# Was müssen wir wissen?

**Von LLM-Expert\*innen:**

* Welche Einschränkungen und Risiken bestehen bei der Anwendung von LLMs auf semi-strukturierte Diagramme wie DFDs?
* Welche Arten von Prompts oder Eingabeformaten eignen sich besonders gut für strukturierte Analyseaufgaben?
* Wie würden Sie die Entscheidungsfindung eines LLMs einer Person ohne ML-Hintergrund verständlich erklären oder begründen?

**Von Sicherheitsexpert\*innen:**

* Was sind typische Schmerzpunkte beim STRIDE-basierten Threat Modeling unter Verwendung von DFDs?
* Wie wichtig ist Erklärbarkeit bei automatisierten Bedrohungsvorschlägen?
* Welche Mindestanforderungen muss ein System erfüllen, um in diesem Kontext als vertrauenswürdig zu gelten?

**Von Softwareentwickler\*innen:**

* Wie verwenden Sie aktuell DFDs oder Threat Modeling in Ihrem Entwicklungsprozess?
* Würde sich ein solches System in Ihre bestehenden Werkzeuge und Arbeitsabläufe integrieren lassen?
* Welche Arten von Ausgaben (Outputs) würden Sie tatsächlich nutzen – und welche eher ignorieren?

# Das Konzept validieren

1. Ist das Vorgehen sinnvoll? Was fehlt? (F1)
2. Wo sind Herausforderungen aus Sicht von LLM / Security / Software Engineering? (F2)
3. Welche Verbesserungen oder Alternativansätze schlagen Sie vor? (F3)

- 1: Ist das Vorgehen sinnvoll? Was fehlt? - 2: Wo sind Herausforderungen? - 3: Welche Verbesserungen oder Alternativen?

# Identifizieren der Anforderungen (Input, Prompts, and Automation)

1. Wann ist menschlicher Input notwendig? Wo explizit? (F5)
2. Welche Prompt-Typen führen zu brauchbarem Output? (F6)
3. Unterschiede zwischen textueller Prompt-Eingabe vs. grafischer (DFD als Bild)? (F7)
4. Chain-of-Thought nutzbar? Wie könnte das aussehen? (F8)

- 5: Wann ist menschlicher Input notwendig? Wo? - 6: Welche Prompt-Typen funktionieren? - 7: Textuelle vs. grafische Eingabe? - 8: Chain-of-Thought nutzbar?

# Verstehen von Vertrauen, Aufsicht und Grenzen

1. Welche Tools sind geeignet – warum? (F4)
2. Wie kann man Vertrauen in ein KI-basiertes System schaffen? (F9)
3. Welche Visualisierungen oder Formate wären hilfreich für die Kommunikation mit dem Tool (z. B. Text, DFD, Flowcharts, Tabellen)? (F10)

- 4: Welche Tools sind geeignet – warum? - 9: Wie schafft man Vertrauen? - 10: Welche Formate sind hilfreich für die Kommunikation?

# Domänenübergreifende Perspektiven abbilden

1. Welche Informationen fehlen Ihnen aktuell noch, um sich ein vollständiges Bild vom geplanten Prozess zu machen? (F11)

- 11 Gibt es etwas, das wir noch nicht angesprochen haben?

# Empfohlene Interface-Sektionen für einen ersten Prototyp

# Eingabebereich

* 1. Hochladen oder Einfügen des DFD (als Bild oder Text)
  2. Optional: DFD-Editor
  3. Auswahl: z. B. STRIDE-Analyse starten

# Threat-Vorschläge

1. Automatisch generierte Liste
2. STRIDE-Kategorie pro Threat
3. Kurzbegründung
4. Vertrauensbewertung
5. Aktionen: akzeptieren, bearbeiten, ablehnen, zur Prüfung markieren

# Erklärungsbereich

Beim Klick auf einen Threat:

→ “Warum wurde dieser vorgeschlagen?”

→ Schrittweise Kausalkette (Chain-of-Thought), wenn verfügbar

# Ausgabe & Export

1. Tabs zur Anzeige:
   * Tabelle der Bedrohungen
   * Annotiertes DFD
   * Zusammenfassung als Text
2. Exportoptionen: PDF, JSON, Markdown, etc.

# Einstellungen / Kontrolle

1. Menschliche Validierung aktivieren/deaktivieren
2. Prompt-Stil auswählen: strukturiert vs. frei
3. Vertrauensintervall oder Confidence-Score anzeigen