StatsBot Presentation
Table of Content

1. StatsBot Demo

1. Next Steps: Statsbot Scalability

1. Q&A
Discover the StatsBot

Golem.ai
OECD
Statistics Canada
Statistique Canada
Key Figures

3 STATS BOTS

1 datagroup | 3 datasets | 3 languages

82% correct answers

final test of 1,515 questions
StatCan StatsBot
Next Steps and Scalability
Golem.ai’s AI is an Award Winning AI
One and Only Symbolic AI for Language Understanding

**Scalability**
Structure is configured once for a language, then re-used for every other language.

**Multilingual**
24 read-to-use languages in Golem.ai. 10 days to add a new language.

**Explainable**
Easily demonstrate any decision taken by the AI.

**Sustainable**
No need for data to configure the product. Very lightweight technology.
Scalability Model: How Datasets and Languages Scale

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset within same data</td>
<td>With each extra dataset the probability of statistics and filters being</td>
</tr>
<tr>
<td>“group”</td>
<td>pre existent in the StatsBot increases, up to 100% can be pre existent</td>
</tr>
<tr>
<td>Dataset in new data</td>
<td>With each extra datagroup the probability of filters being pre existent</td>
</tr>
<tr>
<td>“group”</td>
<td>in the StatsBot increases, up to 100% can be pre existent</td>
</tr>
<tr>
<td>New language</td>
<td>With each extra organisation the probability of statistics and filters</td>
</tr>
<tr>
<td></td>
<td>being pre existent in that language in the StatsBot increases, up to 100%</td>
</tr>
<tr>
<td></td>
<td>can be pre existent</td>
</tr>
</tbody>
</table>
Adding a Language

Adding a language does not require any modification of the model

STEP 1: Add a language in the studio (24 ready-to-use languages in Golem.ai)

STEP 2: Translate the AI configuration into the new language
Adding a Dataset

SDMX DB

ID + keywords / synonyms / languages

matching to existing statistics/filters

AI

integrate validated statistic/filter

matched

propose matching elements for validation
propose missing keywords for validation

unmatched

prefill a new stat/filter... for edition and validation

AI

Studio interface
StatsBot Studio: Your Tool to Manage the StatsBot

Features:

- Add a new dataset
- Add a new filter
- Add a new statistic
- Add a new language
- Maintain/improve existing StatsBots
- KPI dashboard
- Assisted continuous improvement

no code platform to create, maintain and improve StatsBots
StatsBot Target Interface & UX

Population in France

The population in France is 66.99 millions

Query History

Population in France
Scalability Roadmap

**POC**
- R&D validating feasibility

**MVP**
- Implementation proving feasibility & value
  - 5 sprints

**PRODUCTION**
- Implementation providing value
  - 3 sprints

**CONTINUOUS IMPROVEMENT**
- Implementation reaching full potential

- Scalable framework
- Basic UX
- Language handling
- Conversational memory
- Data exploration
- Basis of final UX
- Major data formats
- StatsBot over multiple datasets
- Most relevant features
- Feedback mechanism
- StatsBot studio
- Technical scaling
- New features
- New data formats
- Improvements based on user feedback

sprints are 2 weeks long
Next Steps Depend on You

- If you are interested just contact us: **celia@golem.ai**
- Or Eric Anvar @ oecd

- We shall discuss further:
  - go deeper into tech details
  - full demo
  - ...

- Or work together to deploy your own StatsBot
Follow Us On Our Social Networks
ANNEXES
Workflow overview

AI
Transforms user input into a query by extracting user intention and matching it to the SDMX

AI OUTPUT
Identified components of the user request
- statistics
  - unemployment
  - employment
- filters
  - ref area
  - age group
- query type
  - data query
  - data definition
- query modifiers
  - maximum
  - comparison

user interface

PROCESSOR
Processes queried data according to AI output to deliver a formatted answer and drill-down options

QUERY BUILDER
Validates the user request and queries the corresponding data

DB

.DStat API

DSD

SDMX

1. AI
2. AI OUTPUT
3. user input
4. user interface
5. PROCESSOR
6. QUERY BUILDER
7. DB
Draft Business Model*

Consortium

One-time
StatsBot build

Recurring
StatsBot run

Individual organisations

One-time

Organisation 1
StatsBot instanciation
(N datasets and N languages)

Organisation N
StatsBot instanciation
(N datasets and N languages)

Recurring

Run license proportional to datasets

Scope extension fee

...