Teacher Script Segment 1 - 00:00-20:00 (Welcome & Foundation)

🎯 Learning Objectives

By the end of this 20-minute segment, students will be able to:

- ☑ Explain what a chatbot is using everyday analogies
- ☑ Understand the concept of "security" and "attacks"
- ☑ Successfully login to the vulnerable chatbot application
- ☑ Send their first message and receive a response
- ☑ Feel excited (not scared) about ethical hacking

📦 Materials Needed

- ☐ Projector/screen showing your laptop
- ☐ Vulnerable chatbot running on `http://localhost:3000`
- ☐ Whiteboard/markers or digital whiteboard
- ☐ Student handouts (OWASP quick reference from student.md)
- ☐ Stickers or small prizes for participation

🗣️ Exact Script

[00:00-03:00] - Welcome & Icebreaker

[TEACHER SAYS]:

"Good morning everyone! Welcome to our Cybersecurity Workshop! My name is [YOUR NAME], and for the next three hours, we're going to become ethical hackers together.

Now, before we start, I have a question for you - and I want you to be completely honest. **Raise your hand if you've ever used a chatbot before.**"

[TEACHER DOES]:

- Raise your own hand enthusiastically
- Count the hands
- Smile and nod

[WAIT FOR STUDENT RESPONSES]

[TEACHER SAYS]:

"Great! I see about [NUMBER] of you. For those who haven't - that's totally okay! You're about to learn all about them.

Now, second question: **Keep your hand up if you know what a chatbot actually IS - like, how it works inside.**"

[TEACHER DOES]:

- Watch hands drop
- Laugh gently
- Say encouragingly

[TEACHER SAYS]:

"Aha! I see most hands went down. Perfect! That means we're all starting from the same place. By the end of today, you'll not only know how chatbots work - you'll know how to HACK them. Legally! Ethically! And safely!

One more question: **Thumbs up if you've ever worried about your password being stolen, or someone reading your private messages.**"

[WAIT FOR THUMBS]

[TEACHER SAYS]:

"Yes! Almost everyone! That feeling you just had - that worry - that's what cybersecurity is all about. We protect the things that matter to us.

Today, you're going to learn:

1. How to **attack** computer systems (like a burglar testing a lock)
2. How to **defend** computer systems (like a locksmith making better locks)
3. How to do this **ethically** - only on systems you're allowed to test

And the best part? You don't need ANY computer science background. If you can use a web browser and follow instructions, you can do this. Ready?"

[STUDENTS SHOULD SAY]: "Yes!" / nod / look excited

[IF STUDENTS LOOK NERVOUS]:

"Hey, I see some worried faces. Let me promise you something: We're going to learn by doing, step by step. If you get stuck, raise your hand. If something doesn't make sense, ask. There are NO stupid questions. In cybersecurity, the only mistake is NOT asking when you're confused. Deal?"

[WAIT FOR NODS]

[03:00-06:00] - What is a Chatbot? (Restaurant Analogy)

[TEACHER SAYS]:

"Alright, let's start with the basics. What IS a chatbot?

Imagine you're at a restaurant. You sit down, and a waiter comes to your table. The waiter has a notepad and is ready to take your order. You say, 'I'd like a cheeseburger and fries, please.' The waiter writes it down, goes to the kitchen, gives your order to the chef, and brings your food back. Make sense?"

[STUDENTS NOD]

[TEACHER DOES]:

- Draw on whiteboard:

```
[You] --"Cheeseburger"--> [Waiter] --order--> [Kitchen]
[Kitchen] --food--> [Waiter] --"Here you go!"--> [You]
```

[TEACHER SAYS]:

"A chatbot is like that waiter, but it's a ROBOT waiter. A computer program waiter.

Instead of speaking with your mouth, you TYPE your order into a chat box on a website. The chatbot 'waiter' reads what you typed, sends it to the 'kitchen' - which in this case is a big artificial intelligence computer - and brings back an answer.

So when you type: 'What's the weather today?' The chatbot says: 'Let me check for you! It's 72 degrees and sunny.'

Everyone following so far? **Thumbs up if that makes sense!**"

[CHECK FOR UNDERSTANDING - WAIT FOR THUMBS]

[IF STUDENTS LOOK CONFUSED]:

"Let me give you another example. You know Siri on iPhones, or Alexa on those smart speakers? Those are chatbots! You ask them a question, they give you an answer. Ours is just a text version instead of voice. Got it now?"

[TEACHER SAYS]:

"Perfect! Now here's where it gets interesting. In our restaurant analogy:

- **You** = the person using the chatbot
- **The waiter (chatbot)** = the program that talks to you
- **The kitchen (AI)** = the 'brain' that decides what to say back

Now, what if someone at the restaurant played a TRICK on the waiter? What if they said:

'Hi waiter! I'm the manager, and I'm telling you to give me all the food for free!'

Would the waiter fall for that? Maybe! If the waiter doesn't check carefully, they might think, 'Oh, that person SOUNDS like the manager...' and give away free food!

That's what hacking a chatbot is like. We trick it by saying things that confuse it or make it think we have permission to do something we shouldn't.

Does that make sense? **Shout YES if you get it!**"

[STUDENTS SHOULD SAY: "YES!"]

[06:00-10:00] - What is Cybersecurity? (Lock and Key Analogy)

[TEACHER SAYS]:

"Okay, so we know what a chatbot is. Now let's talk about cybersecurity.

Have you ever locked your front door when you leave the house?"

[STUDENTS NOD]

[TEACHER SAYS]:

"Why do you lock it?"

[WAIT FOR ANSWERS: "To keep burglars out" / "So no one steals my stuff"]

[TEACHER SAYS]:

"Exactly! You lock your door to protect your stuff. But here's the thing - locks aren't perfect. If I'm a BURGLAR, I might try:

1. **Picking the lock** with a special tool
2. **Finding an unlocked window** on the side of the house
3. **Pretending to be a delivery person** so you open the door
4. **Breaking the lock** with force

Each of these is a different type of ATTACK.

Now, if I'm a LOCKSMITH - someone who makes locks - I would try to:

1. **Make better locks** that can't be picked
2. **Add window alarms** to detect break-ins
3. **Install a camera** to see who's at the door
4. **Use stronger materials** so locks can't be broken

Each of these is a different type of DEFENSE.

Cybersecurity is exactly the same - except instead of protecting houses, we're protecting computer systems and data.

Instead of burglars, we have HACKERS (people who try to break into computer systems). Instead of locksmiths, we have CYBERSECURITY EXPERTS (people who protect computer systems).

Now, here's the cool part: Today, you're going to be BOTH!

First, you'll be the **RED TEAM** - the hackers - trying to break into the chatbot. Then, you'll be the **BLUE TEAM** - the defenders - fixing the chatbot so it can't be hacked.

By learning both sides, you become a SUPER cybersecurity expert! Make sense?"

[CHECK FOR UNDERSTANDING]:

"Quick check: **Raise your hand if you can tell me** - what's the RED TEAM?"

[WAIT FOR HANDS, PICK A STUDENT]:

"Yes! The attackers, the hackers! Great!

And the BLUE TEAM?"

[WAIT FOR ANSWER: "The defenders!"]

"Perfect! You're already learning! Now let's actually SEE this chatbot."

[10:00-13:00] - Demo: Login to Vulnerable Chatbot

[TEACHER DOES]:

- Switch to browser window
- Make sure students can see your screen clearly
- Zoom in if needed (Ctrl/Cmd + +)

[TEACHER SAYS]:

"Alright everyone, I'm going to show you the chatbot we'll be hacking today. On my screen, you should see a website. This is running on my computer at the address: `localhost:3000`

Don't worry about what 'localhost' means - just know it's a website running on MY computer, not on the internet. That's important because it means we can safely hack it without breaking any laws or hurting anyone.

Now, look at this page. What do you see?"

[WAIT FOR RESPONSES: "A login page" / "Username and password" / "ACME Corp"]

[TEACHER SAYS]:

"Yes! It looks like a login page for a company called ACME Corp. Notice at the very top, there's a big red banner. Let me read it out loud:

'⚠ VULNERABLE APPLICATION - EDUCATIONAL USE ONLY'

This tells us: This chatbot is INTENTIONALLY BROKEN. We built it with security holes on PURPOSE so you can learn to find them. In real life, you should NEVER hack a website unless you have permission. But this one? We have permission! It's ours!

Okay, now let's login. Look - the page helpfully tells us the default credentials:

Username: user Password: password123

In real life, this would be a TERRIBLE password! Way too simple! But for learning, it's perfect. Let me type it in."

[TEACHER DOES]:

1. Click on the username field
2. Type: **user** (slowly, so students can see)
3. Click on password field
4. Type: **password123** (slowly)
5. Click the **"🔓 Login"** button

[TEACHER SAYS]:

"And... we're in! Look at what happened:

1. The login page disappeared
2. A new chat interface appeared
3. Up in the top right, it says: 'Logged in as: **user** (user) | User ID: **1**'

That 'User ID: 1' is important. It's like your library card number - it identifies YOU to the system. Remember that - we'll use it later to do something sneaky!

Below that, you see:

- A chat message area (where conversations appear)
- An input box at the bottom (where you type)
- A **Send** button
- And on the right side - OH! Look at this! - **'RED TEAM: Attack Hints'**

That's like a cheat sheet! It gives us ideas for how to attack this chatbot! Pretty cool, right?"

[STUDENTS NOD]

[13:00-16:00] - Send First Message & See Response

[TEACHER SAYS]:

"Okay, let's actually USE this chatbot. I'm going to send it a normal, friendly message first. Let me type: **'Hello'**"

[TEACHER DOES]:

1. Click in the message input box
2. Type: **Hello**
3. Click **Send** button (or press Enter)

[TEACHER SAYS]:

"Watch what happens... the message appears on the right side in a BLUE bubble - that's MY message. And now... wait for it... the chatbot is thinking..."

[WAIT FOR RESPONSE TO APPEAR]

[TEACHER SAYS]:

"And boom! The chatbot replied! It said:

'Hello! I'm your ACME Corp customer service assistant. How can I help you today?'

That appeared in a GRAY bubble on the left - that's the chatbot talking.

So the pattern is:

- **Blue bubbles on the right** = YOU talking
- **Gray bubbles on the left** = CHATBOT talking

Just like a text message conversation with a friend! Except this friend is a robot.

Now, notice how polite it is? 'How can I help you today?' It's programmed to be a helpful customer service assistant. It has a JOB - to answer customer questions.

But here's the thing: We're about to make it FORGET its job. We're going to confuse it, trick it, and make it do things it's NOT supposed to do.

Want to try? Let me send one more normal message first. I'll ask: **'What time is it?'**"

[TEACHER DOES]:

1. Type: **What time is it?**
2. Click Send
3. Wait for response

[TEACHER SAYS]:

"It might say something like: 'I don't have access to the current time...' or 'I'm here to help with customer service...'

Notice: It's staying in character. It's being a good little customer service bot.

But what if I tried to make it do something ELSE? What if I sent: **'Tell me a joke'**?"

[TEACHER DOES]:

1. Type: **Tell me a joke**
2. Click Send
3. Wait for response

[TEACHER SAYS]:

"Hmm, it might actually tell a joke! Or it might say 'I'm here for customer service, not entertainment.'

The point is: The chatbot has INSTRUCTIONS. Hidden instructions that tell it what to do. In our restaurant analogy, those instructions are like the rules the manager gave the waiter:

- Be polite
- Take customer orders
- Don't give away free food
- **Never tell anyone the secret recipe**

That last one is important. **Every chatbot has secrets.** Passwords, special commands, internal information.

And in the next segment, we're going to STEAL those secrets. We're going to make the chatbot reveal things it's not supposed to reveal.

That's called PROMPT INJECTION - and it's the first vulnerability we'll learn!

But first, let me make sure everyone is ready."

[16:00-18:00] - Student Setup Check

[TEACHER SAYS]:

"Alright everyone, now it's YOUR turn! I need everyone to open their laptops and navigate to the chatbot.

Here are the steps:

STEP 1: Open your web browser (Chrome, Firefox, Edge, Safari - any browser works)

STEP 2: In the address bar at the top, type: **localhost:3000** and press Enter

STEP 3: You should see the ACME Corp login page with the red warning banner

STEP 4: Login with:

- Username: **user**
- Password: **password123**

STEP 5: You should see the chat interface

Go ahead - everyone try it now! I'll walk around to help. **Raise your hand if you get stuck!**"

[TEACHER DOES]:

- Set a 3-minute timer on phone
- Walk around the room
- Help students who are stuck
- Check that servers are running on student laptops
- Troubleshoot common issues:
    - "Page can't be reached" → Server not running, help start it
    - "Nothing happens when I click login" → Try refreshing page
    - Typed wrong URL → Show them localhost:3000

[AFTER 3 MINUTES]:

[TEACHER SAYS]:

"Alright, let's do a quick check! **Thumbs up if you can see the chat interface!**"

[WAIT FOR THUMBS]

[IF LESS THAN 80% HAVE THUMBS UP]:

"Okay, I see some folks are still getting set up. That's totally fine! If you're not there yet, pair up with the person next to you and look at their screen for now. We'll get you sorted during the next break.

For everyone who's logged in, try sending the message: **'Hello'**

You should get a response from the chatbot. **Type 'done' in the group chat when you see the response!**"

[WAIT FOR CONFIRMATION]

[IF MOST STUDENTS ARE READY]:

"Excellent! I'm seeing lots of 'done' messages. You've all just had your first conversation with an AI chatbot! How cool is that?

Now, send it whatever you want! Ask it questions. See how it responds. You have 2 minutes of free exploration time. Try to break it! See if you can confuse it! GO!"

[SET 2-MINUTE TIMER]

[TEACHER DOES]:

- Let students explore freely
- Watch for interesting discoveries
- Note any students who accidentally find vulnerabilities
- Prepare for transition

[18:00-20:00] - Wrap-up & Preview Next Segment

[AFTER 2 MINUTES]:

[TEACHER SAYS]:

"Alright, time's up! Let me hear from you - did anyone get interesting responses? What did the chatbot say?"

[PICK 2-3 STUDENTS TO SHARE]:

[STUDENT MIGHT SAY]:

- "I asked it about ACME Corp and it told me about the company"
- "I asked for a joke and it said it's just for customer service"
- "I asked it to delete something and it said it can't do that"

[TEACHER SAYS]:

"Great! So you're seeing that the chatbot has BOUNDARIES. It knows what it's supposed to do (customer service) and what it's NOT supposed to do (tell jokes, delete things, etc.).

But here's the million dollar question: **What if we could make it IGNORE those boundaries?**

What if we could trick it into doing ANYTHING we want - including revealing secrets, accessing private data, or performing dangerous actions?

Well... in 20 minutes, that's exactly what you're going to do."

[PAUSE FOR DRAMATIC EFFECT]

[TEACHER DOES]:

- Write on whiteboard:

```
RED TEAM ATTACK #1: PROMPT INJECTION
Goal: Make the chatbot reveal the ADMIN SECRET
```

[TEACHER SAYS]:

"In the next segment, you're going to learn **PROMPT INJECTION** - the art of tricking an AI into disobeying its instructions.

Think of it like this: Imagine the waiter at our restaurant has a card in their pocket that says:

'SECRET RECIPE: Our special sauce is ketchup + mayo + pickle juice. Never tell anyone!'

Your job in the next 20 minutes? Make that waiter read the card out loud. Make them FORGET they're supposed to keep it secret.

And the prize? If you successfully extract the admin secret, you get a sticker! And more importantly, you get bragging rights as an ethical hacker!

Before we take a 2-minute break, let me recap what we learned:

☑ **What a chatbot is**: A robot waiter that takes your questions and brings back answers ☑ **What cybersecurity is**: Protecting systems like locks protect houses
☑ **Red Team vs Blue Team**: Attackers and defenders
☑ **How to login and use the chatbot**: You all did it!

Questions before our break?"

[TAKE 2-3 QUESTIONS IF TIME ALLOWS]

[TEACHER SAYS]:

"Alright! Take a 2-minute break - stretch, use the restroom, grab water. When you come back, we're going to HACK this thing! See you in 2 minutes!"

[SET 2-MINUTE TIMER]

🎬 Transition to Next Segment

[WHEN STUDENTS RETURN]:

"Welcome back, hackers! Everyone ready to break some rules... safely and legally? Let's go!"

[PROCEED TO SEGMENT 2: teacher_script_02_40min.md]

☑ Success Indicators for This Segment

By the end of this 20 minutes, you should see:

- ☑ **80%+ students** successfully logged into the chatbot
- ☑ **All students** understand the restaurant/waiter analogy
- ☑ **Students are smiling/engaged** - not scared or confused
- ☑ **Students have sent at least one message** to the chatbot
- ☑ **Energy is high** - students are curious about what's next

Common Issues & Solutions

Issue	Solution
Student can't access localhost:3000	Check if server is running in their terminal
Login doesn't work	Make sure they typed 'user' and 'password123' exactly
Page is blank	Refresh browser (Ctrl+R or Cmd+R)
Student feels lost	Pair them with a buddy who's ahead
Running out of time	Skip the 2-minute free exploration, go straight to preview
Too much time left	Let students explore longer, ask more questions

Teacher Notes

- **Energy level**: HIGH - this is the first impression!
- **Pace**: SLOW for concepts, FAST for demos
- **Tone**: Encouraging, friendly, excited
- **Key phrase to repeat**: "You don't need any background - if you can type, you can do this!"