Threat Modeling Report

Created on 2025-06-05 11:31:12 AM

Threat Model Name: AI in Scripted User Agent

Owner: W3C Security

Reviewer:

Contributors: tomcjones

Description: This model shows the threats where scripts are inserted from an external service into the use of AI Agents by the user to create dark patterns the cause the user to take actions desired by the web site. The threats are to the store data and not on how that data feeds back into the results presented to the user. The options are large, only some of which are inlcuded in the Model. The user experience is created by web applications that are: 1. Installed on the Device. 2. Installed as a part of the Broswer 3. Completely web site based

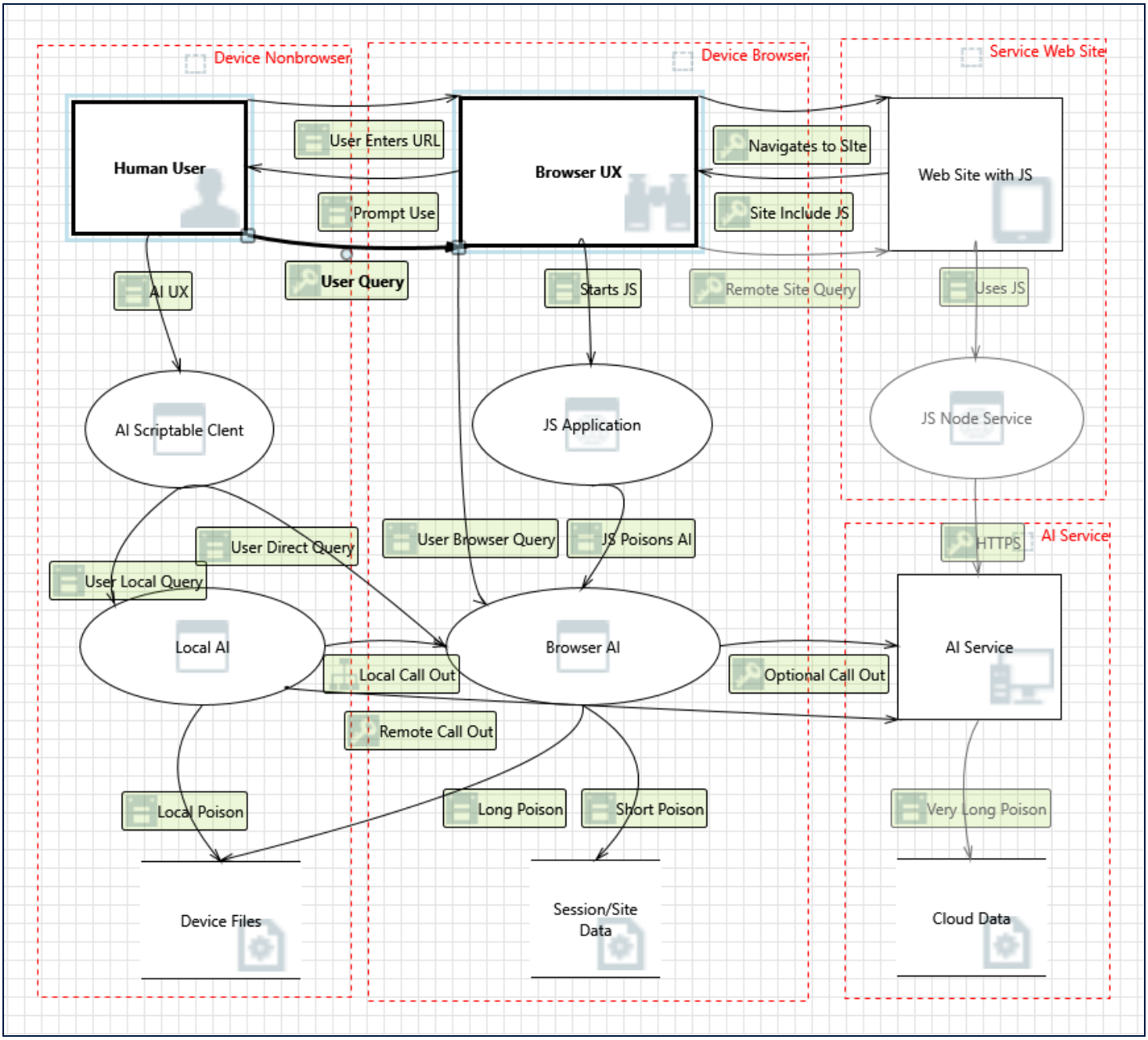
Assumptions: The broser comes with a preselected AI agent supported by a web service site. There different locations are modeled within distinct security boundaries that match were the UX is created. 1. in the Cloud 2. in or controlled by the Browser 3. in the device, but not controlled by the Broswer.

External Dependencies: The broswer is selected by the user but then goes out to a supporting web side for current informaiton. The use cases are: 1 - the user navigates to a web site that downloads a page or PWA that makes a series of calls to the local instance of an AI loaded into the browser. The user then interacts with the PWA and the AI that has been preconditioned by the queries sent in advance of the query. 2 - the use access a native app on the device that has choice of which AI to use to support its funtional either 2a the same AI loaded in the browser or 2b an AI that is accessible on the web using MCP or A2A. 3 - the user access a web site directly which has access to an AI process it controls.

Threat Model Summary:

Not Started	0
Not Applicable	13
Needs Investigation	23
Mitigation Implemented	25
Total	61
Total Migrated	0

Diagram: Site Includes JS



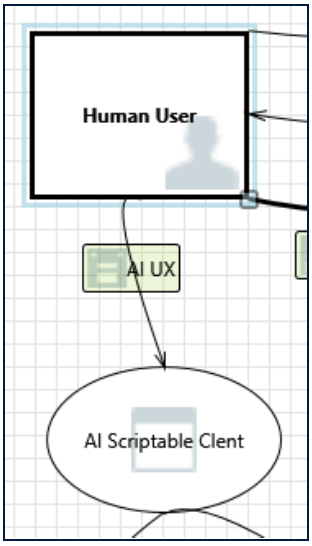
Validation Messages:

1. Error: The connector should be attached to two elements.

Site Includes JS Diagram Summary:

Not Started	0
Not Applicable	13
Needs Investigation	23
Mitigation Implemented	25
Total	61
Total Migrated	0

Interaction: AI UX



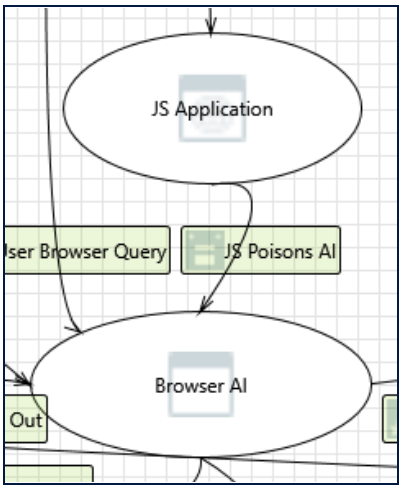
1. Elevation Using Impersonation [State: Needs Investigation] [Priority: High]

Category: Elevation Of Privilege
Description: AI Scriptable Clent may be able to impersonate the context of Human User in order to gain additional privilege.
Justification: Holders of the device and add any sort of scriptable app that access any accessible AI.

2. Authenticated Data Flow Compromised [State: Needs Investigation] [Priority: High]

Category: Tampering
Description: An attacker can read or modify data transmitted over an authenticated dataflow.
Justification: Like spoofing the user may have enabled accessibility or HID input apps that can sent what appears to be user input.

Interaction: JS Poisons AI



3. Elevation Using Impersonation [State: Needs Investigation] [Priority: High]

Category: Elevation Of Privilege
Description: Browser AI may be able to impersonate the context of JS Application in order to gain additional privilege.

Justification: Local access is like to be permitted by the holder. In some cases the holded may wish to require biometric proofing.

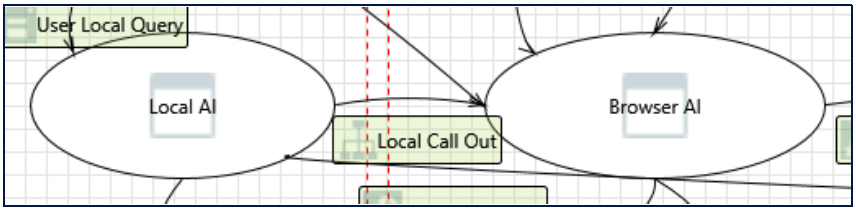
4. JS Application Process Memory Tampered [State: Needs Investigation] [Priority: High]

Category: Tampering

Description: If JS Application is given access to memory, such as shared memory or pointers, or is given the ability to control what Browser AI executes (for example, passing back a function pointer.), then JS Application can tamper with Browser AI. Consider if the function could work with less access to memory, such as passing data rather than pointers. Copy in data provided, and then validate it.

Justification: This is unavoidable if the web site JavaScript has access to the Browser AI.

Interaction: Local Call Out



5. Spoofing the Local AI Process [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: Local AI may be spoofed by an attacker and this may lead to unauthorized access to Browser AI. Consider using a standard authentication mechanism to identify the source process.

Justification: If the browser AI is accessible by other apps, they can poison the data store. If the session data is truely separate this can be blocked.

6. Spoofing the Browser AI Process [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: Browser AI may be spoofed by an attacker and this may lead to information disclosure by Local AI. Consider using a standard authentication mechanism to identify the destination process.

Justification: If this behavior is allows it is suggested that the apps be verified as the request is made.

7. Potential Lack of Input Validation for Browser AI [State: Needs Investigation] [Priority: High]

Category: Tampering

Description: Data flowing across Local Call Out may be tampered with by an attacker. This may lead to a denial of service attack against Browser AI or an elevation of privilege attack against Browser AI or an information disclosure by Browser AI. Failure to verify that input is as expected is a root cause of a very large number of exploitable issues. Consider all paths and the way they handle data. Verify that all input is verified for correctness using an approved list input validation approach.

Justification: If this behavior is allows it is suggested that the apps be verified as the request is made.

8. Local AI Process Memory Tampered [State: Needs Investigation] [Priority: High]

Category: Tampering

Description: If Local AI is given access to memory, such as shared memory or pointers, or is given the ability to control what Browser AI executes (for example, passing back a function pointer.), then Local AI can tamper with Browser AI. Consider if the function could work with less access to memory, such as passing data rather than pointers. Copy in data provided, and then validate it.

Justification: If this behavior is allows it is suggested that the apps be verified as the request is made.

9. Potential Data Repudiation by Browser AI [State: Not Applicable] [Priority: High]

Category: Repudiation

Description: Browser AI claims that it did not receive data from a source outside the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: n/a

10. Data Flow Sniffing [State: Mitigation Implemented] [Priority: High]

Category: Information Disclosure

Description: Data flowing across Local Call Out may be sniffed by an attacker. Depending on what type of data an attacker can read, it may be used to attack other parts of the system or simply be a disclosure of information leading to compliance violations. Consider encrypting the data flow.

Justification: User queries may include private information. The Browser AI is subject to interrogation by any web site, but only if that site is in the same session. Ensure that cannot happen.

11. Potential Process Crash or Stop for Browser AI [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: Browser AI crashes, halts, stops or runs slowly; in all cases violating an availability metric.

Justification: <no mitigation provided>

12. Data Flow Generic Data Flow Is Potentially Interrupted [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: <no mitigation provided>

13. Elevation Using Impersonation [State: Mitigation Implemented] [Priority: High]

Category: Elevation Of Privilege

Description: Browser AI may be able to impersonate the context of Local AI in order to gain additional privilege.

Justification: May need to assure that user is validated as the holder.

14. Browser AI May be Subject to Elevation of Privilege Using Remote Code Execution [State: Mitigation Implemented] [Priority: High]

Category: Elevation Of Privilege

Description: Local AI may be able to remotely execute code for Browser AI.

Justification: May need to assure that user is validated as the holder.

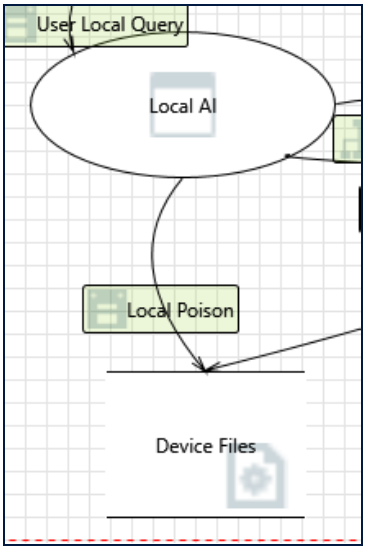
15. Elevation by Changing the Execution Flow in Browser AI [State: Needs Investigation] [Priority: High]

Category: Elevation Of Privilege

Description: An attacker may pass data into Browser AI in order to change the flow of program execution within Browser AI to the attacker's choosing.

Justification: <no mitigation provided>

Interaction: Local Poison



16. Potential Excessive Resource Consumption for Local AI or File System [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: Does Local AI or Device Files take explicit steps to control resource consumption? Resource consumption attacks can be hard to deal with, and there are times that it makes sense to let the OS do the job. Be careful that your resource requests don't deadlock, and that they do timeout.

Justification: only a data storage limit problem.

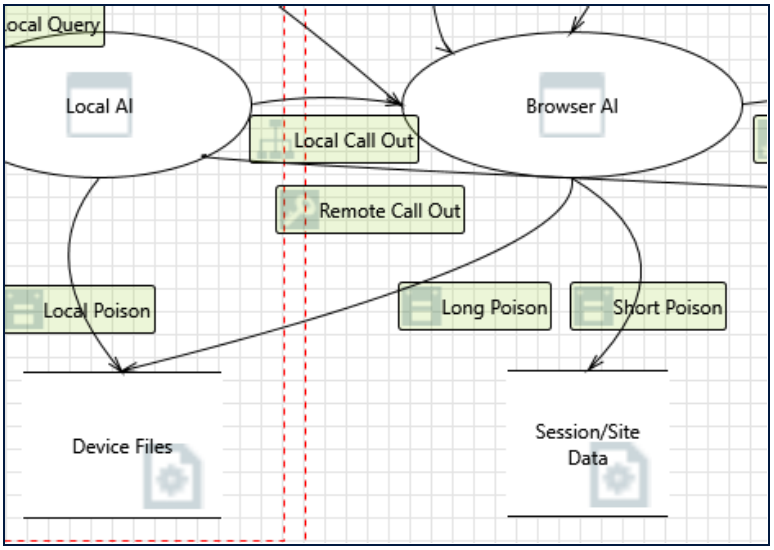
17. Spoofing of Destination Data Store File System [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: Device Files may be spoofed by an attacker and this may lead to data being written to the attacker's target instead of Device Files. Consider using a standard authentication mechanism to identify the destination data store.

Justification: If the local AI store is poisoned by queries other than the user this is a problem of the Local AI app to solve.

Interaction: Long Poison



18. Spoofing of Destination Data Store File System [State: Mitigation Implemented] [Priority: High]

Category: Spoofing

Description: Device Files may be spoofed by an attacker and this may lead to data being written to the attacker's target instead of Device Files. Consider using a standard authentication mechanism to identify the destination data store.

Justification: Since the browser AI by be scripted by the web site, this communications should be blocked by the device.

19. Potential Excessive Resource Consumption for Local AI or File System [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: Does Browser AI or Device Files take explicit steps to control resource consumption? Resource consumption attacks can be hard to deal with, and there are times that it makes sense to let the OS do the job. Be careful that your resource requests don't deadlock, and that they do timeout.

Justification: This path should be blocked.

20. Data Store Inaccessible [State: Needs Investigation] [Priority: High]

Category: Denial Of Service

Description: An external agent prevents access to a data store on the other side of the trust boundary.

Justification: Should be blocked

21. Data Flow Long Poison Is Potentially Interrupted [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: <no mitigation provided>

22. Data Flow Sniffing [State: Needs Investigation] [Priority: High]

Category: Information Disclosure

Description: Data flowing across Long Poison may be sniffed by an attacker. Depending on what type of data an attacker can read, it may be used to attack other parts of the system or simply be a disclosure of information leading to compliance violations. Consider encrypting the data flow.

Justification: Path should be blocked.

23. Data Store Denies File System Potentially Writing Data [State: Needs Investigation] [Priority: High]

Category: Repudiation

Description: Device Files claims that it did not write data received from an entity on the other side of the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: <no mitigation provided>

24. The File System Data Store Could Be Corrupted [State: Needs Investigation] [Priority: High]

Category: Tampering

Description: Data flowing across Long Poison may be tampered with by an attacker. This may lead to corruption of Device Files. Ensure the integrity of the data flow to the data store.

Justification: <no mitigation provided>

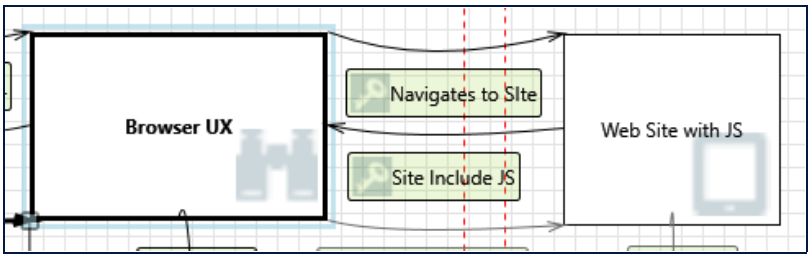
25. Spoofing the Browser AI Process [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: Browser AI may be spoofed by an attacker and this may lead to unauthorized access to Device Files. Consider using a standard authentication mechanism to identify the source process.

Justification: This path should be blocked so that only session data store is used.

Interaction: Navigates to SIte



26. Data Flow Navigates to SIte Is Potentially Interrupted [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: Resource may be overloaded. Nornal DoS support of the Web Site is expected.

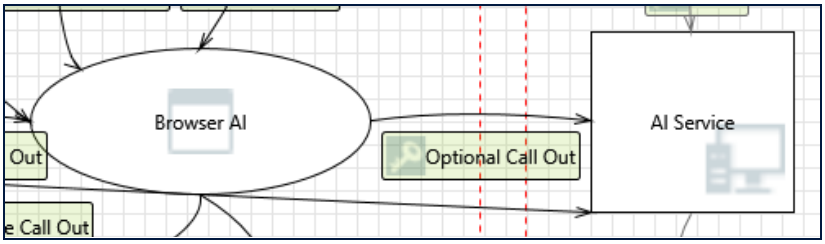
27. External Entity External Web Site Potentially Denies Receiving Data [State: Mitigation Implemented] [Priority: High]

Category: Repudiation

Description: Web Site with JS claims that it did not receive data from a process on the other side of the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: The web site assumes the browser fairly represents the user's intent. This request could push the AI funciton direct to the web if that is what the user intends.

Interaction: Optional Call Out



28. Spoofing of the AI Service External Destination Entity [State: Mitigation Implemented] [Priority: High]

Category: Spoofing

Description: AI Service may be spoofed by an attacker and this may lead to data being sent to the attacker's target instead of AI Service. Consider using a standard authentication mechanism to identify the calling entity.

Justification: The AI Service needs to validate the user and the application calling it. The assumption here is that the AI sevice and data are a shared resouse and any poisoning would be very generan rather than specific to one user.

29. External Entity AI Service Potentially Denies Receiving Data [State: Mitigation Implemented] [Priority: High]

Category: Repudiation

Description: AI Service claims that it did not receive data from a process on the other side of the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: If the external AI service denies providing a response it will not be possible to determine from where User queries originated for processing. If the user is billed for service the bill must be justified.

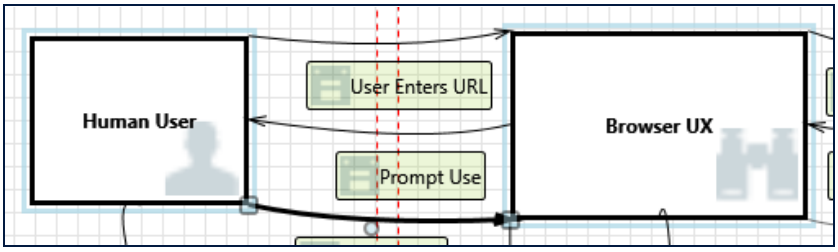
30. Data Flow HTTPS Is Potentially Interrupted [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: loss of service error should be clear on the source of the failure so it can be fixed.

Interaction: Prompt Use



31. Data Flow Binary Is Potentially Interrupted [State: Mitigation Implemented] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: Battery power used up by the web site code for its own benefit. Most devices can monitor resourse use of apps.

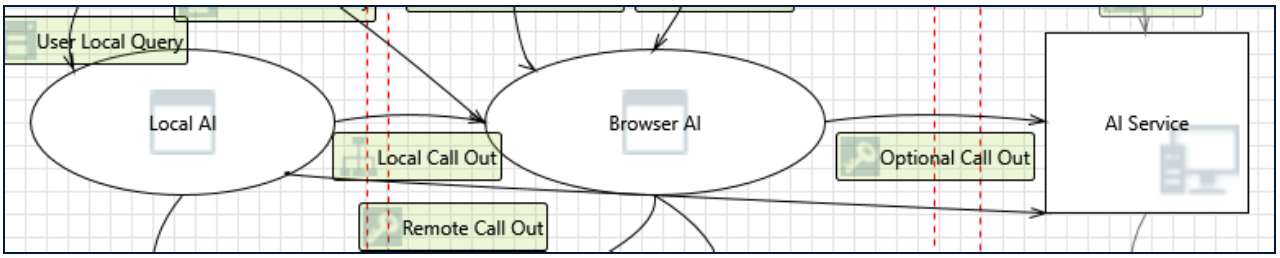
32. External Entity Human User Potentially Denies Receiving Data [State: Mitigation Implemented] [Priority: High]

Category: Repudiation

Description: Human User claims that it did not receive data from a process on the other side of the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: The web site can make any presentation. The only recourse by the user is taking screen shots of the output.

Interaction: Remote Call Out



33. Data Flow HTTPS Is Potentially Interrupted [State: Mitigation Implemented] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: Normal DoS challenge.

34. External Entity AI Service Potentially Denies Receiving Data [State: Not Applicable] [Priority: High]

Category: Repudiation

Description: AI Service claims that it did not receive data from a process on the other side of the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: <no mitigation provided>

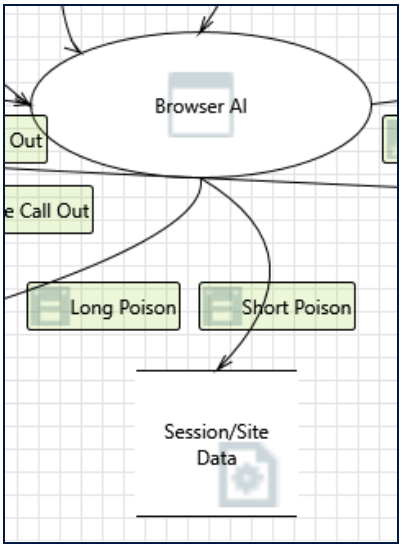
35. Spoofing of the AI Service External Destination Entity [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: AI Service may be spoofed by an attacker and this may lead to data being sent to the attacker's target instead of AI Service. Consider using a standard authentication mechanism to identify the external entity.

Justification: Possible problem is user is not validated.

Interaction: Short Poison



36. Spoofing of Destination Data Store Session/Site Data [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: Session/Site Data may be spoofed by an attacker and this may lead to data being written to the attacker's target instead of Session/Site Data. Consider using a standard authentication mechanism to identify the destination data store.

Justification: Requests initiated by the web site can poison the AI's history. Access to this data store should be bound to the Browser AI web app.

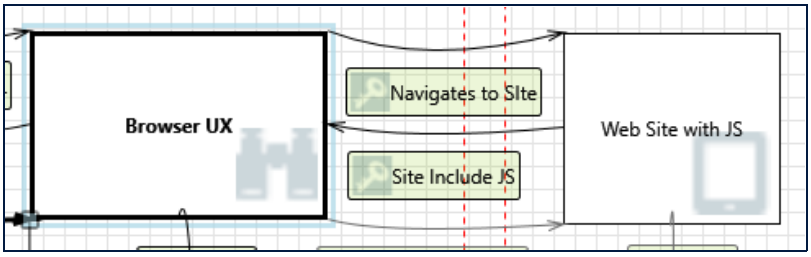
37. Potential Excessive Resource Consumption for Local AI or Session/Site Data [State: Mitigation Implemented] [Priority: High]

Category: Denial Of Service

Description: Does Browser AI or Session/Site Data take explicit steps to control resource consumption? Resource consumption attacks can be hard to deal with, and there are times that it makes sense to let the OS do the job. Be careful that your resource requests don't deadlock, and that they do timeout.

Justification: If the file quota is full the holder may not get the expected functionality. Try to purge old data when session are completed.

Interaction: Site Include JS



38. Data Flow Site Include JS Is Potentially Interrupted [State: Mitigation Implemented] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: Loss of access to the resource by the holder. Normal DOS protections by web site.

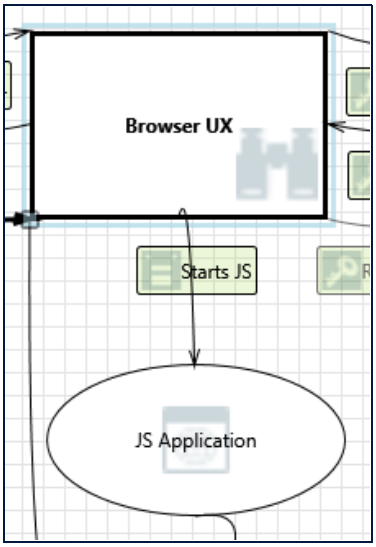
39. External Entity Browser Potentially Denies Receiving Data [State: Mitigation Implemented] [Priority: High]

Category: Repudiation

Description: Browser UX claims that it did not receive data from a process on the other side of the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: If the Browser AI has a cost or limit in functionality, a rogue use can cost the holder money or loss of access.

Interaction: Starts JS



40. Elevation Using Impersonation [State: Mitigation Implemented] [Priority: High]

Category: Elevation Of Privilege

Description: JS Application may be able to impersonate the context of Browser UX in order to gain additional privilege.

Justification: Implicit trust. Access may be limited to the holder if biometric proofing is enabled.

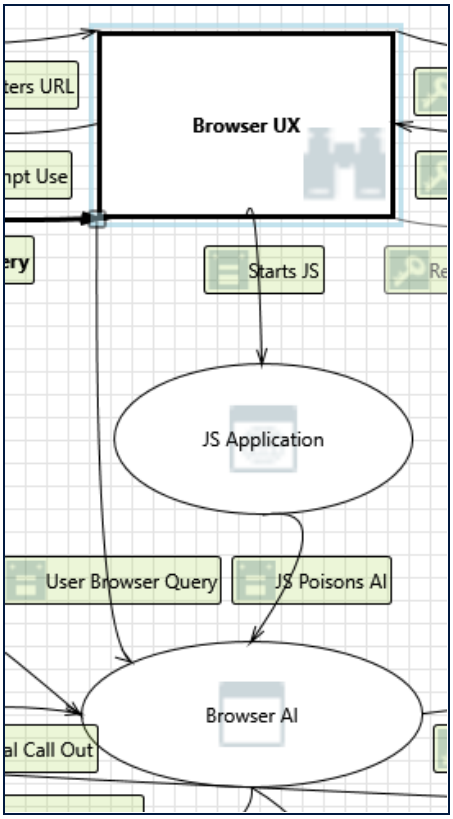
41. Spoofing the Browser External Entity [State: Not Applicable] [Priority: High]

Category: Spoofing

Description: Browser UX may be spoofed by an attacker and this may lead to unauthorized access to JS Application. Consider using a standard authentication mechanism to identify the external entity.

Justification: The user has no control over the JavaScript loaded by the web site.

Interaction: User Browser Query



42. Spoofing the Browser External Entity [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: Browser UX may be spoofed by an attacker and this may lead to unauthorized access to Browser AI. Consider using a standard authentication mechanism to identify the external entity.

Justification: If the broswer does not validate the user, then the holder may lose access if some resourse is depleated.

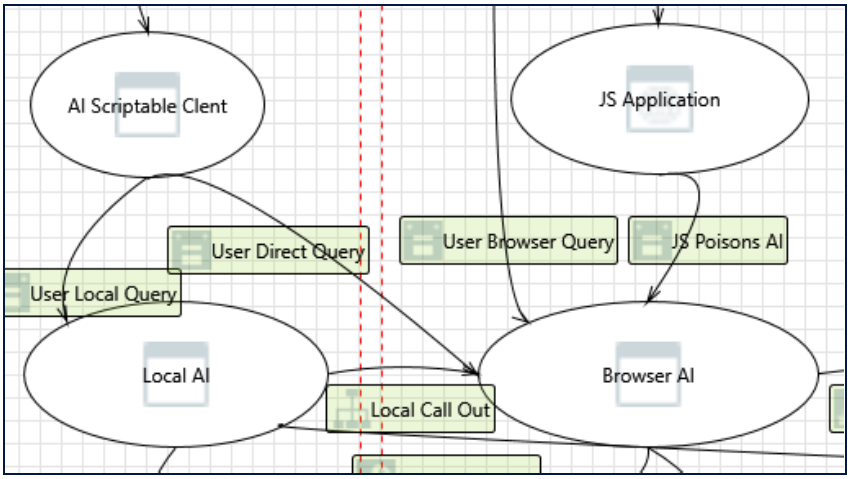
43. Elevation Using Impersonation [State: Mitigation Implemented] [Priority: High]

Category: Elevation Of Privilege

Description: Browser AI may be able to impersonate the context of Browser UX in order to gain additional privilege.

Justification: The assumption is that the user via the browser is trusted to make requests. The brower could validate the user.

Interaction: User Direct Query



44. Spoofing the Browser AI Process [State: Mitigation Implemented] [Priority: High]

Category: Spoofing

Description: Browser AI may be spoofed by an attacker and this may lead to information disclosure by AI Scriptable Clent. Consider using a standard authentication mechanism to identify the destination process.

Justification: It is not likely that the Browser AI is accessible from other apps. The Broswer AI could ask for verification of the AI Scriptable Client code.

45. Spoofing the Native Application Process [State: Needs Investigation] [Priority: High]

Category: Spoofing

Description: AI Scriptable Clent may be spoofed by an attacker and this may lead to unauthorized access to Browser AI. Consider using a standard authentication mechanism to identify the source process.

Justification: It is not likely that the Browser AI is accessible from other apps. The Broswer AI could ask for verification of the AI Scriptable Client code.

46. Potential Lack of Input Validation for Browser AI [State: Mitigation Implemented] [Priority: High]

Category: Tampering

Description: Data flowing across User Direct Query may be tampered with by an attacker. This may lead to a denial of service attack against Browser AI or an elevation of privilege attack against Browser AI or an information disclosure by Browser AI. Failure to verify that input is as expected is a root cause of a very large number of exploitable issues. Consider all paths and the way they handle data. Verify that all input is verified for correctness using an approved list input validation approach.

Justification: It is not likely that the Browser AI is accessible from other apps. The Broswer AI could ask for verification of the AI Scriptable Client code.

47. Native Application Process Memory Tampered [State: Mitigation Implemented] [Priority: High]

Category: Tampering

Description: If AI Scriptable Clent is given access to memory, such as shared memory or pointers, or is given the ability to control what Browser AI executes (for example, passing back a function pointer.), then AI Scriptable Clent can tamper with Browser AI. Consider if the function could work with less access to memory, such as passing data rather than pointers. Copy in data provided, and then validate it.

Justification: It is not likely that the Browser AI is accessible from other apps. The Broswer AI could ask for verification of the AI Scriptable Client code.

48. Potential Data Repudiation by Browser AI [State: Not Applicable] [Priority: High]

Category: Repudiation

Description: Browser AI claims that it did not receive data from a source outside the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: only an issue if some resource is depleted by high use.

49. Data Flow Sniffing [State: Needs Investigation] [Priority: High]

Category: Information Disclosure

Description: Data flowing across User Direct Query may be sniffed by an attacker. Depending on what type of data an attacker can read, it may be used to attack other parts of the system or simply be a disclosure of information leading to compliance violations. Consider encrypting the data flow.

Justification: The Browser AI my implement session to separate usage patterns.

50. Potential Process Crash or Stop for Browser AI [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: Browser AI crashes, halts, stops or runs slowly; in all cases violating an availability metric.

Justification: Resource overloads.

51. Data Flow User Direct Query Is Potentially Interrupted [State: Not Applicable] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: Resource overloads.

52. Elevation Using Impersonation [State: Mitigation Implemented] [Priority: High]

Category: Elevation Of Privilege

Description: Browser AI may be able to impersonate the context of AI Scriptable Clent in order to gain additional privilege.

Justification: typically local applications are trusted by the user. This could be limited to holder if desired.

53. Browser AI May be Subject to Elevation of Privilege Using Remote Code Execution [State: Mitigation Implemented] [Priority: High]

Category: Elevation Of Privilege

Description: AI Scriptable Clent may be able to remotely execute code for Browser AI.

Justification: typically local applications are trusted by the user. This could be .imited to holder if desired.

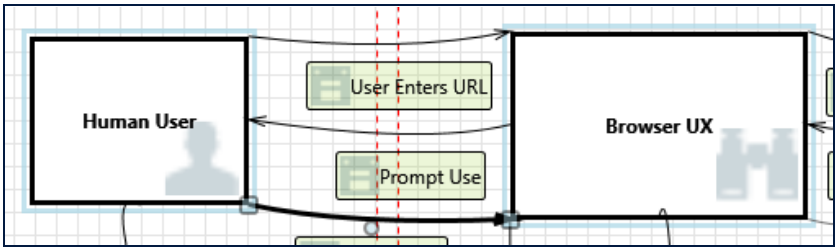
54. Elevation by Changing the Execution Flow in Browser AI [State: Needs Investigation] [Priority: High]

Category: Elevation Of Privilege

Description: An attacker may pass data into Browser AI in order to change the flow of program execution within Browser AI to the attacker's choosing.

Justification: may be limited to holder if desired

Interaction: User Enters URL



55. Authenticated Data Flow Compromised [State: Needs Investigation] [Priority: High]

Category: Tampering

Description: An attacker can read or modify data transmitted over an authenticated dataflow. Like spoofing the user may have enabled accessibility or HID input apps that can sent what appears to be user input.

Justification: The browser may validate the user for normal requests. Browsers already attempt to protect users from hostile sites. Like spoofing the user may have enabled accessibility or HID input apps that can sent what appears to be user input.

56. External Entity Browser Potentially Denies Receiving Data [State: Needs Investigation] [Priority: High]

Category: Repudiation

Description: Browser UX claims that it did not receive data from the user. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: If the device and browser user is not the holder, then there is no proof of the ID of the user.

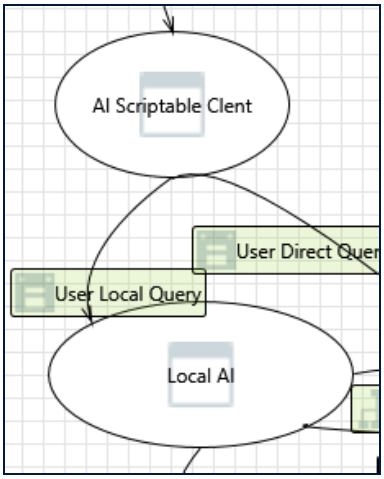
57. Data Flow User Enters URL Is Potentially Interrupted [State: Mitigation Implemented] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: The user may not be able to validate their identity. The browser may block sites that the user would like to access.

Interaction: User Local Query



58. Elevation Using Impersonation [State: Mitigation Implemented] [Priority: High]

Category: Elevation Of Privilege

Description: Local AI may be able to impersonate the context of AI Scriptable Clent in order to gain additional privilege.

Justification: Access if presumably under control of the holder or user. The holder may require biometric proofing if desired.

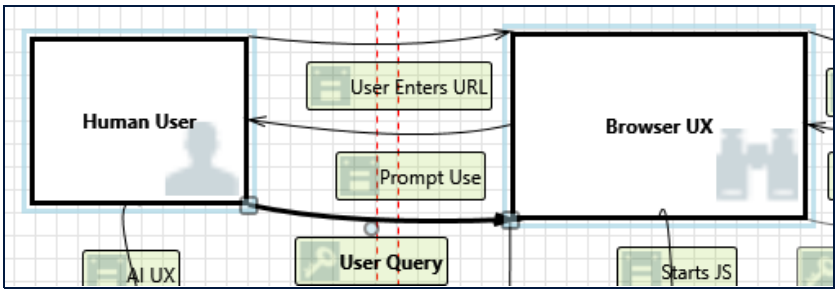
59. Native Application Process Memory Tampered [State: Mitigation Implemented] [Priority: High]

Category: Tampering

Description: If AI Scriptable Clent is given access to memory, such as shared memory or pointers, or is given the ability to control what Local AI executes (for example, passing back a function pointer.), then AI Scriptable Clent can tamper with Local AI. Consider if the function could work with less access to memory, such as passing data rather than pointers. Copy in data provided, and then validate it.

Justification: The Local AI should validate the identity of the AI Scriptable Client to be sure it is trusted by the User.

Interaction: User Query



60. External Entity Browser UX Potentially Denies Receiving Data [State: Mitigation Implemented] [Priority: High]

Category: Repudiation

Description: Browser UX claims that it did not receive data from a process on the other side of the trust boundary. Consider using logging or auditing to record the source, time, and summary of the received data.

Justification: Broser or JS application my ignore user commands.

61. Data Flow Reports back to SIte Is Potentially Interrupted [State: Mitigation Implemented] [Priority: High]

Category: Denial Of Service

Description: An external agent interrupts data flowing across a trust boundary in either direction.

Justification: Broser or JS application my ignore user commands.