

# Open Data and AI

**Nikesh Balami**

CEO - Open Knowledge Nepal

[www.neekes.com.np](http://www.neekes.com.np)



AI

- sounds sexy
- gets us money from VCs
- what we all hope is the future

Machine Learning

- the only real “AI”
- traditionally an academic discipline
- not concerned with real-world software

Data Science

- applies machine learning to create actual products
- deals with real-world complexity

# Data Science - The Connection

**Data Science** helps AIs figure out solutions to problems by linking similar data for future use and allows to find appropriate and meaningful information from those huge pools faster and more efficiently.


**An example:** Facebook's facial recognition system



~~Data Science~~  
**Open Data**

# Open Data

Data can be defined as information in its raw, pre-analyzed form, such as numbers, words, or pictures. Open data, on the other hand, ***“refers to data that is made available in a machine readable format and shared publicly so that it is free to access, use and reuse for any purpose”***. To explain it simply, with open data:

- The data itself should be accessible freely through the medium of internet such as through websites, data portals, etc.
  - The available data should be usable and reusable without any legal restriction.
  - Using and performing operations on the data to add value through analysis, visualization, developing applications and more should be free.
  - Information being made available in a machine-readable format to enable computer-based reuse.
- 

# Why Open Data?

- Innovation, Efficiency and Transparency
- The Many Mind Principle - The best thing to do with your data will be thought of by someone else.
- Fixing is Faster with Open Data - To many eyes all bugs are shallow.



**Open Data should become the new open source;**

Shared, Clean, Enriched Data is one of the key ingredients to real innovation

# Technology to work we need a lot of data

We all will benefit from opening more data to the world so multiple teams and startups that might be working on new ideas may use it to make a smarter AI and find the answers much needed by our society.





# International Scenario - Open Gov Data fit for Machine Learning

## Local Government

Inicio Misiones

Open Government Data Wien (OGD)

Open data - City of Brussels

Open Data - Brisbane City Council

Open data - Salford City Council

Sunderland City Council : Local Public Data

Welcome to the London Datastore | London DataStore

Leeds City Council - Open Data

Home - DataGM - Data Greater Manchester

Open Data | Derby City Council

Council data - Brighton & Hove City Council

Open Data - Birmingham City Council

Aberdeen City Council Open Data

Open Data - City of Waterloo

Open Data catalogue | City of Vancouver

Open Data Home - Open Data - Home | City of Toronto

City of Prince George - Open Data Catalogue

Open Data Ottawa | City of Ottawa

Open Data Catalogue - City of Red Deer

Open Data | City of Niagara Falls, Canada

Open Data Catalogue | City of Nanaimo

Mississauga.ca - Residents - Publications and Open Data Catalogue

City of Medicine Hat Open Data Catalogue

Kamloops open data

Open Data Catalogue Kelowna

City of Hamilton - Open Data

City of Fredericton - Open Data Home

City of Edmonton Open Data Catalogue

City of Somerville, MA

Data.Seattle.Gov | Seattle's Data Site

City of Scottsdale

Welcome - Santa Cruz Open Data

Data | San Francisco

Open Raleigh - The Official City of Raleigh Portal

## National Governments and States

Portal de Obligaciones de Transparencia

Junta de Andalucía - Datos abiertos

Reutilización de la Información del Sector Público | Reutilización de la Información de los Servicios Públicos

Portal de Datos Abiertos de JCCM

Ayuntamiento de Zaragoza. Datos de Zaragoza Reutilización

Dades obertes Lleida - Ajuntament de Lleida

ISTAC | El ISTAC

Dades Obertes. Generalitat de Catalunya

Dades Obertes CAIB

Reutilización de la Información del Sector Público en Gijón

Open Data Euskadi ataria, Eusko Jaurlaritzaren datu publikoen irekitzea

Data for Hawaii | data.hawaii.gov

Florida Has A Right To Know

Open.Georgia.gov

Commonwealth Data Point

Open Data | data.maryland.gov

Connecticut Transparency Website

RI.gov: Open Data

NYS Data Center

Maine.gov DataShare

State of Alabama - Open.alabama.gov

Open Government for the State of Tennessee

Ohio.gov | Government | State Facts and History

OpenDoor - Kentucky

Data.Illinois.gov | Open Illinois

SOM - Michigan Data Store

Louisiana Transparency and Accountability Portal

data.mo.gov | State of Missouri Data Portal

DATAShare | data.iowa.gov

Minnesota open data // your portal for Minnesota data transparency

Open Data Texas

Welcome to Oklahoma's Official Web Site

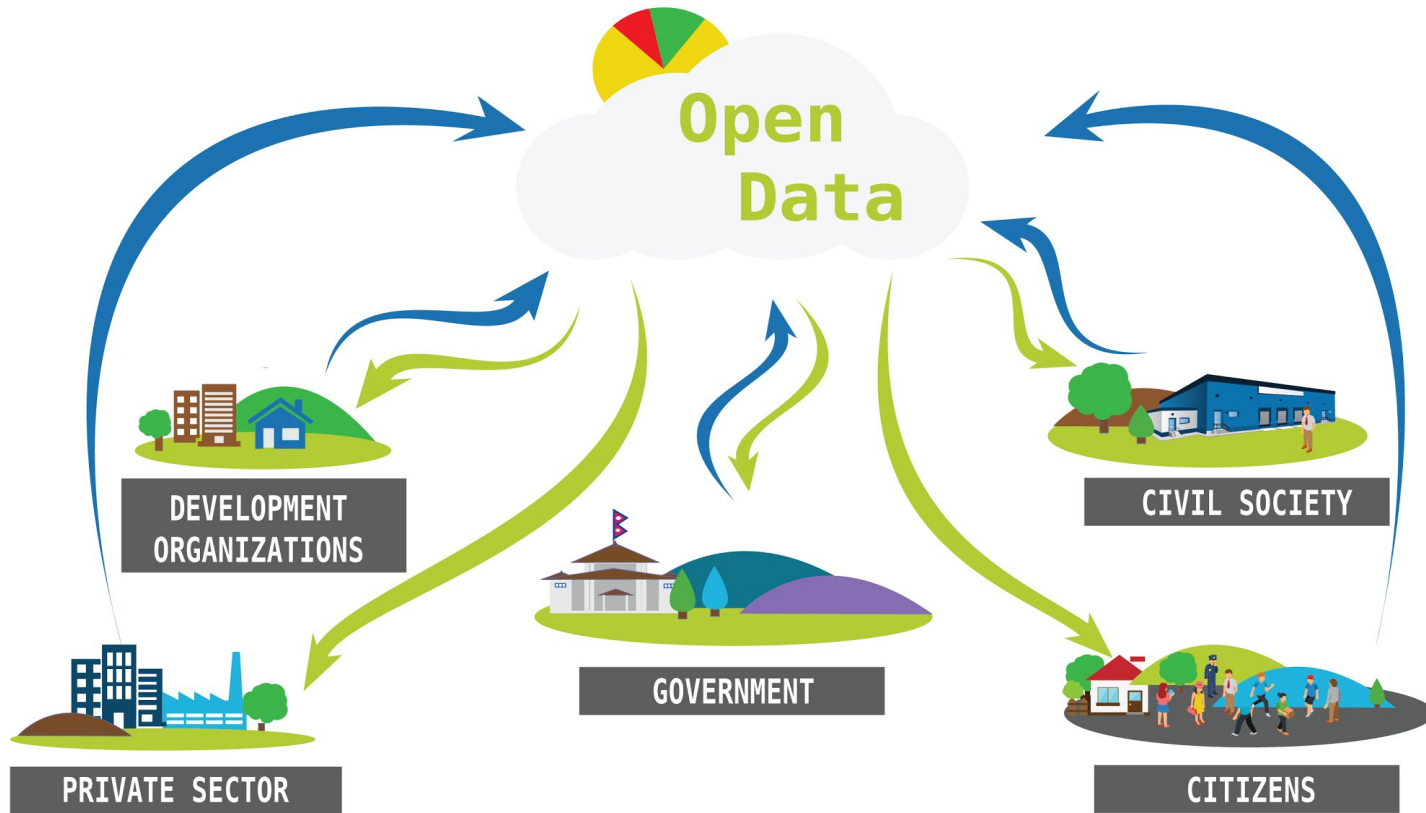
**Open Data creates power for the many not the few**



**But in Nepal**  
It's a children's game



# Nepal's Open Data Ecosystem



# Problem

**First:** Very less data production and unstructured sharing

**Second:** ~~Few~~ machine readable format and Good Data





# FRICTIONLESS DATA

STANDARDS AND TOOLING

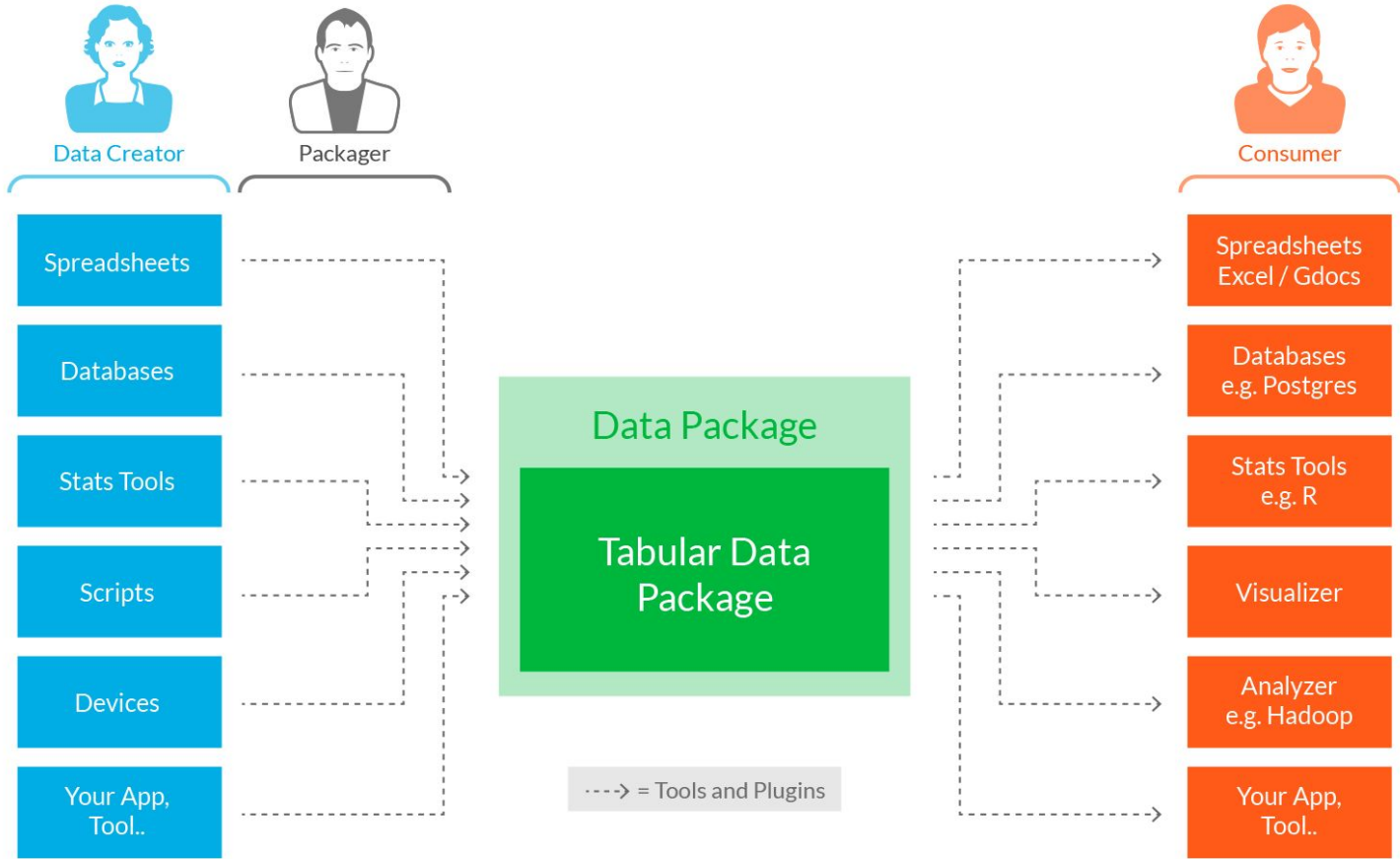
# What

Frictionless Data is about removing the friction in working with data

By developing a set of [tools](#), [standards](#), and [best practices](#) for publishing data. The heart of Frictionless Data is the [Data Package standard](#), a containerization format for any kind of data based on existing practices for publishing open-source software.

**URL:** <http://frictionlessdata.io>







**Go Explore:** <http://frictionlessdata.io/tools/>

TOOLS

---

# Tools, Integrations, Libraries, and Platforms

- Tools and Integrations
- Libraries
- Platforms
- Other
- Representations
- Directory



# Think Again

1. Focused
2. Web-oriented
3. Distributed
4. Open
5. Built Around Existing Tooling
6. Simple



Explore Nepal

**Solve the Problem**

Don't let others face the same struggle which you are facing





**Thank You!!**