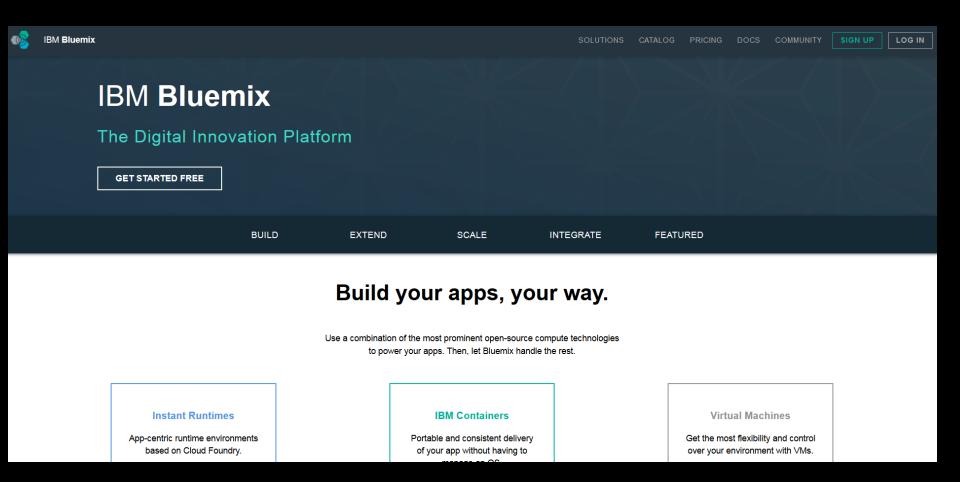
IBM Bluemix Watson Dialog Services

Tomáš Macek, IBM Prague R&D lab

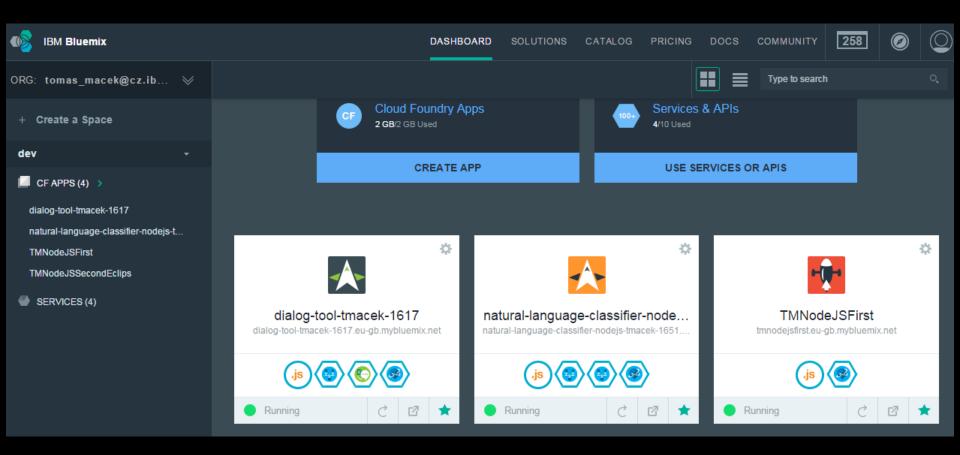
for students CVUT FEL, April 27 2016

Overview

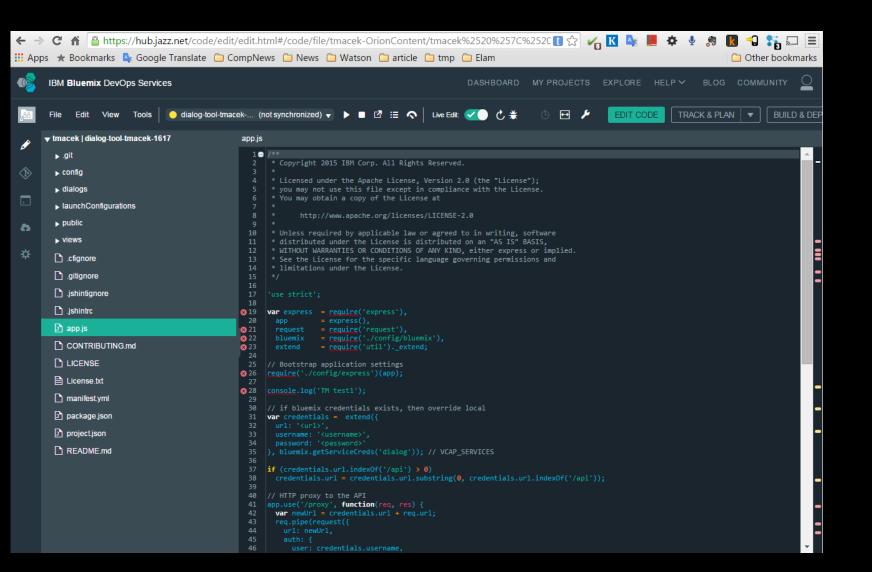
- Bluemix PaaS introduction to the platform
- Prague R&D lab
- Who is Watson?
- NL related services
 - Dialog,
 - -ASR,
 - TTS
- UI specifics for NLP in cloud

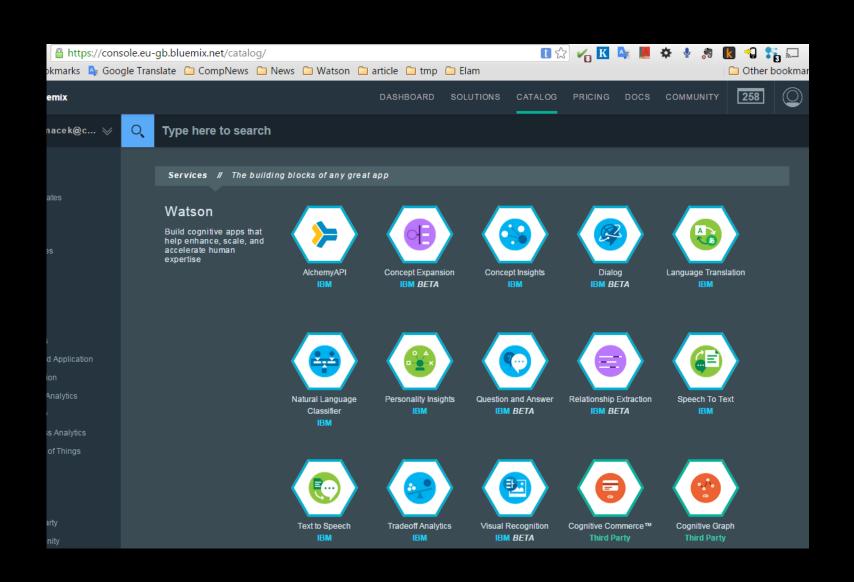


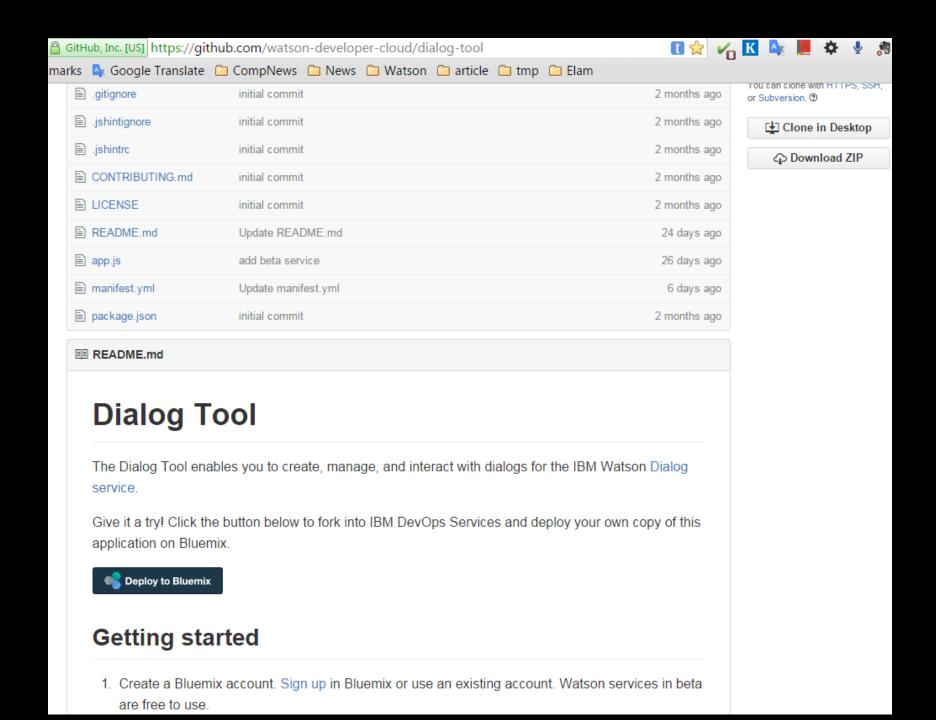
Bluemix



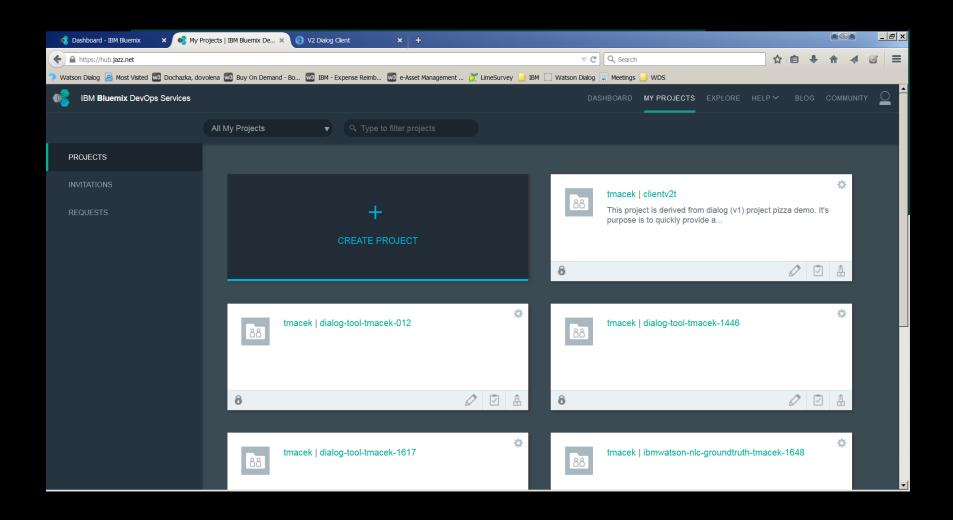
- Service vs application,
- credentials



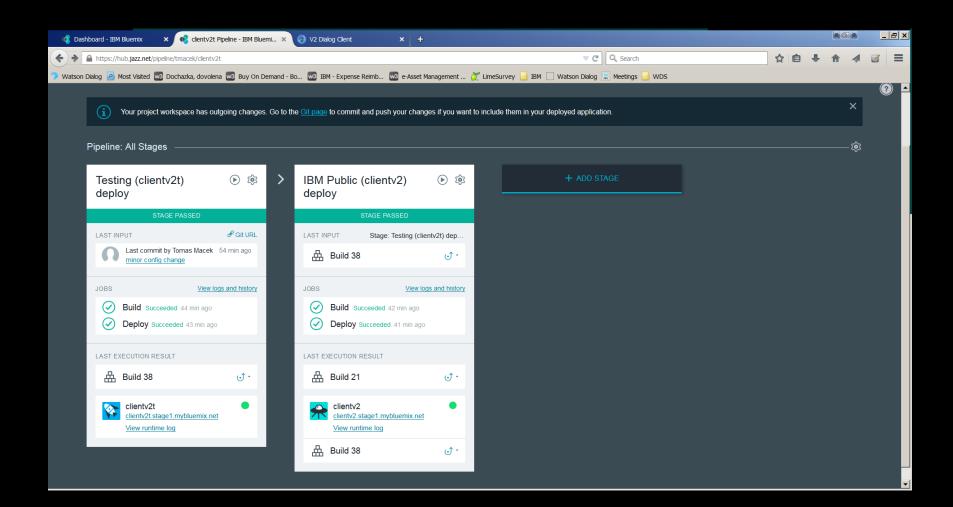




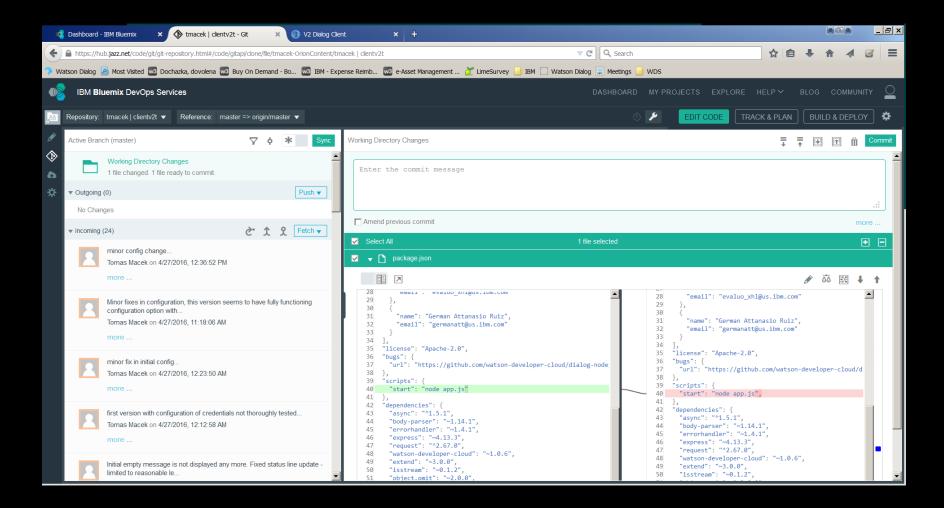
PROJECT MANAGEMENT

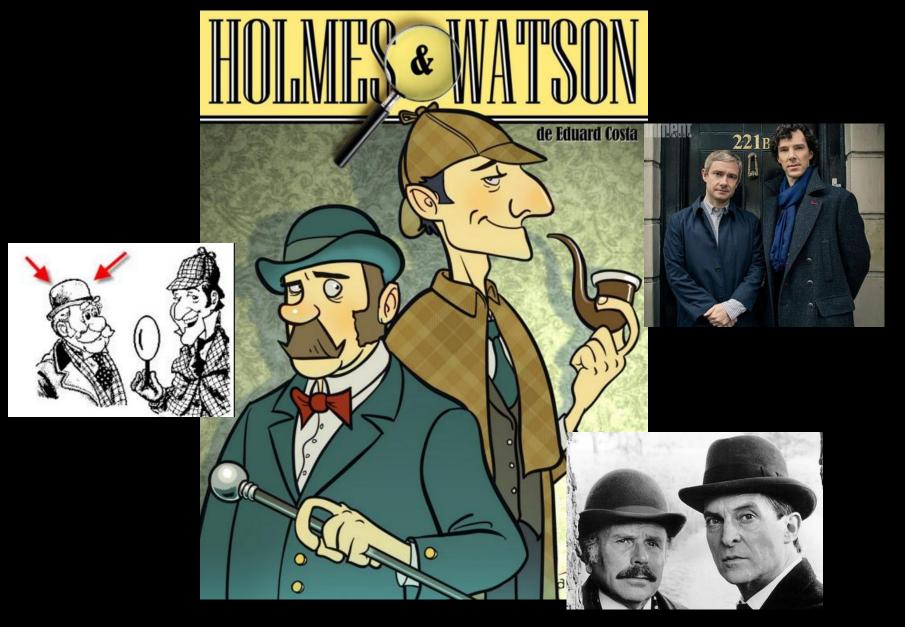


STAGE



GIT INTEGRATION





Who is Watson: Dr. Watson is the friend and assistant of the greatest detective of all ages



Born Thomas John Watson

February 17, 1874

Campbell, New York, U.S.

Died June 19, 1956 (aged 82)

New York City, New York, U.S.

Occupation Chairman and CEO of IBM 1914-

1956, business

THINK". Watson began using "THINK" to motivate, or inspire,

Chairman and CEO of IBM



T. J. Watson research Center

On February 14, 2011, IBM Watson made history



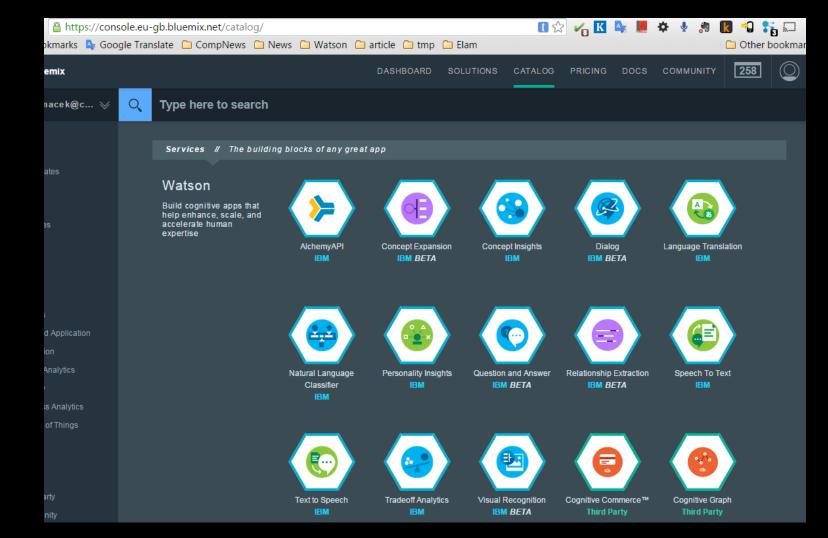
IBM computer which made the history

IBM Prague R&D Lab Watson Dialog Services





Selected Watson services











Natural Language Classifier

The IBM Watson Natural Language Classifier service applies deep learning techniques to make predictions about the best predefined classes for short sentences or phrases. The classes can trigger a corresponding action in an application, such as directing a request to a location or person, or answering a question. After training, the service returns information for texts that it hasn't seen before. The response includes the name of the top classes and confidence values.

API Overview

<u>Documentation</u>

Fork on Github

Try the service

Ask a question about the weather

Watch the Natural Language Classifier categorize your weather-related question. In this demo, the classifier is trained to determine whether the question is related to temperature or weather conditions. The output includes the top classification and a confidence score.

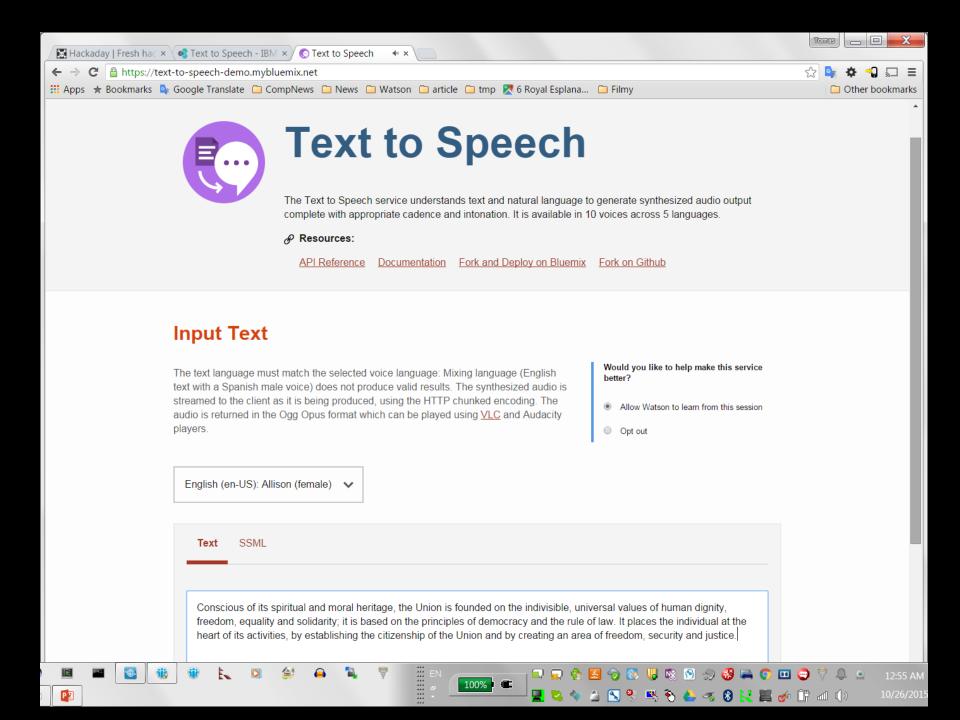
Enter a weather question or Try a sample question below

Ask

Sample questions

Is it hot outside? What is the expected high for today? Will it be foggy tomorrow morning? Should I prepare for sleet? Will there be a storm today?

The classifier often scores well with terms that it hasn't been trained on. In the sample questions, the words "sleet," "foggy," and "typhoon" are not part of the training data, yet the classifier correctly handles questions about them.





Speech to Text

The IBM Watson Speech to Text service uses speech recognition capabilities to convert English, Spanish, Brazilian Portuguese and Mandarin speech into text.



API Reference Documentation Fork on Github Fork and Deploy on Github

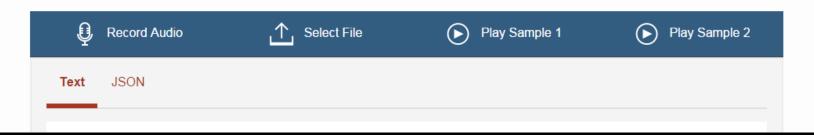
Transcribe Audio

Use your microphone (compatible only with <u>Google Chrome</u> and <u>Mozilla Firefox</u>), upload pre-recorded audio (WAV containing uncompressed audio or FLAC file formats), drag and drop recorded audio onto the page, or use the audio samples provided. The returned text includes metadata that provides the timestamps for start and end times of individual words, confidence scores behind those words, and alternative hypotheses for phrases.

Would you like to help make this service better?

- Allow Watson to learn from this session
- Opt out

US English broadband model (16KHz)



CONVERSATIONAL APP (IN THE WATSON DIALOG SERVICE)

• What it is

- Artificial agent that takes dialog turns with the user
- Runtime Engine + Development
 Environment
- Text in, text out
- Can be used in conjunction with
 Speech-To-Text and Text-To-Speech

What it is not

- 100% accurate
- Full replacement for human assistants
- IDE for authoring full UX
- Stateless FAQ or question answering system

S: Welcome to *Utility 1*. What can I help you with today?

U: Hello

S: Hi there!

U: What are my payment options?

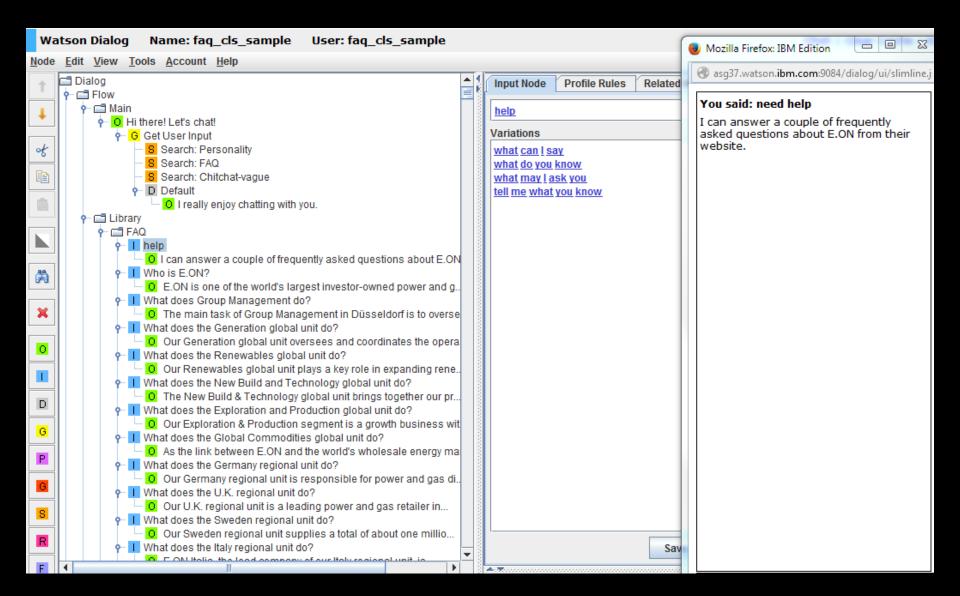
S: There are many ways to pay your bill;

- Credit Card
- Cash
- Fixed or Direct Debit

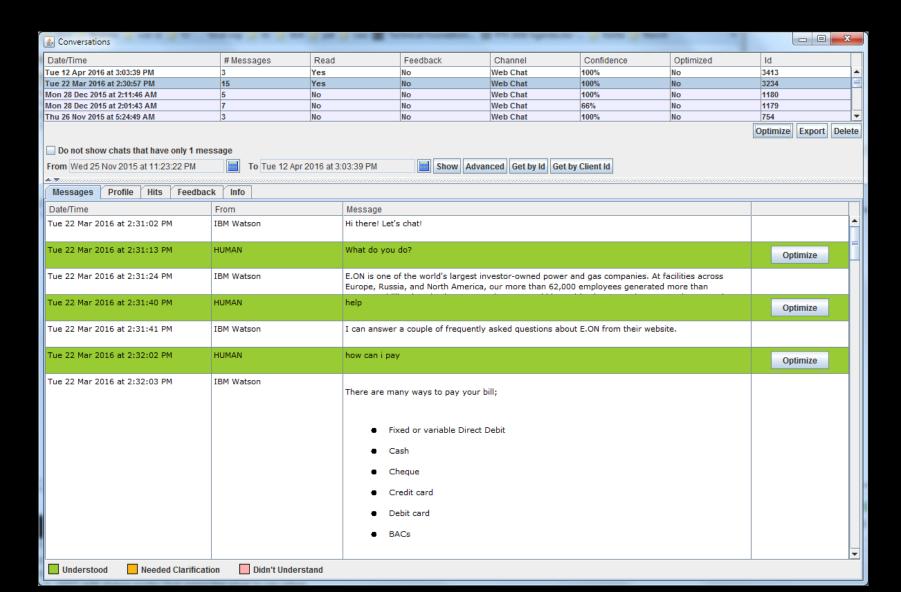
U: Pay by card

S: You can pay by credit/debit card using your online account, or without it on our website.

DIALOG DEVELOPMENT ENVIRONMENT – FLOW EDITOR



DIALOG IDE – VIEW CONVERSATIONS, OPTIMIZE



WHY CREATE CONVERSATIONAL APPS?

- Enable customer / employee self-service through multiple channels
- Ready 24x7, no queues for human operators
- Even several % of traffic handled by WDS save human labor from routine tasks suitable for automation to complex tasks that really require human in the loop
- Consistent answers, can be "retrained" instantly, needs little tutoring
- Typical Tasks for Dialog
 - FAQ with disambiguation
 - Chit-Chat
 - Transactional and Form-filling tasks
 - Problem-solving
 - Mixture of the above
- Delivery channels
 - Mobile or desktop, typing and/or clicking
 - Speech, GUI, Robot (multi-modal)
 - Phone

HOW DOES IT WORK 1/2

- Combination of scripted and statistical techniques
- Flow and
 understanding done
 using a "visual
 programming
 language" of dialog
 - input nodes with matching rules
 - search nodes that group input nodes
 - output nodes
 - profile checks (conditions)
 - profile rules (variable sets)
 - get user input (explicit)
 - goto's

nodes



Rule-based input understanding

- input nodes contain simple matching patterns like
 - "\$ car parking" matches any sentence that contains "car parking"
 - "park * car" matches "park my car", "park my friends car" etc.
- predictable, but large patterns sets are hard to manage

HOW DOES IT WORK 2/2

- Combination of scripted and statistical techniques
- Statistical input understanding
 - instead of patterns, input nodes contain training samples
 - trained model classifies user input to the best matching input node (class)
 - often outperforms rules from moderate samples set sizes
 - generalizes better, easier to author
 - behind the scenes, Dialog integrates with the NL Classifier Service (dark feature)

```
How do I find car parking?
  Campus! -> How do I find car parking?

    $ (CAMPUS ENT)={Campus}
                                              U: "Parking at Burwood"
        P Profile Check: Campus = Burwood

◆ I $ short term

◆ I $ long term

             O Do you want short-term or long-term parking at {Campus}?<...</p>
            Profile Check: Campus = Warrnambool
                O <span class="hidden">Your question was about parking at {Camp...
           P Profile Check: Campus = Waterfront
                O <span class="hidden">Your question was about parking at {Camp...
           P Profile Check: Campus = Waurn Ponds
                O <span class="hidden">Your question was about parking at {Camp...

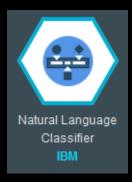
    Geelong

    Deakin has two campuses at Geelong, which one do you want to p...

        G Get User Input
             G GoTo: Campus! -> How do I find car parking?
  Off-campus

    O There are many parking spots and garages around Melbourne, but...

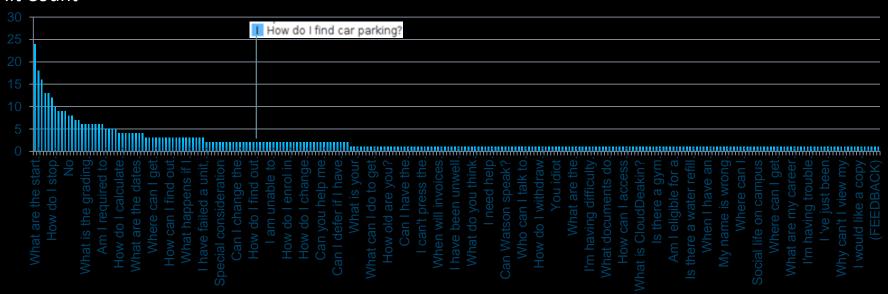
          G GoTo: Which campus do you want to park at?<mct:input>Burwoo...
  P Profile Check: Campus Is Blank
     👇 <mark>0</mark> Which campus do you want to park at?<mct:input>Burwoo...
        G Get User Input
             G GoTo: Campus! -> How do I find car parking?
  P Profile Check: Campus Has Value
     O <span class="hidden">The campus {Campus} has been detected pr...
          G GoTo: Campus! -> How do I find car parking?
```



STATISTICAL VS. RULE-BASED UNDERSTANDING

- An Input Node represents an expected user question (intent, class). Multiple, different user questions match to the same, single Input node.
 - Contextual: Different Input Nodes active during various stages of Conversation
- User phrase is matched to an Input Node either based on
 - Matching Rules like " \$ parking " (phrase contains "parking")
 - Example phrases used to train the Natural Language Classifier running in the background

Hit Count



Input nodes ordered by frequency of being hit within a period of time

STATISTICAL VS. RULE-BASED UNDERSTANDING

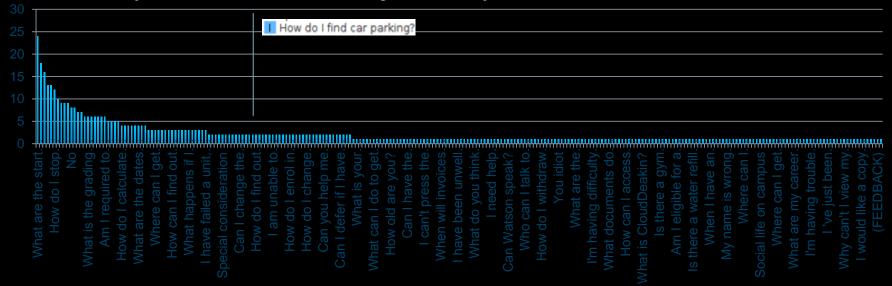
- An Input Node represents an expected user question (intent, class). Multiple, different user questions match to the same, single Input node.
 - Contextual: Different Input Nodes active during various stages of Conversation
- User phrase is matched to an Input Node either based on
 - Matching Rules like " \$ parking " (phrase contains "parking")
 - Example phrases used to train the Natural Language Classifier running in the background

Hit Count

Fat Head: Long Tail:

Statistical Classifier best here

Matching Rules or Info Retrieval best



Hybrid: high-precision rules cover "exact matches" and take precedence, rest is trained

Input nodes ordered by frequency of being hit within a period of time

WHAT IS DIALOG MANAGEMENT?

- **Dialog manager (DM)** is the core component of any conversational system. It solicits input from the user (ASR+NLU, text, GUI), retrieves the content (DB, QA systems,) and decides on the best response to the user based on certain dialog strategy
- 4 key metrics (need optimal mix of all)
 - Task completion rate (maximize)
 - Task completion time (minimize)
 - User perception/satisfaction (maximize)
 - User cognitive demand/load (minimize)
- Directed-dialogs used today
 - System completely controls the initiative (user just answers system's (simple) questions) to minimize impact of speech recognition errors
- Prompt:
 You mean Mary Lou
 at 9am tomorrow, right?

 action=new_meeting
 time = 9 (am | pm)
 invitees = her

 schedule a meeting
 with her at nine

- Mixed-initiative dialogs are next generation
 - Both user and system can drive the dialog
 - User can provide complex, context dependent responses
 - Feels quite natural but has many challenges

THE NEED OF DIALOG

- The desire of dialog management extend beyond information retrieval
- We see dialog as mean for interacting with all kinds of knowledge
 - databases, ontologies, knowledge graphs, unstructured data...
 - on desktops, tablets, phones, cars, robots, IOT...
 - affective, situated, multi-modal, multi-party, multi-app, cross-device, iot
 - various tasks
 - Persuasion
 - Diagnosis/remediation
 - Recommendation
 - Procedural (i.e. transactional) tasks
 - Tutoring
- New deployment models for dialog systems
 - New ways of integrating dialog management with existing applications
 - Entirely new kinds of dialog-based applications
- We work towards a flexible general purpose dialog system for the Watson family of products and services

DIALOG IN CONTEXT

Car

Ellipsis Anaphora K2Q and in life insurance? discount car insurance What is What is covered in covered in What discounts are available **How much does it cost?** life insurance in car insurance How much does car insurance cost? Insurance? **Complex Previous Spurious** Focus Question **Questions** OK very good how do I get one? What is the cost and Is theft covered? how do I get car insurance discounts available? Is theft covered What is the cost of car insurance **Under car insurance** What discounts are available in car insurance **Everything costs so much now a days** how much does this cost? Over Specification/ how much does car insurance cost?

Watson Dialog API

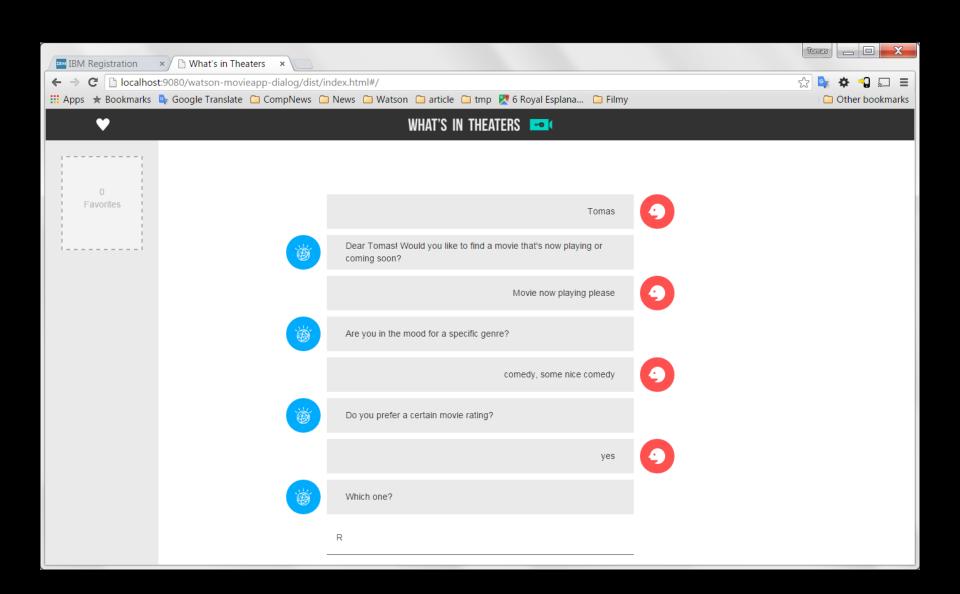
The WCDS API is a HTTP API that allows another system to make requests to applications configured on a WCDS instance.

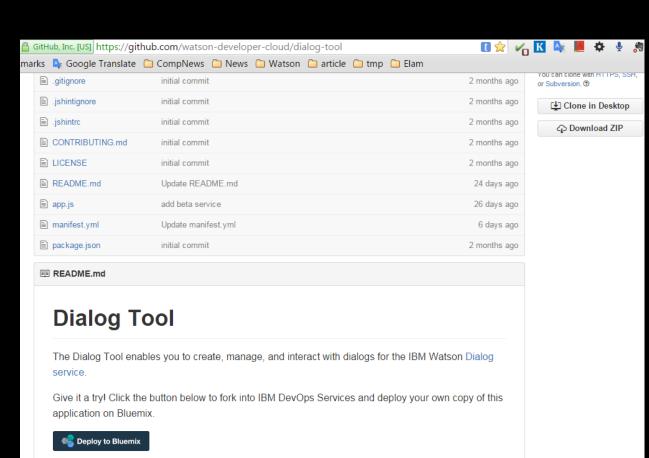
Content : Upload, Update, or Delete existing dia	Show/Hide List Operations Expand Operations
рит /dialogs/{dialog_id}/content	Used to update content for specified nodes.
GET /dialogs/{dialog_id}/content	Get content for all Input and Output nodes.
POST /dialogs	The dialog_id returned is used for subsequent calls to the API
GET /dialogs	Provides a list of Dialog applications associated with a service instance.
DELETE /dialogs	Used to close entire Dialog instance. This permanently removes all associated data.
РИТ /dialogs/{dialog_id}	Upload an existing dialog file to the system.
/dialogs/{dialog_id}	Close an existing dialog file. This permanently removes all associated data.
Conversation : Start new conversations, Obtain responses, Find Session Data	
	Show/Hide List Operations Expand Operations
/dialogs/{dialog_id}/conversation Used to obtain a response from the system for a submitted input message. Also used to start new conversations.	
GET /dialogs/{dialog_id}/conversation	Used to obtain chat session data dump for a given date.
Profile : Get or Set profile variables	Show/Hide List Operations Expand Operations
/dialogs/{dialog_id}/profile	Set the values for profile variables.

WHAT'S IN THEATERS

What can I call you?

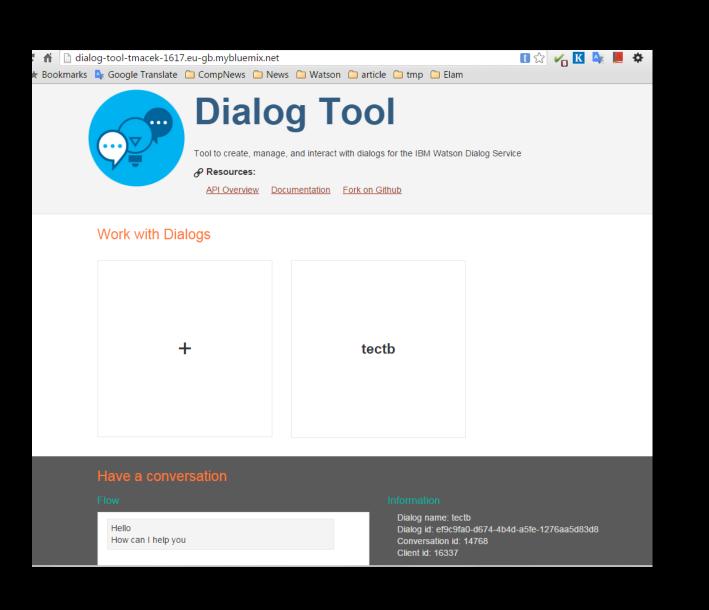
Tomas Chat





Getting started

 Create a Bluemix account. Sign up in Bluemix or use an existing account. Watson services in beta are free to use.



COGNITIVE AVATAR

Customer: Technology Exploration Center, Software Group

Description:

- Talking head with synchronized lips and numerous face expressions
- Body gesture recognition
 User waves to get attention
 User moves head forward to mute the system
- Implements six selected dialogue domains (small talk, weather, time, name days, local space navigation, education).
- WDS backend component permits fast authoring and maintenance
- Grammar + Remote dictation; Dictation + NLU
- Remote microphone, microphone array, techniques of opening microphone based on noise and state of dialogue
- Situation awareness (number of people and ambient nose level considered)
- Proactive attention request activities

(no people around) It is 12:00, time for lunch (shows menu on the screen) (person passing by) Hi, do you know .. (stops talking when no attention is drawn)



Hi, this is John Hi John

AIML

NLU

What is the weather forecast It is going to be sunny in low 30?

And next day?

It will rain Tomorrow.

Where can I find a rest room?

The rest room is at the end of the corridor, do not forget your batch (shows location on the screen)

Dialog Management

http://www.ibm.com/cz/research_en/

THANK YOU!