

Inside our AI Pipeline



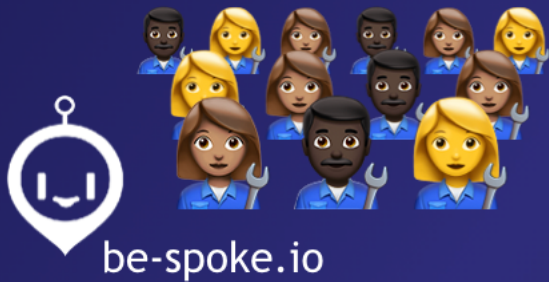
Putting the Natural Back Into Artificial
Intelligence to Move at Light Speed

Christine Gerpheide
Bespoke, VP of Engineering
chris@be-spoke.io

about me



amazon



BESPOKE



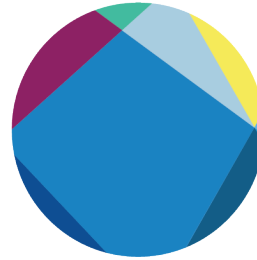
#HopperDownUnder

to move fast ➡ experiment



HUMANS

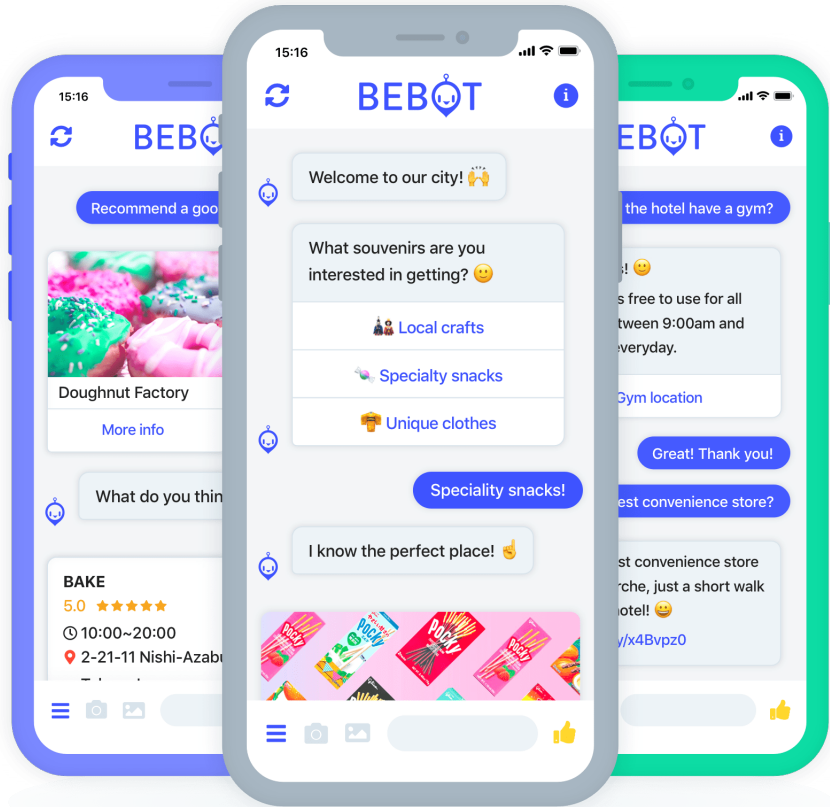
Leverage humans before
coding effort



CODE HOOKS

Leverage bad code before
more coding effort

Bebot by Bespoke



CUSTOM CHATBOTS at airports, train stations, hotels, and cities across Japan

FAQS, RESTAURANTS, EMERGENCY INFORMATION, AND MORE
When automation isn't available, human operators respond

> 30,000 USERS EVERY DAY getting help as a tourist

Humans define example questions and answers.

Answer:

"Check out is at 9AM"

Question phrasings:

"when is check out?"

"When can I check-out?"

"How late is checkout?"

01

BESPOKE

02

Train an ML model on the question phrasings.

Label, aka Class:

WHEN_IS_CHECK_OUT

WHERE_IS_THE_HOTEL

...

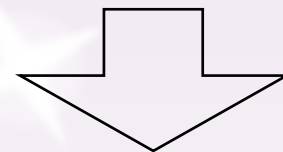
Data:

(Example questions for each label)

When a new user query arrives, predict which question it represents.

"ohi when could I maybe check out?"

Thanks :)"



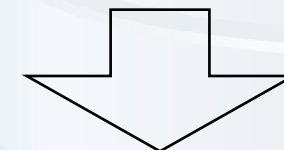
WHEN_IS_CHECK_OUT

03

04

Look up the pre-defined answer for that question.

WHEN_IS_CHECK_OUT

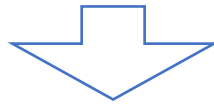


"Check out is at 9AM"

02

Train a machine learning model on the question phrasings.

```
where_is_the_hotel : [
    'Where is my hotel?',
    'Hotel location?'
]
when_is_checkout_time : [
    'When is check-out time?',
    'When do I need to check out?'
]
```

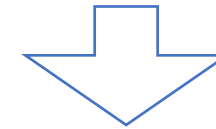


```
where_is_the_hotel : [
    'hotel' : 2,
    'location' : 1,
    'when' : 0
    ...
]
when_is_checkout_time : [
    'when' : 2,
    'check' : 2,
    'location' : 0
    ...
]
```

Which class is most similar?



“ohi how get 2 hotel kthx _(ツ)_/”



```
[
    'ohi' : 1,
    'how' : 1,
    '\_(ツ)_/' : 1
    ...
]
```

where_is_the_hotel: 85%
when_is_checkout_time: 32%

03

When a new user query arrives, predict which question it represents.

04

Look up a pre-defined answer

01

Define example question and answers

BESPOKE



Experiment
Quickly test out an idea and see if it
actually improves your product.

Welcome to Japan! What can I help you with? 😊

 To Tokyo

 Food

 Lounges

 Shopping

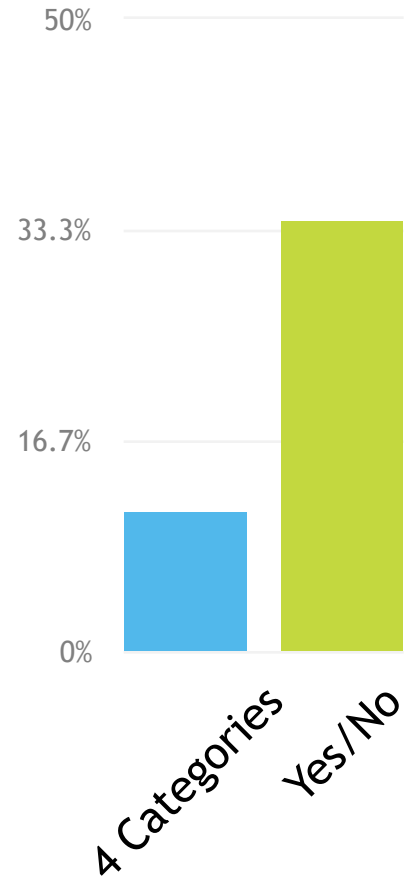


Welcome to Narita 🙌 I'm your AI travel buddy! Do you need flight info? 📍

 Yes

 No

Used the bot at least once

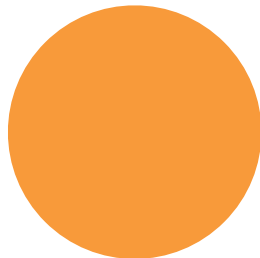


humans



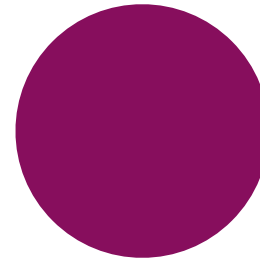
DESIGN CONVERSATIONS

Humans are good at sensical conversation flows.



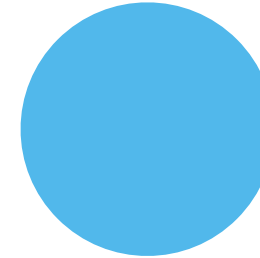
DEFINE EXPERIMENTS

Make hypotheses about human behavior and test them.



PERFORM TASKS NOT YET AUTOMATED

When considering a new feature, have humans chat with users first.



EVALUATE AUTOMATED REPLIES

Operators give you immediate feedback on algorithm performance.

Welcome to Narita. 🙌 I'm your AI travel buddy! Do you need flight info? 📌

✈️ Yes

😬 No

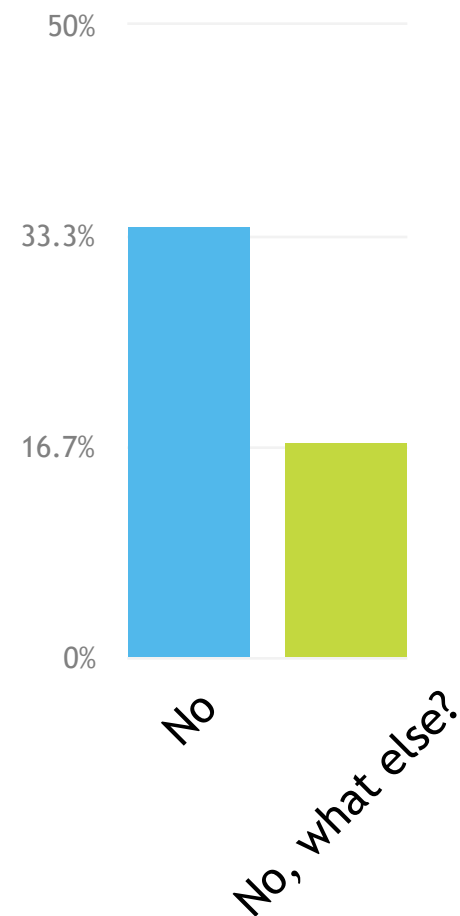


Welcome to Narita. 🙌 I'm your AI travel buddy! Do you need flight info? 📌

✈️ Yes

😬 No, what else can you do?

Used the bot at least once





Code hook

An integration point that allows you to insert a different code path, without altering the rest of the application.

“i want to go to a pokémon centre”

pokémon ➡ pokemon
(convert to ASCII)

ok 🐅 thanks! ➡ ~~okthanks!~~
ok thanks!

Accuracy
improvement


20%

13.3%

6.7%

0%





```
class MyApp {  
    function getAnswer(String question) {  
  
        // normal algorithm  
        result = new NaiveBayesAlgorithm().predict(question);  
        return result;  
    }  
}
```

```
class MyApp {  
  
    function getAnswer(String question) {  
  
        if (NEW_ALGORITHM_ENABLED) {  
            return new NewAlgorithm().predict(question);  
        }  
  
        // normal algorithm  
        result = new NaiveBayesAlgorithm().predict(question);  
        return result;  
    }  
}
```

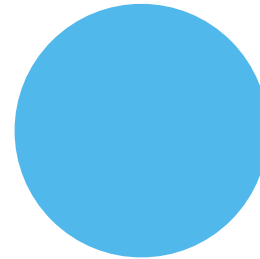


code hooks



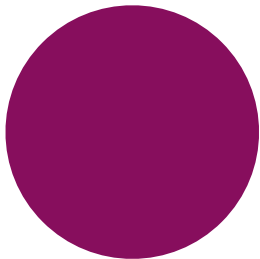
RUN EXPERIMENTAL ALGORITHMS

After benchmark test sets, enable in production for a short period of time.



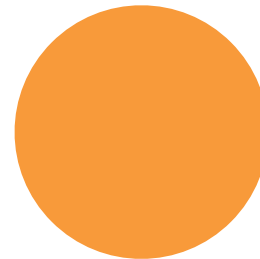
RUN UNOPTIMIZED CODE

Ugly code? Memory-intensive code? See if it works before improving it.



COMPARE ALGORITHMS

Log the results of both the old and the new algorithms.

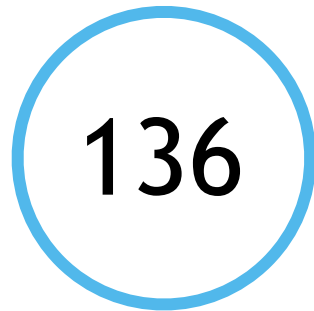


USE NEW TECHNOLOGIES

Code hooks can run independently of your existing stack.

ML pipeline + humans and code hooks

To support experimentation and bring guests the most natural user experience.



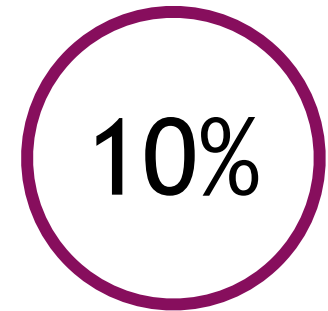
Experiments
run in 2019H1



Average
experiments
run per day



Average days
between a new
idea and trying
it in prod



Hotel guests make
reservations
(via human operators)

Thanks!

Putting the Natural Back Into Artificial
Intelligence to Move at Light Speed

Christine Gerpheide
Bespoke, VP of Engineering
chris@be-spoke.io

We're hiring in Tokyo!

- Chatbot R&D Lead (ML/NLP)
- Engineers
- ...and more

bebot.io/jobs

