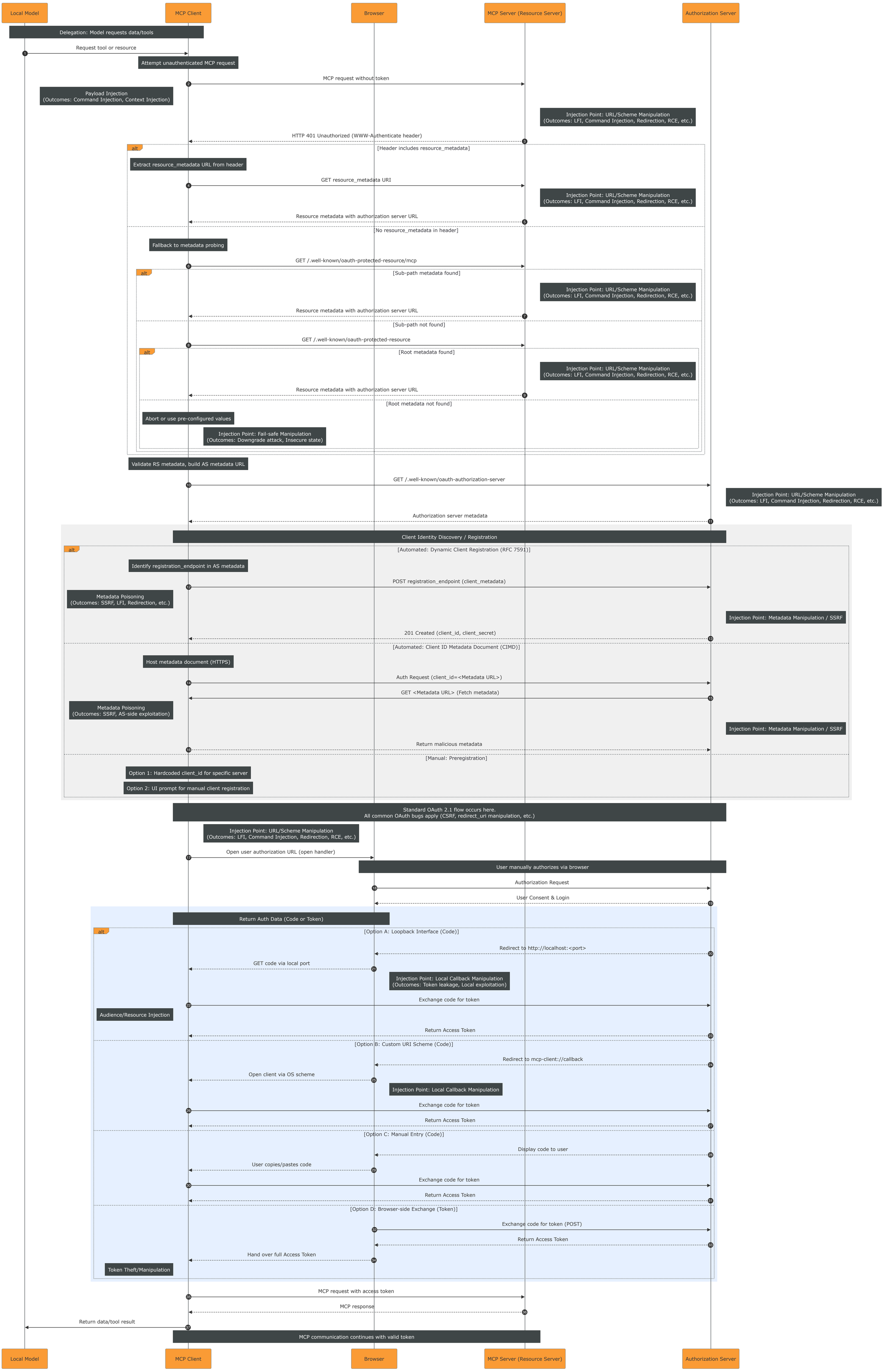
Local Model
MCP Client
Browser
MCP Server (Resource Server)
Authorization Server
Delegation: Model requests data/tools
Request tool or resource
Attempt unauthenticated MCP request
MCP request without token
Payload Injection
(Outcomes: Command Injection, Context Injection)
Injection Point: URL/Scheme Manipulation
(Outcomes: LFI, Command Injection, Redirection, RCE, etc.)
HTTP 401 Unauthorized (WWW-Authenticate header)
alt
[Header includes resource_metadata]
Extract resource_metadata URL from header
GET resource_metadata URI
Injection Point: URL/Scheme Manipulation
(Outcomes: LFI, Command Injection, Redirection, RCE, etc.)
Resource metadata with authorization server URL
[No resource_metadata in header]
Fallback to metadata probing
GET /.well-known/oauth-protected-resource/mcp
alt
[Sub-path metadata found]
Injection Point: URL/Scheme Manipulation
(Outcomes: LFI, Command Injection, Redirection, RCE, etc.)
Resource metadata with authorization server URL
[Sub-path not found]
GET /.well-known/oauth-protected-resource
alt
[Root metadata found]
Injection Point: URL/Scheme Manipulation
(Outcomes: LFI, Command Injection, Redirection, RCE, etc.)
Resource metadata with authorization server URL
[Root metadata not found]
Abort or use pre-configured values
Injection Point: Fail-safe Manipulation
(Outcomes: Downgrade attack, Insecure state)
Validate RS metadata, build AS metadata URL
GET /.well-known/oauth-authorization-server
Injection Point: URL/Scheme Manipulation
(Outcomes: LFI, Command Injection, Redirection, RCE, etc.)
Authorization server metadata
Client Identity Discovery / Registration
alt
[Automated: Dynamic Client Registration (RFC 7591)]
Identify registration_endpoint in AS metadata
POST registration_endpoint (client_metadata)
Metadata Poisoning
(Outcomes: SSRF, LFI, Redirection, etc.)
Injection Point: Metadata Manipulation / SSRF
201 Created (client_id, client_secret)
[Automated: Client ID Metadata Document (CIMD)]
Host metadata document (HTTPS)
Auth Request (client_id=<Metadata URL>)
GET <Metadata URL> (Fetch metadata)
Metadata Poisoning
(Outcomes: SSRF, AS-side exploitation)
Injection Point: Metadata Manipulation / SSRF
Return malicious metadata
[Manual: Preregistration]
Option 1: Hardcoded client_id for specific server
Option 2: UI prompt for manual client registration
Standard OAuth 2.1 flow occurs here.
All common OAuth bugs apply (CSRF, redirect_uri manipulation, etc.)
Injection Point: URL/Scheme Manipulation
(Outcomes: LFI, Command Injection, Redirection, RCE, etc.)
Open user authorization URL (open handler)
User manually authorizes via browser
Authorization Request
User Consent & Login
Return Auth Data (Code or Token)
alt
[Option A: Loopback Interface (Code)]
Redirect to http://localhost:<port>
GET code via local port
Injection Point: Local Callback Manipulation
(Outcomes: Token leakage, Local exploitation)
Exchange code for token
Audience/Resource Injection
Return Access Token
[Option B: Custom URI Scheme (Code)]
Redirect to mcp-client://callback
Open client via OS scheme
Injection Point: Local Callback Manipulation
Exchange code for token
Return Access Token
[Option C: Manual Entry (Code)]
Display code to user
User copies/pastes code
Exchange code for token
Return Access Token
[Option D: Browser-side Exchange (Token)]
Exchange code for token (POST)
Return Access Token
Hand over full Access Token
Token Theft/Manipulation
MCP request with access token
MCP response
Return data/tool result
MCP communication continues with valid token
Local Model
MCP Client
Browser
MCP Server (Resource Server)
Authorization Server