Big Data Obtenção de Dados

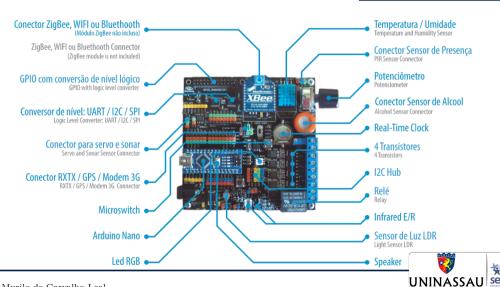


Prof. Flávio Murilo de Carvalho Leal Centro Universitário de Juazeiro do Norte Universasau



- ▶ Bancos de Dados Relacionais e Não-Relacionais
- ► Sensores IoT (Internet of Things)
- ► Redes Sociais
- ► Logs de Servidores Web
- ▶ Dados Públicos e Governamentais









CINCO MOTIVOS PARA A ABERTURA DOS DADOS

- Contribuição da sociedade com serviços inovadores ao cidadão
- Aprimoramento na qualidade dos dados governamentais
- Viabilização de novos negócios

PARA QUE

Podem ser usados pelos cidadãos e pelas organizações da sociedade para verificar, esclarecer, fiscalizar e acompanhar as informações produzidas pelo governo

Além de promover transparência, a divulgação dos dados permite fiscalizar os órgãos do poder público e apoiar a formulação de projetos privados e também de novas políticas públicas



QUAIS DADOS NÃO PODEM SER ABERTOS?

Não podem ser abertos os dados que estiverem sujeitos a restrições de sigilo ou de privacidade, conforme definidos na Lei de Acesso à Informação. Nestes casos, os dados não são catalogados no Portal Brasileiro de Dados Abertos



Lei de Acesso à Informação

Fonte: Portal de Dados Abertos



- ▶ Qualidade dos Dados: Ruído, dados faltantes, duplicados
- ► Privacidade e Segurança dos Dados
- ► Escalabilidade das Soluções de Coleta
- ► Integração de Fontes Diversas



- ► Hadoop: Framework para processamento distribuído de grandes volumes de dados
- ▶ Spark: Motor de processamento rápido para Big Data
- ▶ Kafka: Plataforma de streaming de dados
- ▶ Flume: Serviço para coleta de logs

