## Arquitetura de Integração: OSINT e Investigação

### 1. Visão Geral

O objetivo é integrar as funcionalidades de OSINT e investigação de forma modular e escalável ao sistema existente. A arquitetura proposta visa:

* **Coleta de Dados:** Utilizar as ferramentas selecionadas (theHarvester, SpiderFoot, Shodan) para coletar informações de fontes abertas.
* **Processamento e Análise:** Centralizar os dados coletados e aplicar técnicas de análise para identificar conexões e padrões relevantes.
* **Visualização:** Apresentar os resultados de forma clara e intuitiva na interface gráfica, utilizando a biblioteca de visualização escolhida (NetworkX com Matplotlib/Graphviz).
* **Interação com o Usuário:** Permitir que o usuário guie o processo de investigação, fornecendo feedback e direcionando as próximas etapas.

### 2. Componentes da Arquitetura

1. **Módulo de Coleta de Dados:**
   * **Submódulo de Ferramentas Externas:** Integração com as APIs das ferramentas selecionadas (theHarvester, Shodan) para automatizar a coleta de dados.
   * **Submódulo de Web Scraping:** Desenvolvimento de scripts para extrair informações de fontes web que não possuem APIs (ex: redes sociais, fóruns, blogs).
   * **Submódulo de Importação de Dados:** Funcionalidade para importar dados de arquivos locais (ex: CSV, TXT) ou de outras fontes.
2. **Módulo de Processamento e Análise de Dados:**
   * **Normalização e Padronização:** Tratamento dos dados coletados para um formato consistente.
   * **Extração de Entidades:** Identificação de entidades relevantes (pessoas, organizações, locais, etc.) e seus relacionamentos.
   * **Análise de Sentimento (Opcional):** Avaliação do tom e sentimento em textos coletados para identificar potenciais ameaças ou opiniões relevantes.
   * **Detecção de Padrões:** Utilização de algoritmos para identificar padrões e anomalias nos dados.
3. **Módulo de Visualização e Interação:**
   * **Interface Gráfica (GUI):** Exibição dos resultados da análise de forma clara e interativa.
   * **Visualização de Grafos:** Utilização da biblioteca escolhida (NetworkX com Matplotlib/Graphviz) para representar as conexões entre entidades.
   * **Interação do Usuário:** Permitir que o usuário explore os dados, filtre informações e forneça feedback para refinar as análises.
4. **Módulo de Gerenciamento de Casos:**
   * **Criação e Gerenciamento de Casos:** Permitir que o usuário crie e gerencie casos de investigação, agrupando informações e descobertas.
   * **Geração de Relatórios:** Geração de relatórios customizáveis com os resultados da investigação.

### 3. Fluxo de Dados e Interação

1. O usuário inicia uma nova investigação ou seleciona uma existente.
2. O usuário define os alvos e os tipos de informações a serem coletadas.
3. O sistema utiliza os módulos de coleta de dados para buscar informações em fontes abertas e na web.
4. Os dados coletados são processados e analisados pelo módulo de processamento e análise.
5. Os resultados são apresentados na interface gráfica, com destaque para as informações mais relevantes e as conexões identificadas.
6. O usuário interage com os dados, explorando os grafos, filtrando informações e adicionando anotações.
7. O sistema utiliza o feedback do usuário para refinar as análises e buscar novas informações.
8. O usuário pode gerar relatórios com as descobertas e conclusões da investigação.

### 4. Considerações Adicionais

* **Escalabilidade:** A arquitetura deve ser projetada para lidar com grandes volumes de dados e permitir a adição de novas ferramentas e funcionalidades no futuro.
* **Segurança:** É fundamental garantir a segurança dos dados coletados e armazenados, bem como a privacidade dos indivíduos envolvidos.
* **Usabilidade:** A interface do usuário deve ser intuitiva e fácil de usar, mesmo para usuários com pouca experiência em análise de dados.

Este plano de arquitetura servirá como base para o desenvolvimento da próxima fase do projeto.