digital
taxonomy
Smart Software for Market Research

Robots to the Rescue:
Decoding the Real Voice of the Customer Using Leading-edge AI

Tim Brandwood, CTO

We make smarter software for MR

codeit
Coding. Only Better
from digital taxonomy

- Automated coding of open-ended text answers
- Unique "blended" model combines multiple AI techniques
- Better productivity, speed and quality

loadit
Data Processing Made Simple
from digital taxonomy

- ETL for mere mortals!
- Automate repetitive data processing tasks with a simple drag and drop workflow
- Fully customisable and extensible
- Extensive data transformation tools

digital taxonomy
Smart Software for Market Research

Codeit: Dealing With Text Data

Q: Why did you not enjoy your stay at Hotel X?

A: Staff were rude, sheets were dirty, noisy stag party. I hate Hotel X

- ✓ Staff behaviour
- ✓ Issues with linen
- ✓ Problems with other residents
- ✓ General negative brand attitude

30% Staff behaviour

28% Issues with linen

15% Problems with other residents

8% General negative brand attitude

Wouldn't it be great...

"big screen amazing sound system"

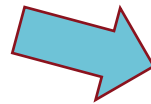*"Large screen and good sound systems"*

"Louder noise, bigger screen, more immersive"

"The occasion, sound, and screen size"

"The sound is much better and the wide screen is great"

"Larger screen and atmosphere"

Verbatim comments

Insights

Wouldn't it be great...

"big screen amazing sound system"
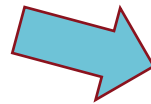
"Large screen and good sound systems"

"Louder noise, bigger screen, more immersive"

"The occasion, sound, and screen size"

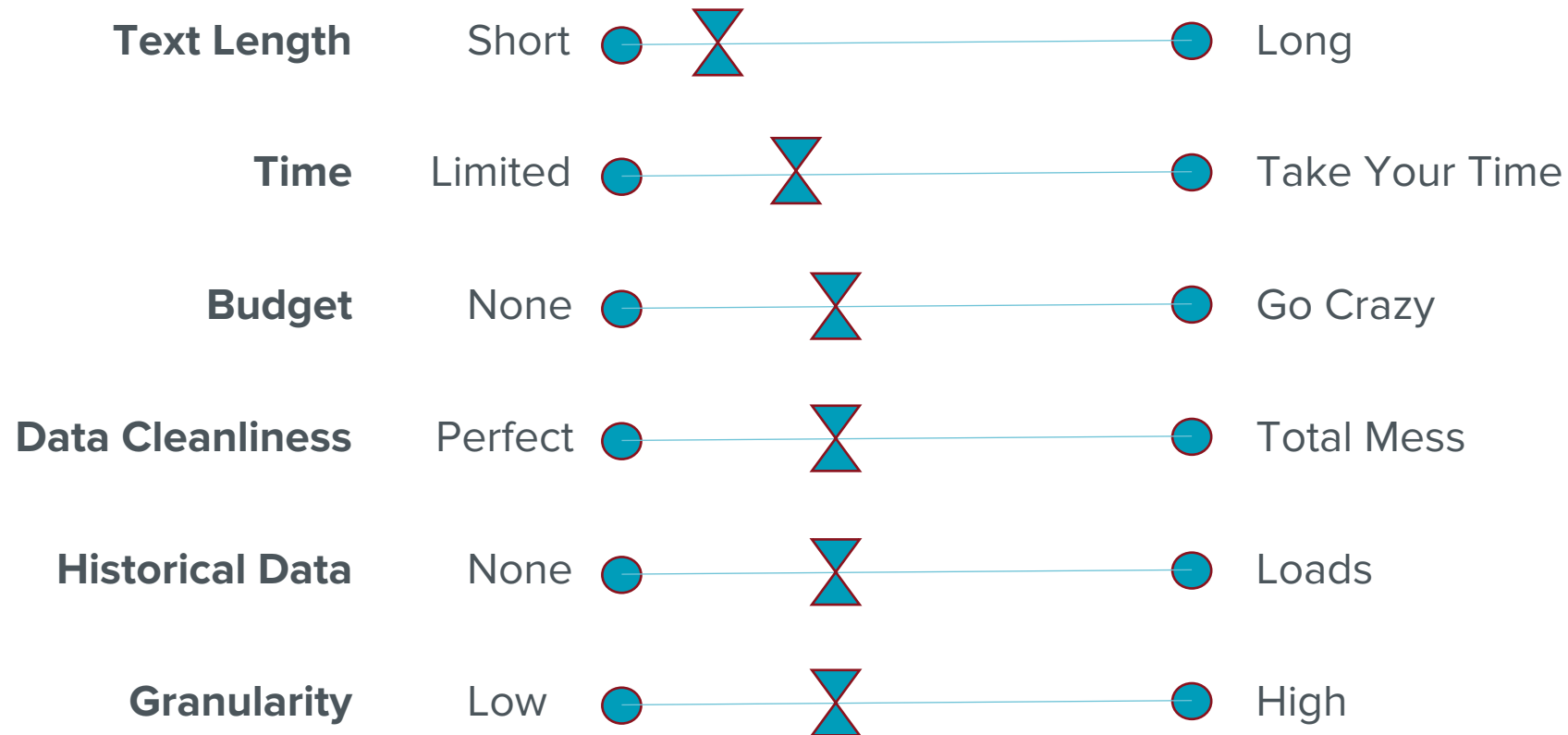*"The sound is much better and the wide screen is great"*

"Larger screen and atmosphere"

Verbatim comments

1. big screen
2. Large screen
3. bigger screen
4. screen size
5. wide screen
6. Larger screen
7. good sound systems
8. Louder noise
9. more immersive
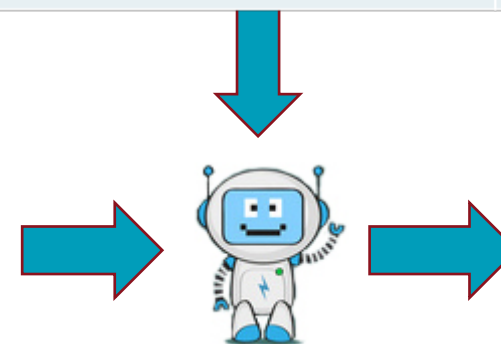10. atmosphere

The Real-World Is Complicated

Text Length   Short   ●━━━━━━━━━━━━━━━●   Long

Time   Limited   ●━━━━━━━━━━━━━━━●   Take Your Time

Budget   None   ●━━━━━━━━━━━━━━━●   Go Crazy

Data Cleanliness   Perfect   ●━━━━━━━━━━━━━━━●   Total Mess

Historical Data   None   ●━━━━━━━━━━━━━━━●   Loads

Granularity   Low   ●━━━━━━━━━━━━━━━●   High

Using Historical Data

Training Data

Verbatim	Codes
"Bigger screen better sound"	2;3
"Larger screen and atmosphere"	2;1
"the surround sound and screen size"	3;2
"the big screen"	2
"Love the big screen"	2
...	...

"big screen amazing sound system"
"Large screen and good sound systems"
"Louder noise, bigger screen, more immersive"
"The occasion, sound, and screen size"
"The sound is much better and the wide screen is great"
"Larger screen and atmosphere"

1. Atmosphere
2. Bigger screen
3. Better Sound
4. Immersive
5. Night out

Verbatim comments

Code labels

Using Historical Data

Training Data

"big screen amazing sound system"
"Large screen and good sound systems"
"Louder noise, bigger screen, more immersive"
"The occasion, sound, and screen size"
"The sound is much better and the wide screen is great"
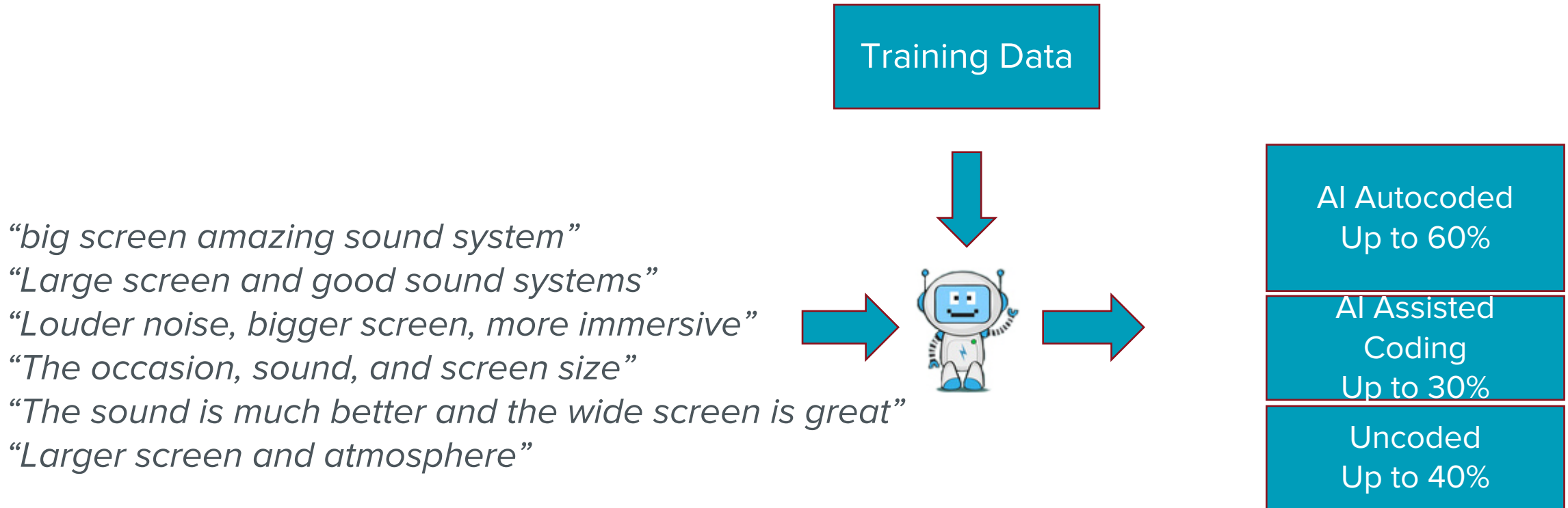"Larger screen and atmosphere"

Verbatim comments

AI Autocoded
Up to 60%

AI Assisted
Coding
Up to 30%

Uncoded
Up to 40%

Codeit - AI Assisted Coding

Codeit – Uncoded Items

Training Data

"Quadrophonic sound adds terrific atmosphere to a movie and a gisnt screen rungs it closer to home!!"

?

"The hole atmosphire of watching it with everyone else. The same wow and Achs"

Verbatim comments
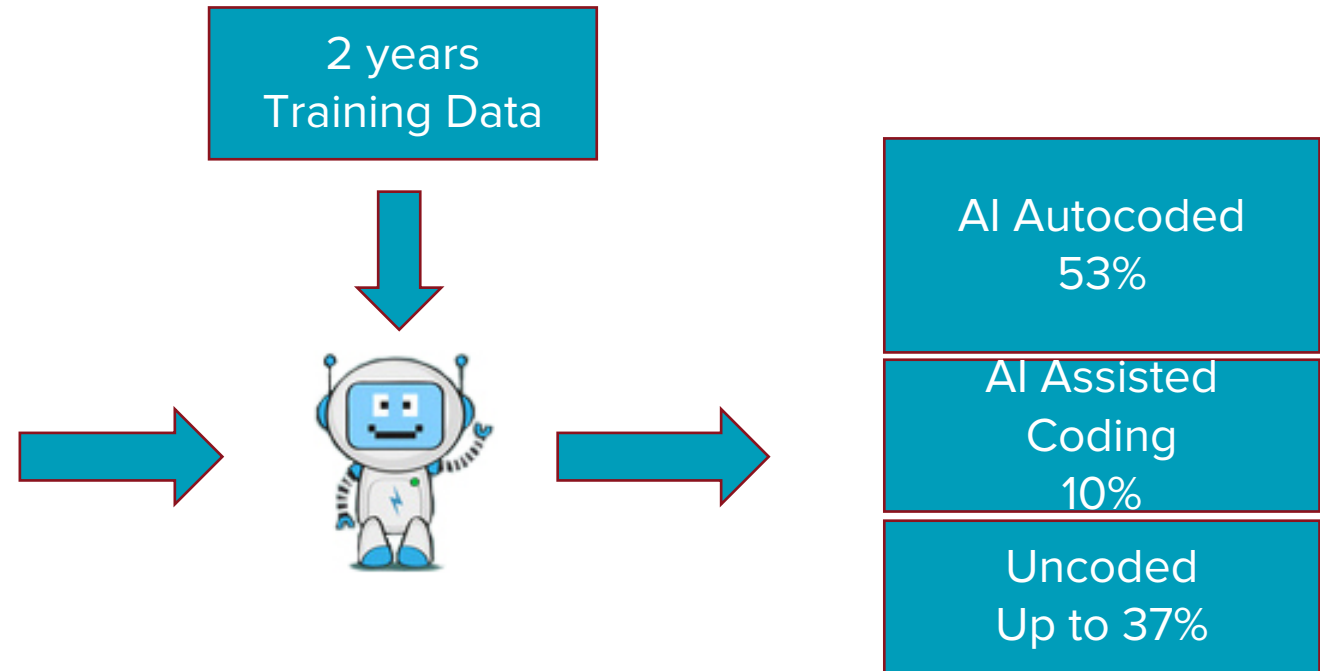
Client Case Study: NEC Group

NEC Exhibition Centre, Birmingham

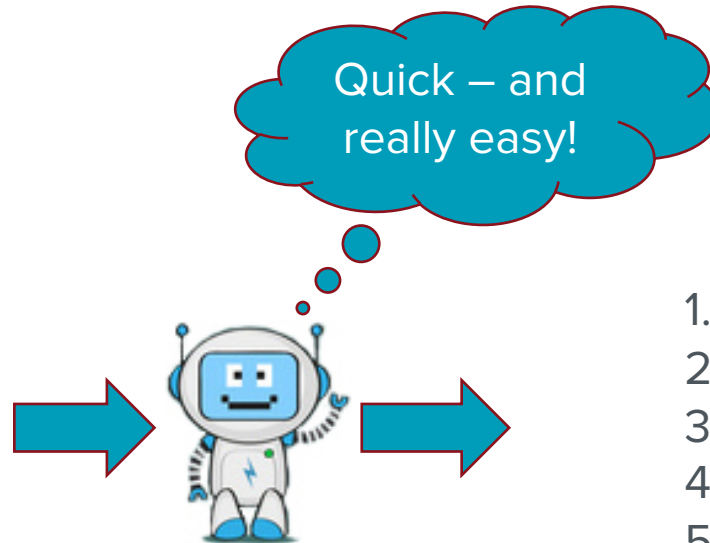★★★★★

WE VALUE YOUR FEEDBACK
TELL US HOW DID WE DO

Please note: We do NOT share your information with third parties, we do NOT share your email addresses or contact details with sponsors or any third parties, and we do NOT run exclusive 'sponsored' emails on behalf of third parties.

2 years
Training Data

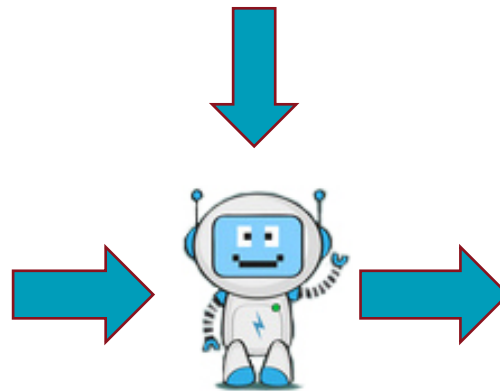AI Autocoded
53%

AI Assisted
Coding
10%

Uncoded
Up to 37%

Short Responses – The "Long Tail"

Enhance the AI by defining
text matching rules

```
mc|donal|m.?+c\sdon
```

"mcdonald's"
"Mac Donald"
"mcdonald"
"mc donald"
"mc"
"mc'donald"
"mec donald"
"mec donalw"
"Mec dondalds"
"mecdonalds"

1. McDonalds

Free-text   Coded output

Avoiding Overmatching

Avoid overmatching by
defining exceptions

```
new
```

"I like the new hotel"
"The new facilities are great"
"The new menu is delicious"
"The best hotel in New York"
"The new décor is brilliant"
*"New Delhi needs more places
like this!"*

1. New
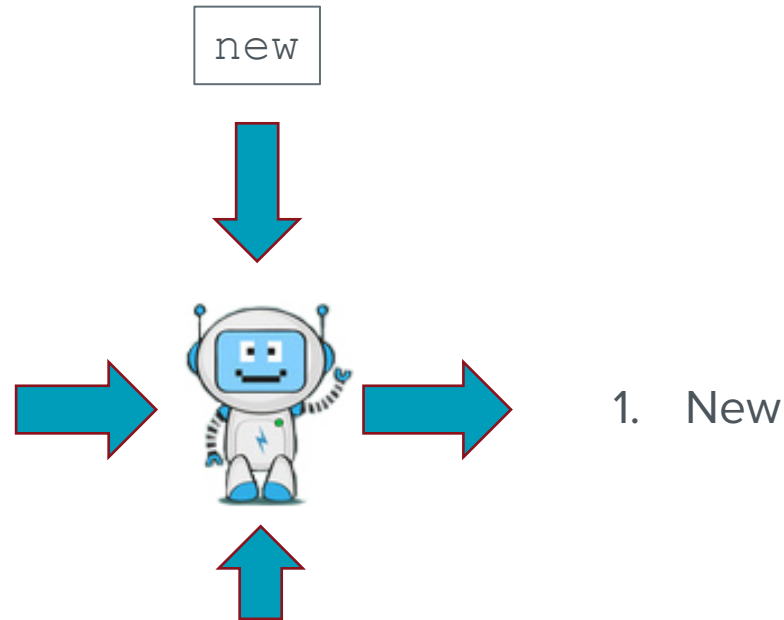
```
EXCEPT: new\s+?(york|zealand|delhi|orleans)
```

Free-text   Coded output

Using Text Analytics

No Training
Data

"big screen amazing sound system"
"Large screen and good sound systems"
"Louder noise, bigger screen, more immersive"
"The occasion, sound, and screen size"
"The sound is much better and the wide screen is great"
"Larger screen and atmosphere"

Verbatim comments
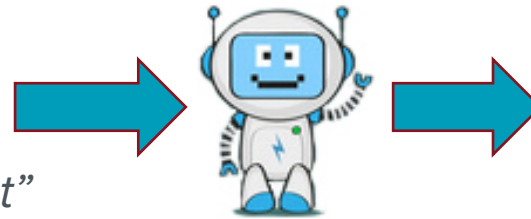
1. big screen
2. Large screen
3. bigger screen
4. screen size
5. wide screen
6. Larger screen
7. good sound systems
8. Louder noise
9. more immersive
10. atmosphere

Coded output

The Codeit "Blended AI" Approach

VERBATIM TEXT
They give good service, but it was expensive and the staff were rude.

TEXT MATCHING
Simple Text Matching against previous examples

RULES MATCHING
User-defined text patterns

TEXT ANALYTICS
Topic and Sentiment Extraction

MACHINE LEARNING
Trained Machine Learning algorithm(s)

HUMAN CODERS
Human Coders to help train the AI

CODED RESPONSE
Code 3: Good service
Code 7: Poor value/Too expensive
Code 10: Staff attitude/Issues

Case Study – BVA BDRC

- 50,714 customer comments rating Hotels

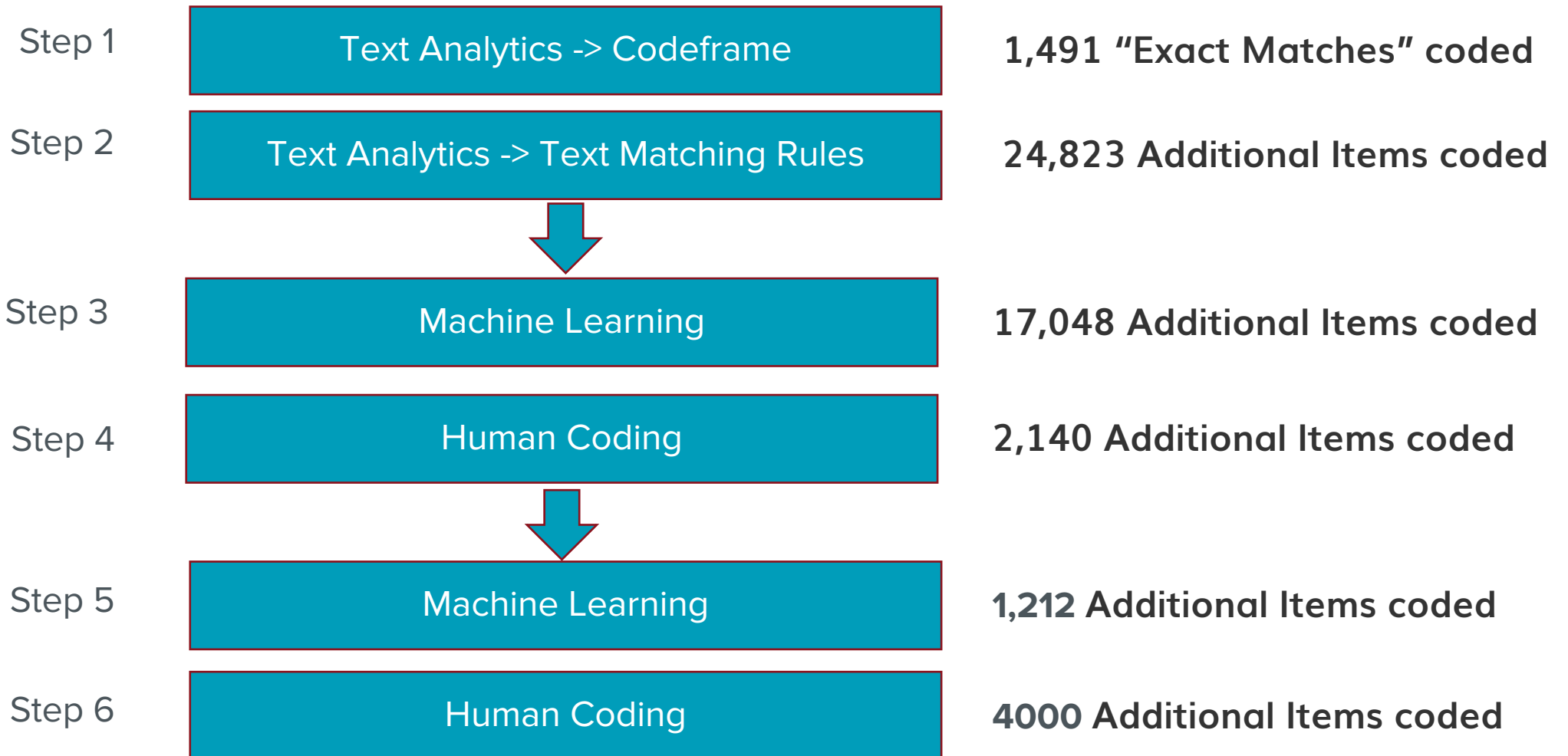- Tight time and budget constraints

- Applied the Codeit "Blended" approach

Case Study: Codeit "Blended" Solution

Step 1	Text Analytics -> Codeframe	1,491 "Exact Matches" coded
Step 2	Text Analytics -> Text Matching Rules	24,823 Additional Items coded
Step 3	Machine Learning	17,048 Additional Items coded
Step 4	Human Coding	2,140 Additional Items coded
Step 5	Machine Learning	1,212 Additional Items coded
Step 6	Human Coding	4000 Additional Items coded

What next for AI in the Survey World?

- *"Life begins at a billion examples"*

- *Current training data and AI model is siloed*

- *Quantity of data in each silo is relatively small*

Coder → Survey → AI Model → Training Data
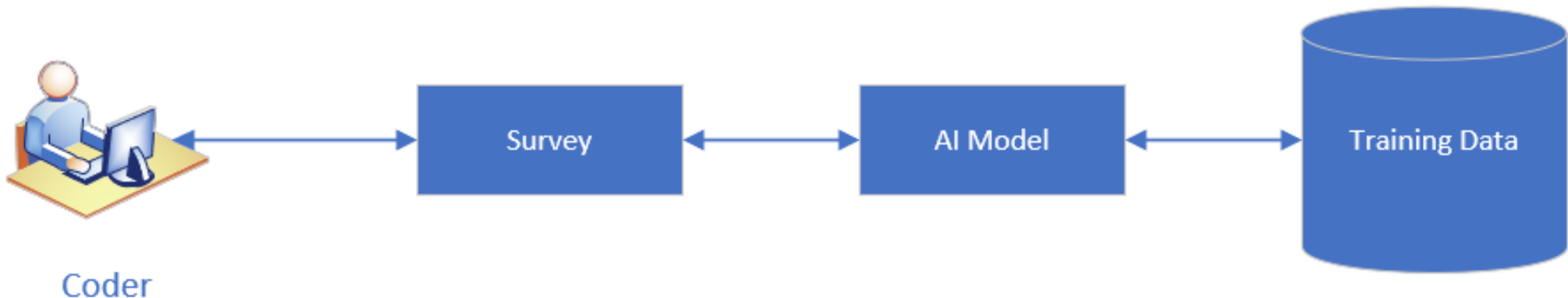
What next for AI in the Survey World?

- Codeit architecture can "pool" data to enhance AI training

- Pooling data within a company, across companies or sector-wide would turbocharge AI for Market Research

Conclusions

- It is a mistake to look for a single "silver bullet"

- The best results are achieved by blending techniques to suit the task at hand

- It is a mistake to think of AI as simply replacing humans.

    For now, AI is optimized if it is used to assist humans rather than trying to do everything for them.

- This "new wave" of AI is only 5 years old ... Plenty of room for further innovation in models and techniques

digital taxonomy
Smart Software for Market Research

Thank you!

Tim Brandwood
tim@digitaltaxonomy.co.uk
www.digitaltaxonomy.co.uk