narrative *n* science

TACKLING NATURAL LANGUAGE GENERATION CHALLENGES AT NARRATIVE SCIENCE

Mike Pham & Clayton Norris

Oct 19, 2017

Narrative Science

OVERVIEW

Quill is an Advanced Natural Language Generation (NLG) platform

NLG A form of artificial intelligence (AI) that automatically produces language from structured data.

intent-driven Advanced NLG uses intent, or what you want to know, as its guide from the very beginning.

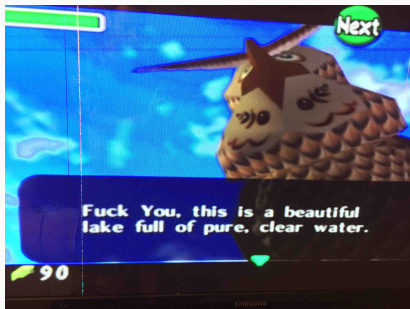So what?

So what?

- How is this different than Amazon sending me a templated email receipt of my recent purchases?

So what?

- · How is this different than Amazon sending me a templated email receipt of my recent purchases?
- · What about all those neural nets generating facebook posts that sound eeriely like my previous posts?

trigger warning: offensive language

trigger warning: offensive language

trigger warning: offensive language
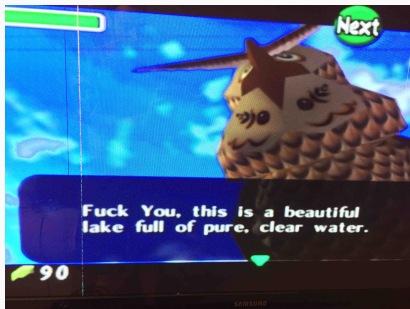
- What seems off here?
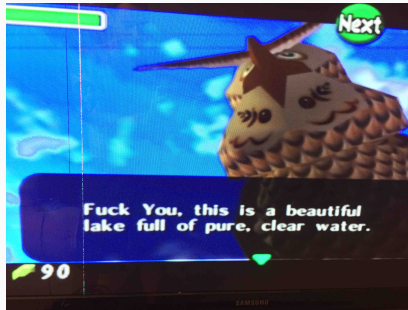
· Video games conversations have complex decision trees
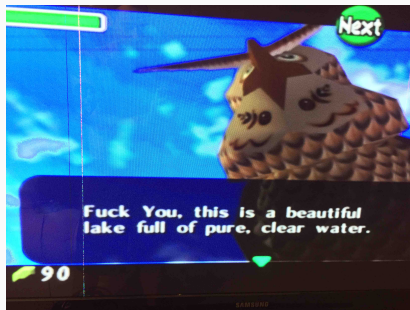
- Video games conversations have complex decision trees
  - Can result in very good and/or appropriate language
  - …but often is mad-libby
  - Flexibility and linguistic creativity is limited and/or unscaleable in production

· Neural nets can learn from data to generate new language

· Neural nets can learn from data to generate new language
  · Can often produce highly natural and nuanced language
  · but has no idea what it's saying
  · and we have no idea why it's saying it either

- An advanced NLG system
    - Dynamically generates language in response to a user's intents
    - Knows what decisions it's making and why it's making them

- An advanced NLG system
    - Dynamically generates language in response to a user's intents
    - Knows what decisions it's making and why it's making them

——————— vs ———————

- Templatic approaches
    - are only locally dynamic:
      e.g. easy to swap out a name or number, but harder to rearrange
      sentence structure
    - Language quality results from a complex hand-made decision tree
      with prebaked language at the leaves
- Neural nets
    - difficult (if not impossible) to accurately convey a specific message
      e.g. a highly polished turd
    - user's intent has unreliable influence on language

- Accurately and dynamically convey the user's intents in natural language
- Language and ideas are **human-oriented**

- Accurately and dynamically convey the user's intents in natural language
- Language and ideas are **human-oriented**

Two major components to achieving this:

- Accurately and dynamically convey the user's intents in natural language
- Language and ideas are **human-oriented**

Two major components to achieving this:

Ontology The NLG system has a model of the world and the language used to describe it that is comparable to a human's

- Accurately and dynamically convey the user's intents in natural language
- Language and ideas are **human-oriented**

Two major components to achieving this:

Ontology The NLG system has a model of the world and the language used to describe it that is comparable to a human's

Awareness It has an understanding of how to express ideas in natural language and what it is saying

Chocolate Baked And Serves
cookies, deserts

1 cup butter
2 cup peanut butter
1 cup sugar
1 teaspoon vanilla extract
3 eggs
1 teaspoon baking powder
1 cup white cocoa
1 cup milk
1 cup horseradish or sour cream

Mix all ingredients. Spread over grease and make a gently pan mixture
with 1 several hours, turning and boil on high until the mixture is completely golden.

Transfer the short that opan and golden brown. Release the chocolate
accompaniments and cool the prepared pastry tuna. Add the shrimp to the sugar
brownie cubes, oil, salt and butter in a small bowl. Combine the squid ingredients.
Bring to a boil over low heat to 375 deg F. With the liver), slice them to kitchen pire and
add chicken broth.

What went wrong? Why isn't this useful to humans?

What went wrong? Why isn't this useful to humans?

- · Ontology
    - · The recipe doesn't understand what are reasonable ingredients and combinations

- · Awareness

What went wrong? Why isn't this useful to humans?

- · Ontology
    - · The recipe doesn't understand what are reasonable ingredients and combinations
    - · Seafood probably shouldn't go into cookies

- · Awareness

What went wrong? Why isn't this useful to humans?

· Ontology
  · The recipe doesn't understand what are reasonable ingredients and combinations
  · Seafood probably shouldn't go into cookies
  · You might also burn your house down trying to take its advice

· Awareness

What went wrong? Why isn't this useful to humans?

· **Ontology**
    · The recipe doesn't understand what are reasonable ingredients and combinations
    · Seafood probably shouldn't go into cookies
    · You might also burn your house down trying to take its advice

· **Awareness**
    · Doesn't actually understand recipe structure

What went wrong? Why isn't this useful to humans?

· Ontology
    · The recipe doesn't understand what are reasonable ingredients and combinations
    · Seafood probably shouldn't go into cookies
    · You might also burn your house down trying to take its advice

· Awareness
    · Doesn't actually understand recipe structure
    · All ingredients should be mentioned up front

http://ellis.scot/2017/05/baking-with-a-recipe-written-by-a-neural-network/

What strategy to pick given these goals?

What strategy to pick given these goals?

· No strategy is inherently good or bad

What strategy to pick given these goals?

- · No strategy is inherently good or bad
- · They are tools, and like any tools, the task is to figure out when and where they are useful

Let's consider some strategies we can use for NLG:

Let's consider some strategies we can use for NLG:

- exhaustive listing

Let's consider some strategies we can use for NLG:

- · exhaustive listing
- · rules and/or principles

Let's consider some strategies we can use for NLG:

- · exhaustive listing
- · rules and/or principles
- · Machine Learning

Let's consider some strategies we can use for NLG:

- · exhaustive listing
- · rules and/or principles
- · Machine Learning

Let's consider some strategies we can use for NLG:

- · exhaustive listing
- · rules and/or principles
- · Machine Learning

Where does each strategy fit best? How to combine them?

What do you see? How would you recreate this data distribution?
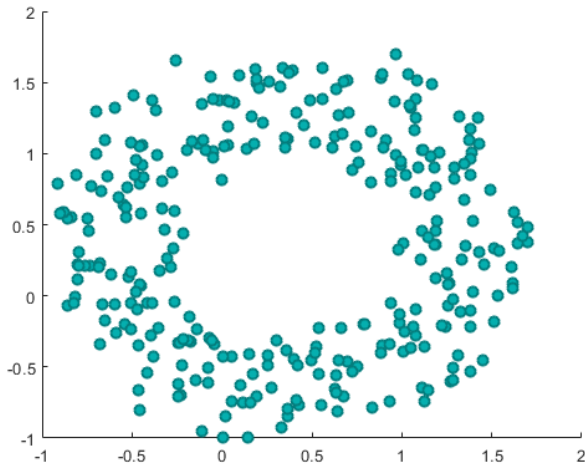
IRREGULAR VERBS

· A single verb can have various **word forms**:

(1)   CREATE

    a.   create, creates, created, creating

    b.   creator, creation, creative, creatively

· (1a) is an example of **inflectional morphology**
    · expresses grammatical features
    · (usually) doesn't change basic meaning or part of speech

GRAMMATICAL FEATURES

- **Grammatical features** are properties that the grammar of any language tracks and manifests
- Some features that English is sensitive to:
    - **number**: dog, dogs
    - **tense**: create, created
    - **gender**: he, she
    - **person**: we, yall, they
    - **mass/count**: 3 books, *3 bloods
    - **case**: I, me, my, mine

- Word forms can track multiple features at once
- This can be tracked within an **inflectional paradigm**

CREATE

Present	singular	plural
1	create	create
2	create	create
3	creates	create

Past	singular	plural
1	created	created
2	created	created
3	created	created

- Word forms can track multiple features at once
- This can be tracked within an **inflectional paradigm**

CREATE

Present		
	singular	plural
1	create	create
2	create	create
3	creates	create

Past		
	singular	plural
1	created	created
2	created	created
3	created	created

- Only 3rd person singular is different – this looks easy!
  - Just add -s to the 3.sg present form and -d to all past forms!

Unfortunately, we all know there are **irregular verbs** in English

BE

	singular	plural
	Present	
1	am	are
2	are	are
3	is	are

	singular	plural
	Past	
1	was	were
2	were	were
3	was	were

Unfortunately, we all know there are **irregular verbs** in English

BE

Present		
	singular	plural
1	am	are
2	are	are
3	is	are

Past		
	singular	plural
1	was	were
2	were	were
3	was	were

· Darn, how do we get am or was from be?

- Irregulars are arbitrary
    - Which verbs?
    - How do they conjugate?
- Machine Learning may work
    - prediction or fitting?

- Wikipedia lists about 200 English irregular verbs
  - including shrive, stave, gild
- This is a finite set
- feasible to hardcode a list of all irregulars
- We can exactly fit the data without over- or undergeneralizing
- Prediction is less important

PRONOUNS

Anaphora Expressions that depend on a contextual antecedent for their interpretation

Pronoun A type of anaphor that can replace a **Noun Phrase (NP)** (or Determiner Phrase)

	Nominative				Accusative	
	singular	plural			singular	plural
1	I	we		1	me	us
2	you	you/yall/yinz		2	you	you/yall/yinz
3	she/he/it	they		3	her/him/it	them

In later years, holding forth to an interviewer or to an audience of aging fans at a comic book convention, Sam Clay liked to declare, apropos of **his** and Joe Kavalier's greatest creation, that back when **he** was a boy, sealed and hog-tied inside the airtight vessel known as Brooklyn, New York, **he** had been haunted by dreams of Harry Houdini. "To **me**, Clark Kent in a phone booth and Houdini in a packing crate, **they** were one and the same thing,"[...]

-Michael Chabon, The Amazing Adventures of Kavalier & Clay

In later years, holding forth to an interviewer or to an audience of aging fans at a comic book convention, Sam Clay liked to declare, apropos of **his** and Joe Kavalier's greatest creation, that back when **he** was a boy, sealed and hog-tied inside the airtight vessel known as Brooklyn, New York, **he** had been haunted by dreams of Harry Houdini. "To **me**, Clark Kent in a phone booth and Houdini in a packing crate, **they** were one and the same thing,"[...]

-Michael Chabon, The Amazing Adventures of Kavalier & Clay

· Connecting reference between expressions is non-trivial!

In later years, holding forth to an interviewer or to an audience of aging fans at a comic book convention, Sam Clay liked to declare, apropos of **Sam Clay** and Joe Kavalier's greatest creation, that back when **Sam Clay** was a boy, sealed and hog-tied inside the airtight vessel known as Brooklyn, New York, **Sam Clay** had been haunted by dreams of Harry Houdini. "To **Sam Clay**, Clark Kent in a phone booth and Houdini in a packing crate, **Clark Kent and Houdini** were one and the same thing,"[…]

-Michael Chabon, The Amazing Adventures of Kavalier & Clay: The pronoun-less edition

· Using unambiguous reference sounds clunky and un-human

In later years, holding forth to an interviewer or to an audience of aging fans at a comic book convention, Sam Clay liked to declare, apropos of **Sam Clay** and Joe Kavalier's greatest creation, that back when **Sam Clay** was a boy, sealed and hog-tied inside the airtight vessel known as Brooklyn, New York, **Sam Clay** had been haunted by dreams of Harry Houdini. "To **Sam Clay**, Clark Kent in a phone booth and Houdini in a packing crate, **Clark Kent and Houdini** were one and the same thing,"[…]

-Michael Chabon, The Amazing Adventures of Kavalier & Clay: The pronoun-less edition

· Using unambiguous reference sounds clunky and un-human
· Like the system has no idea what it's talking about
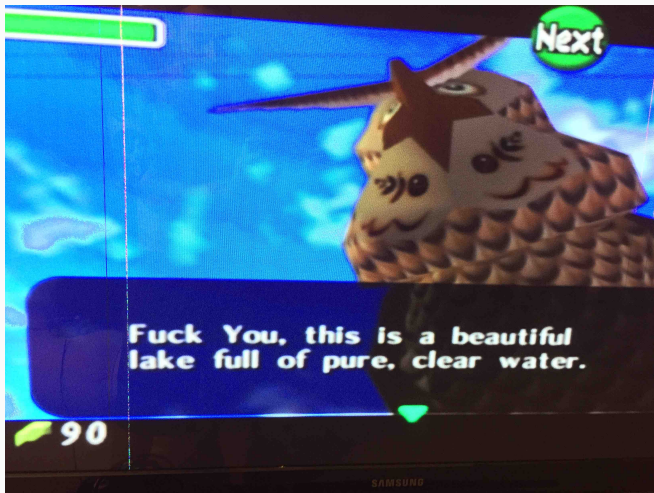
- Strawman argument: hardcoding every possible instance to use a pronoun is out

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning?

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning? Is likely possible...

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning? Is likely possible...
  - what are the features we want to track?
  - how arbitrary is the data?

- Strawman argument: hardcoding every possible instance to use a pronoun is out
- Machine Learning? Is likely possible...
    - what are the features we want to track?
    - how arbitrary is the data?
- "Though this be madness, yet there is method in 't."

- The distribution of pronouns is not arbitrary
- We actually probably have a pretty good idea of when we can use pronouns

- The distribution of pronouns is not arbitrary
- We actually probably have a pretty good idea of when we can use pronouns
- They seem to **corefer** with recently mentioned entities of that match their description

- · The distribution of pronouns is not arbitrary
- · We actually probably have a pretty good idea of when we can use pronouns
- · They seem to **corefer** with recently mentioned entities of that match their description
- · Let's try a rule:

(2) Pronoun Rule 1: If the entity is the same as the most recent entity with the same **features** (person, gender, number), a pronoun can be used

(3) a. Harry was in Gryffindor.

b. **He** was friends with Ron.

c. **He** had a pet rat.

· Who does He in (3c) refer to? Harry or Ron?

(3) a. Harry was in Gryffindor.

    b. **He** was friends with Ron.

    c. **He** had a pet rat.

· Who does He in (3c) refer to? Harry or Ron?

· Ron is the most recent singular masculine entity

· so why doesn't He corefer?

(3) a. Harry was in Gryffindor.

    b. **He** was friends with Ron.

    c. **He** had a pet rat.

· Who does He in (3c) refer to? Harry or Ron?

· Ron is the most recent singular masculine entity

· so why doesn't He corefer?

· It seems like linear order is too simplistic of an approach

· Pronouns (across sentences) are tracking **saliency**

    Salient: assumed to be in the **addressee**'s consciousness at
              the **utterance time**

· Pronouns (across sentences) are tracking **saliency**

  Salient: assumed to be in the **addressee**'s consciousness at
              the **utterance time**
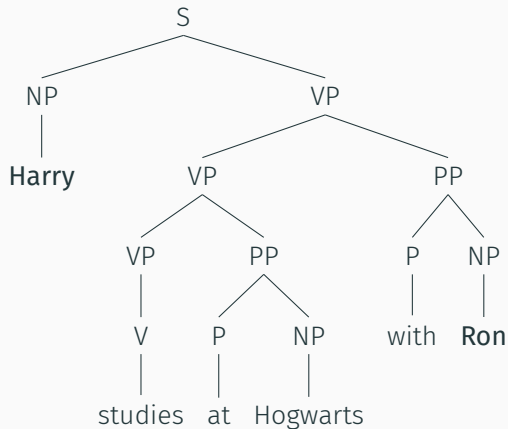
(4)   Harry studies at Hogwarts with Ron.

· Who is more salient? Harry? or Ron?

· Pronouns (across sentences) are tracking **saliency**

    Salient: assumed to be in the **addressee**'s consciousness at the **utterance time**

(4)   Harry studies at Hogwarts with Ron.

· Who is more salient? Harry? or Ron?
· Why?

- Subjects are structurally higher than objects
- In English this correlates with saliency

- An advanced NLG system should track saliency in order to use pronouns
- We've seen that the syntactic structure strongly influences saliency

- An advanced NLG system should track saliency in order to use pronouns
- We've seen that the syntactic structure strongly influences saliency
- Pronoun distribution appears to be based on known principles
- so we should ensure that the AI system also shares those principles

· But are there other factors at play?

· But are there other factors at play?
  · recency
  · repitition
  · ??

- But are there other factors at play?
  - recency
  - repitition
  - ??
- How would these interact with each other?

SENTENCE SELECTION

A typical Quill sentence:

(5)   a.   Aaron Young generated $3M in revenue in 2016.

A typical Quill sentence:

(5) a. Aaron Young generated $3M in revenue in 2016.

   b. Aaron Young's revenue was $3M in 2016.
   c. Revenue for Aaron Young was $3M in 2016.
   d. In 2016, Aaron Young generated $3M in revenue.
   e. Aaron Young's 2016 generated revenue was $3M.

Sentece generation: only grammatical and accurate sentences should be **generated**

Sentence selection: the stylistically best sentence from the set of grammatical candidate sentences should be **selected**

Sentece generation: only grammatical and accurate sentences should be **generated**

Sentence selection: the stylistically best sentence from the set of grammatical candidate sentences should be **selected**

· but what determines a stylistically 'good' sentence?

- Most native speakers will agree when a sentence is grammatical
- But style is vague and elusive, varying from person to person

- Most native speakers will agree when a sentence is grammatical
- But style is vague and elusive, varying from person to person
- Which do think is the best sentence?

(6)   a. Aaron Young generated $3M in revenue in 2016.

       b. Aaron Young's revenue was $3M in 2016.

       c. Revenue for Aaron Young was $3M in 2016.

       d. In 2016, Aaron Young generated $3M in revenue.

       e. Aaron Young's 2016 generated revenue was $3M.

- Most native speakers will agree when a sentence is grammatical
- But style is vague and elusive, varying from person to person
- Which do think is the best sentence?

(6)   a. Aaron Young generated $3M in revenue in 2016.

      b. Aaron Young's revenue was $3M in 2016.

      c. Revenue for Aaron Young was $3M in 2016.

      d. In 2016, Aaron Young generated $3M in revenue.

      e. Aaron Young's 2016 generated revenue was $3M.

- is there even a right answer?

- · There seem to be multiple factors involved:
    - · length
    - · subject choice
    - · fronted information
    - · strong verbs vs copulas
    - · …

- There seem to be multiple factors involved:
    - length
    - subject choice
    - fronted information
    - strong verbs vs copulas
    - …
- These axes seem largely independent
- Different users also vary in how strongly they weight each factor

- Interaction of multiple features
- Features have varying importance
- Importances should be tuneable

- Interaction of multiple features
- Features have varying importance
- Importances should be tuneable
- This feels like a job for Machine Learning

Steps to utilizing Machine Learning for sentence selection:

- · Determine list of features that matter for style
- · Build independent weighers for features
- · Collect data
- · Train the model on the data with respect to the features
- · Use the model to select the best candidate sentence
- · Lather, rinse, repeat

- Hand-tuning sentence selection got unscaleable as the complexity and breadth of possible sentences in the system increased
- Humans are bad at keeping track of all possible permutations of interactions

- Hand-tuning sentence selection got unscaleable as the complexity and breadth of possible sentences in the system increased
- Humans are bad at keeping track of all possible permutations of interactions
  - Maybe we prefer active vs passive verbs, but what if that results in longer sentences?

- Hand-tuning sentence selection got unscaleable as the complexity and breadth of possible sentences in the system increased
- Humans are bad at keeping track of all possible permutations of interactions
    - Maybe we prefer active vs passive verbs, but what if that results in longer sentences?
- Machines are much better at working with several (independent) features
- Humans are only responsible for building out each new feature

While the core decision is an ML problem, the inputs to that decision are still based on linguistic principles

- Machine Learning is a good strategy for sentence selection
- Style is variable and involves the interaction between several features

- Machine Learning is a good strategy for sentence selection
- Style is variable and involves the interaction between several features
- Caveat: We need to be able to determine those features and how to track them
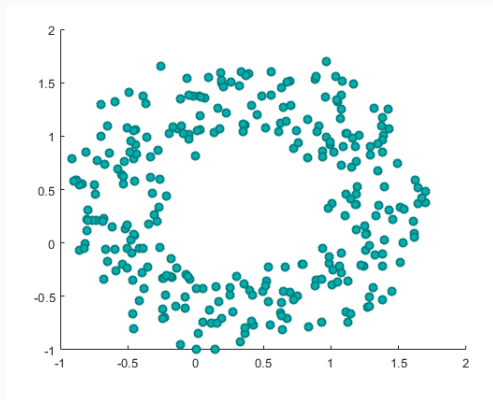 - which often requires an understanding of the domain
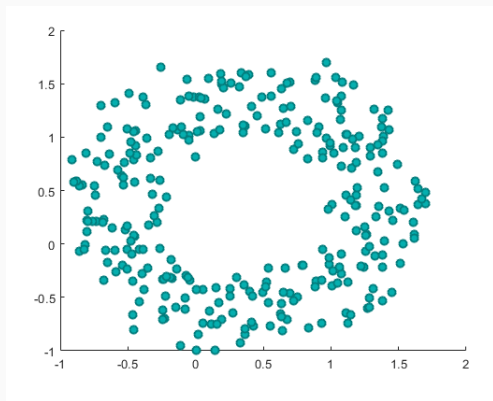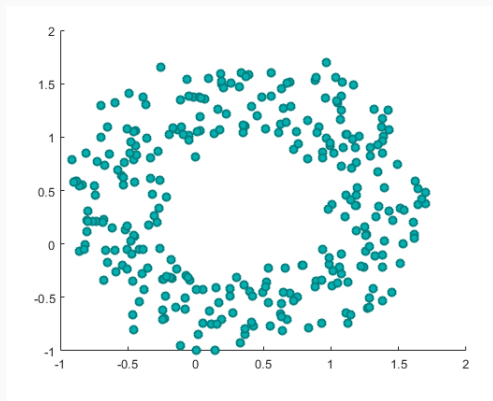
CONCLUSION

What do you see?
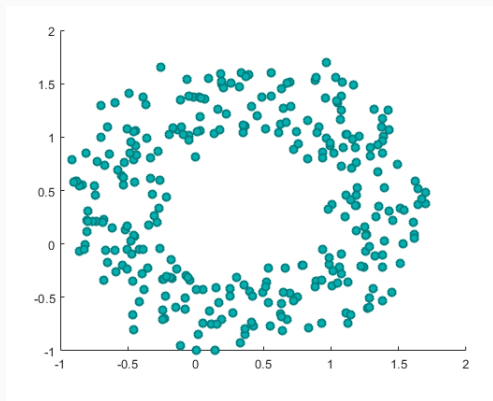
What do you see?

· **Irregular verbs:** discrete points

What do you see?

- **Irregular verbs:** discrete points
- **Pronouns**: conceptual circle → messy data

What do you see?

- **Irregular verbs:** discrete points
- **Pronouns**: conceptual circle → messy data
- **Sentence selection**: messy data → conceptual circle

Problems are often multi-faceted:

- · Verb inflection does have regular rules
- · Antecedent saliency for pronominal reference may have multiple factors
- · Sentence selection features require principled analysis

Things we've learned creating advanced NLG:

- · Strategies for tackling problems should not (always) be monolithic
- · Utilize whatever tools you have

Things we've learned creating advanced NLG:

- · Strategies for tackling problems should not (always) be monolithic
- · Utilize whatever tools you have
- · but make sure those strategies are contingent on thoroughly assessing the nature of the problems

Things we've learned creating advanced NLG:

- · Strategies for tackling problems should not (always) be monolithic
- · Utilize whatever tools you have
- · but make sure those strategies are contingent on thoroughly assessing the nature of the problems
- · which often requires having domain knowledge

Things we've learned creating advanced NLG:

- · Strategies for tackling problems should not (always) be monolithic
- · Utilize whatever tools you have
- · but make sure those strategies are contingent on thoroughly assessing the nature of the problems
- · which often requires having domain knowledge
    - · go learn about what others have done in your field

Things we've learned creating advanced NLG:

- · Strategies for tackling problems should not (always) be monolithic
- · Utilize whatever tools you have
- · but make sure those strategies are contingent on thoroughly assessing the nature of the problems
- · which often requires having domain knowledge
    - · go learn about what others have done in your field
    - · from various perspectives: e.g. linguistics, comp sci, journalism,...

THANK YOU!
QUESTIONS?

Get the source of this theme and the demo presentation from

github.com/matze/mtheme

The theme itself is licensed under a Creative Commons
Attribution-ShareAlike 4.0 International License.