

The background of the slide features a complex network graph. It consists of numerous nodes, represented by small circles in various colors including blue, green, orange, purple, and teal. These nodes are interconnected by a dense web of thin, dark grey lines, representing edges. The overall layout of the nodes and edges is somewhat circular and dense, filling the upper and middle portions of the slide. The text is overlaid on this background.

# Sistema de Recomendação de Músicas

---

Desenvolvimento e Funcionamento

# INTRODUÇÃO

---

- Este projeto desenvolve um sistema de recomendação de músicas baseado em atributos textuais e numéricos.
- O modelo sugere músicas relevantes utilizando machine learning e análise de dados.

# BIBLIOTECAS

---

- Pandas: Manipulação de grandes volumes de dados, essencial para organizar e processar informações musicais.
- Scikit-learn: Modelagem de dados, vetorização de texto e normalização numérica para análise eficiente.
- Matplotlib e Numpy: Visualização gráfica e operações matemáticas avançadas para representar tendências musicais.

# PROCESSO DE LIMPEZA

---

- Tratamento de Valores Ausentes: Substituição de valores numéricos por medianas
- Padronização de Categorias: Uniformização de gêneros musicais e artistas para garantir consistência.
- Normalização: Escalonamento de dados numéricos para o intervalo  $[0, 1]$

# FUNCIONAMENTO DO CÓDIGO

---

- Combinação de Dados: Transformação de atributos textuais com TF-IDF e normalização de atributos numéricos como energia e popularidade.
- Redução Dimensional: Utilização de PCA para reduzir a complexidade dos dados combinados.
- Cálculo de Similaridade: Uso de similaridade de cosseno para identificar músicas relacionadas.