Event Listeners DOM Breakpoints Properties Accessibility

Filter :hov .cls +

element.style {
}

margin –
border –

content.js

```
var title = "";
var author = "";
var content = "";

/* CNN */
if (hostname == "edition.cnn.com") {
    title = $("h1.pg-headline").text();
    author = $("span.metadata__byline__author").text();

    $("section.zn-body-text > div.l-container").children().each(function () {
        if ($(this)[0].className == "el__leafmedia el__leafmedia--sourced-paragraph") {
            content = content.concat("<p>", $(this).text(), "</p>");
        } else if ($(this)[0].className == "zn-body__paragraph speakable") {
            content = content.concat("<p>", $(this).text(), "</p>");
        }
    });
    $("div.zn-body__read-all").children().each(function () {
        if ($(this)[0].className == "zn-body__paragraph") {
            if ($(this).children("h3").get(0) != null) {
                content = content.concat("<br/><h4>", $(this).text(), "</h4><br/>");
            } else {
                content = content.concat("<p>", $(this).text(), "</p>");
            }
        } else if ($(this)[0].className == "el__embedded el__embedded--fullwidth") {
            content = content.concat("<br/><div style=\"width:500px; word-wrap:break-word; font-size:80%;\">");
            content = content.concat("<img src=\"", $(this).children("div").children("div").children("img").attr("src"), "\"");
            content = content.concat(" width=500px style=\"padding-bottom:0.5em;\" /><br/>");
            content = content.concat($(this).children("div").children("div").children("div.media__caption").text(), "</div><br/>");
        }
    });
    if ($("p.zn-body__footer").text() != "") {
        content = content.concat("<p><i>", $("p.zn-body__footer").text(), "</i></p>");
    }
}

/* New York Times */
if (hostname == "www.nytimes.com") {
```

# NEWS ARTICLE PARSING: JSON

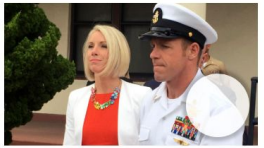
```
a_json = {
    "title": title,
    "author": author,
    "content": content,
    "origin": url
}
```

# NEWS ARTICLE PARSING: COMPLEXITY

- Subtitles, advertisements, extensive reading…
- Children of children of children…
- Inconsistent design and structure.

**WIFE OF NAVY SEAL EDDIE GALLAGHER SAYS SHE, CHILDREN 'HAD A GOOD, LONG CRY' AFTER HUSBAND'S PRE-TRIAL RELEASE**

But, Rugh chided Navy investigators who refused to testify Friday about who authorized the tracking scheme, saying the "lack of candor or cooperation in this process, I think, could be huge as a sign of culpability." The defense discovered the tracking code hidden in a suspicious logo of an American flag with a bald eagle perched on the scales of justice beneath Czaplak's signature.



**Navy SEAL Eddie Gallagher released from custody ahead of murder trial**

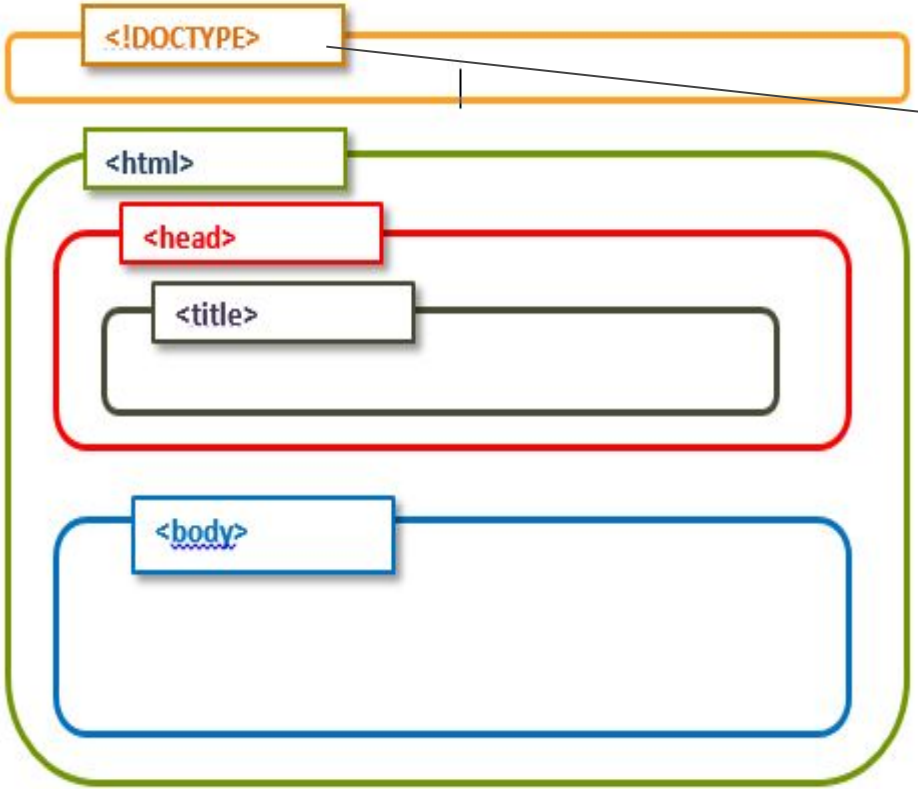
The Navy SEAL accused of killing an injured teen ISIS prisoner in Iraq has been freed from custody as a judge decides the fate of the case; Dan Springer reports.

Rugh indicated Friday he was misled about the tracking effort. He said investigators told him privately they planned to embed code in what he believed to be a court document to help them find the source of leaks but the judge said he didn't have the power to authorize such a tactic and wasn't told they planned to target emails sent to the defense lawyers or a journalist.

Gallagher is accused of stabbing to death a 15-year old ISIS fighter in Iraq in 2017, shooting two civilians in the same year and opening fire on crowds, all charges to which he has pleaded not guilty.

**CLICK HERE TO GET THE FOX NEWS APP**

# FRONT-END DESIGN




-Must be the very first thing in your HTML document.

-Not an HTML tag.

-An instruction to the web browser about what version of HTML the page is written in.

# Inside <head> </head>

- The <link> tag defines a link between a document and an external resource.
- In HTML the <link> tag has no end tag.
- The <style> tag is used to define style information for an HTML document.

<style>

.classname{styling}  example: 

#ID{styling}  css  example:
```
#titletext {
    text-align:center; }
```

</style>

# Inside <body> </body>

- The <body> element contains all the contents of an HTML document, such as text, hyperlinks, images, tables, lists, etc.
- The  tag defines a division or a section in an HTML document.
- Often used as a container for other HTML elements to style them with CSS or to perform certain tasks with JavaScript.

example:

```
<div class="container-fluid">
        <div class="row">
                <div class="col-10">

...

                </div>

        </div>

</div>
```

# Bootstrap Grid System

Bootstrap's grid system allows up to 12 columns across the page.

If you do not want to use all 12 columns individually, you can group the columns together to create wider columns:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| span 1 | span 1 | span 1 | span 1 | span 1 | span 1 | span 1 | span 1 | span 1 | span 1 | span 1 | span 1 |
| span 4 | | | | span 4 | | | | span 4 | | | |
| span 4 | | | | span 8 | | | | | | | |
| span 6 | | | | | | span 6 | | | | | |
| span 12 | | | | | | | | | | | |

Bootstrap's grid system is responsive, and the columns will re-arrange automatically depending on the screen size.

# Inside <script> </script>

- The <script> tag is used to define a client-side script (JavaScript).
- The <script> element either contains scripting statements, or it points to an external script file through the src attribute.

Example: <script src="https://code.jquery.com/jquery-3.4.1.min.js"></script>

Javascript:

```
function Large() {

document.getElementById("text1").style.fontSize="large";
}
```

# DEMONSTRATION

# THANK YOU

News
Prettifier