

Wie kam es zu Bullshit?



Special Questions and techniques

Christian Gätcke
Institut für Informatik
FU Berlin
18.01.2013

- 1. Motivation
 - Was sind Special Questions überhaupt?
- 2. Fragetypen und ihre Auswertung
 - Fill in the Blank
 - Multiple Choice
 - Puzzle
- 3. Verbesserungen
- 4. Maschinelles Lernen

- 1. Motivation
 - Was sind Special Questions überhaupt?
- 2. Fragetypen und ihre Auswertung
 - Fill in the Blank
 - Multiple Choice
 - Puzzle
- 3. Verbesserungen
- 4. Maschinelles Lernen

- Was sind Special Questions?

- Was sind Special Questions?
 - Fragen nach Metaphern
 - Wortspielereien
 - Um die Ecke denken

- Was sind Special Questions?
 - Fragen nach Metaphern
 - Wortspielereien
 - Um die Ecke denken
- Allgemein: Fragen, welche eine gesonderte Pipeline zur Erkennung und Auswertung benötigen!

- Was sind Special Questions?
 - Fragen nach Metaphern
 - Wortspielereien
 - Um die Ecke denken
- Allgemein: Fragen, welche eine gesonderte Pipeline zur Erkennung und Auswertung benötigen!
- Warum betrachten wir das überhaupt?

- Was sind Special Questions?
 - Fragen nach Metaphern
 - Wortspielereien
 - Um die Ecke denken
- Allgemein: Fragen, welche eine gesonderte Pipeline zur Erkennung und Auswertung benötigen!
- Warum betrachten wir das überhaupt?
 - Trotz geringer Prozentsätze z.T. 1/6 einer Runde
 - Interessant für die zukünftige Verwendung von Watson

- 1. Motivation
 - Was sind Special Questions überhaupt?
- 2. Fragetypen und ihre Auswertung
 - Fill in the Blank
 - Multiple Choice
 - Puzzle
- 3. Verbesserungen
- 4. Maschinelles Lernen

- Unterteilung in am häufigsten aufgetretene Kategorien

- Unterteilung in am häufigsten aufgetretene Kategorien
 - welche aber teilweise Präfix haben

- Unterteilung in am häufigsten aufgetretene Kategorien
 - welche aber teilweise Präfix haben
- Fragetypen
 - Fill in the Blank
 - Multiple Choice
 - Puzzle

- Unterteilung in am häufigsten aufgetretene Kategorien
 - welche aber teilweise Präfix haben
- Fragetypen
 - Fill in the Blank
 - Multiple Choice
 - Puzzle
- Antworttypen

- Unterteilung in am häufigsten aufgetretene Kategorien
 - welche aber teilweise Präfix haben
- Fragetypen
 - Fill in the Blank
 - Multiple Choice
 - Puzzle
- Antworttypen
 - direkte Antworten auf die Frage

- Unterteilung in am häufigsten aufgetretene Kategorien
 - welche aber teilweise Präfix haben
- Fragetypen
 - Fill in the Blank
 - Multiple Choice
 - Puzzle
- Antworttypen
 - direkte Antworten auf die Frage
 - aber auch Neologismen/Wortspielereien
 - Anagramme
 - Zahlen

Fragetypen-Fill in the Blank

- Fill in the Blank

- Fill in the Blank
 - z.B.:Kategorie „Only one Vowel“ Frage:Proverbially, you can be „flying“ this or be this „ and dry“!

- Fill in the Blank
 - z.B.:Kategorie „Only one Vowel“ Frage:Proverbially, you can be „flying“ this or be this „ and dry“!
 - Antwort:“High“

- Fill in the Blank
 - z.B.:Kategorie „Only one Vowel“ Frage:Proverbially, you can be „flying“ this or be this „ and dry“!
 - Antwort:“High“
- Auseinandernehmen der Frage

- Fill in the Blank
 - z.B.:Kategorie „Only one Vowel“ Frage:Proverbially, you can be „flying“ this or be this „ and dry“!
 - Antwort:“High“
- Auseinandernehmen der Frage
 - eigentlich 3 Kategorien

- Fill in the Blank
 - z.B.:Kategorie „Only one Vowel“ Frage:Proverbially, you can be „flying“ this or be this „ and dry“!
 - Antwort:“High“
- Auseinandernehmen der Frage
 - eigentlich 3 Kategorien
 - wir betrachten aber nur FITB

- Fill in the Blank
 - z.B.:Kategorie „Only one Vowel“ Frage:Proverbially, you can be „flying“ this or be this „ and dry“!
 - Antwort:“High“
- Auseinandernehmen der Frage
 - eigentlich 3 Kategorien
 - wir betrachten aber nur FITB
- Zerlegung
 - „flying _____“(this)
 - (this)“_____ and dry“

- Fill in the Blank
 - Bsp 2: Kategorie:Double,Double Frage:“The classic name for the double promotories at the Strait of Gibraltar is „the Pillars of“ this mythic hero.

- Fill in the Blank
 - Bsp 2: Kategorie:Double,Double Frage:“The classic name for the double promotories at the Strait of Gibraltar is „the Pillars of“ this mythic hero.
 - Antwort: Herkules

- Fill in the Blank
 - Bsp 2: Kategorie:Double,Double Frage:“The classic name for the double promotories at the Strait of Gibraltar is „the Pillars of“ this mythic hero.
 - Antwort: Herkules
- Antwortaufbau

- Fill in the Blank
 - Bsp 2: Kategorie:Double,Double Frage:“The classic name for the double promotories at the Strait of Gibraltar is „the Pillars of“ this mythic hero.
 - Antwort: Herkules
- Antwortaufbau
 - „this“ weist wieder auf FITB hin

- Fill in the Blank
 - Bsp 2: Kategorie:Double,Double Frage:"The classic name for the double promotories at the Strait of Gibraltar is „the Pillars of“ this mythic hero.
 - Antwort: Herkules
- Antwortaufbau
 - „this“ weist wieder auf FITB hin
 - Aufteilung in mythic hero und Strait of Gibraltar

- Fill in the Blank
 - Bsp 2: Kategorie:Double,Double Frage:"The classic name for the double promotories at the Strait of Gibraltar is „the Pillars of“ this mythic hero.
 - Antwort: Herkules
- Antwortaufbau
 - „this“ weist wieder auf FITB hin
 - Aufteilung in mythic hero und Strait of Gibraltar
 - FITB-Kandidaten werden generiert

- Multiple Choice
 - Sucht aus einer vorgegebenen Lösungsmenge das richtige raus

- Multiple Choice
 - Sucht aus einer vorgegebenen Lösungsmenge das richtige raus
 - Bsp: Kategorie: Branson, Hanson or Manson: This singer combines murderer and „Girls best friend“

- Multiple Choice
 - Sucht aus einer vorgegebenen Lösungsmenge das richtige raus
 - Bsp: Kategorie: Branson, Hanson or Manson: This singer combines murderer and „Girls best friend“
 - Antwort: Manson

- Multiple Choice
 - Sucht aus einer vorgegebenen Lösungsmenge das richtige raus
 - Bsp: Kategorie: Branson, Hanson or Manson: This singer combines murderer and „Girls best friend“
 - Antwort: Manson
- Und wie?

- Multiple Choice
 - Sucht aus einer vorgegebenen Lösungsmenge das richtige raus
 - Bsp: Kategorie: Branson, Hanson or Manson: This singer combines murderer and „Girls best friend“
 - Antwort: Manson
- Und wie?
 - Frage wird als MC klassifiziert

- Multiple Choice
 - Sucht aus einer vorgegebenen Lösungsmenge das richtige raus
 - Bsp: Kategorie: Branson, Hanson or Manson: This singer combines murderer and „Girls best friend“
 - Antwort: Manson
- Und wie?
 - Frage wird als MC klassifiziert
 - mögliche Antworten werden Identifiziert
 - weitergabe an MC-Antwort-Generator

- Puzzle

- Puzzle
 - Hier kommen die Präfixe besonders zum Tragen

- Puzzle
 - Hier kommen die Präfixe besonders zum Tragen
 - Kann Watson auch lustig sein?

- Puzzle
 - Hier kommen die Präfixe besonders zum Tragen
 - Kann Watson auch lustig sein?
 - Kann Watson um die Ecke denken?

- Lexikalische Analyse

- Lexikalische Analyse
 - Eigene Constraint-Klassen

- Lexikalische Analyse
 - Eigene Constraint-Klassen
 - Verschiedene Subkategorien

- Lexikalische Analyse
 - Eigene Constraint-Klassen
 - Verschiedene Subkategorien
 - Kombinierbar

- Lexikalische Analyse
 - Eigene Constraint-Klassen
 - Verschiedene Subkategorien
 - Kombinierbar
- Beispiele für constraints

<i>Constraint class</i>	<i>Sample category</i>	<i>Sample answer</i>
Alliteration	ALLITERATIVE ARTISTS	Pablo Picasso
Blank	ON ____	On Deck
DoubleLetter	DOUBLE "H"	Fishhook
CommonEnding	-OID	Celluloid
NLetterRepetitions	"U"2	Futures
Rhyming	RHYMES WITH JOCK	Frock
SubStringDisjunction	"CHURCH" AND "STATE"	Churchkey
SyllableCount	MONOSYLLABLES	Ain't

- Anwendungsbeispiel:

- Anwendungsbeispiel:
 - Kategorie: „Four-Letter-Words“ Frage: „Fornication under consent of the King“

- Anwendungsbeispiel:
 - Kategorie: „„F“our-Letter-Words“ Frage: „Fornication under consent of the King“
 - Antwort: „Fu*k“

- Anwendungsbeispiel:
 - Kategorie: „F“our-Letter-Words“ Frage: “Fornication under consent of the King“
 - Antwort: „Fu*k“
- Nicht immer eindeutig
 - teilweise über mehrere Wörter verstreut
 - oder aber phonetisch
 - Bsp: The l“one“liest in“to“(lies: two[2])

- Anwendungsbeispiel:
 - Kategorie: „F“our-Letter-Words“ Frage: “Fornication under consent of the King“
 - Antwort: „Fu*k“
- Nicht immer eindeutig
 - teilweise über mehrere Wörter verstreut
 - oder aber phonetisch
 - Bsp: The l“one“liest in“to“(lies: two[2])
 - Lösung: Maschinelles Lernen aus aufgedeckten Lösungen

Wo ist ihr Humor, Watson?-Sherlock Holmes

- Aufgrund der Menge an Wortspielen beschränkt

- Aufgrund der Menge an Wortspielen beschränkt
 - explizit auf Worte in Anführungszeichen

- Aufgrund der Menge an Wortspielen beschränkt
 - explizit auf Worte in Anführungszeichen
 - „direkte“ indirekte Hinweise
- Substring, Instance, Synonym und Relation

- Aufgrund der Menge an Wortspielen beschränkt
 - explizit auf Worte in Anführungszeichen
 - „direkte“ indirekte Hinweise
- Substring, Instance, Synonym und Relation
 - Substring: Frage enthält einen Teil der Lösung
 - Bsp: In this game, the „devil“ might shed some tears!

- Aufgrund der Menge an Wortspielen beschränkt
 - explizit auf Worte in Anführungszeichen
 - „direkte“ indirekte Hinweise
- Substring, Instance, Synonym und Relation
 - Substring: Frage enthält einen Teil der Lösung
 - Bsp: In this game, the „devil“ might shed some tears!
 - Antwort: “Devil” may Cry

- Aufgrund der Menge an Wortspielen beschränkt
 - explizit auf Worte in Anführungszeichen
 - „direkte“ indirekte Hinweise
- Substring, Instance, Synonym und Relation
 - Substring: Frage enthält einen Teil der Lösung
 - Bsp: In this game, the „devil“ might shed some tears!
 - Antwort: “Devil” may Cry
 - Synonym: Frage enthält ein Synonym für die gesuchte Lösung
 - Bsp: Harry Potter fights this person, „who shall not be named“

- Aufgrund der Menge an Wortspielen beschränkt
 - explizit auf Worte in Anführungszeichen
 - „direkte“ indirekte Hinweise
- Substring, Instance, Synonym und Relation
 - Substring: Frage enthält einen Teil der Lösung
 - Bsp: In this game, the „devil“ might shed some tears!
 - Antwort: “Devil” may Cry
 - Synonym: Frage enthält ein Synonym für die gesuchte Lösung
 - Bsp: Harry Potter fights this person, „who shall not be named“
 - Antwort: “Voldem....”(psssssssst)

- Instance: Ein Teil der Frage enthält die übergeordnete Kategorie der Lösung
 - Bsp: This >>spread<< of hazelnut, nougat and milk is also known as „the brown gold“!

- Instance: Ein Teil der Frage enthält die übergeordnete Kategorie der Lösung
 - Bsp: This >>spread<< of hazelnut, nougat and milk is also known as „the brown gold“!
 - Antwort: Nutella (Übergeordnete Kategorie: Spread-Brottaufstrich)

- Instance: Ein Teil der Frage enthält die übergeordnete Kategorie der Lösung
 - Bsp: This >>spread<< of hazelnut, nougat and milk is also known as „the brown gold“!
 - Antwort: Nutella (Übergeordnete Kategorie: Spread-Brottaufstrich)
- Relation: Aus der Frage und den Hinweisen ergibt sich eine (in)direkte Verbindung
 - Bsp: Even if they wanted to, Lucy Koh, Kwon Oh Hyun and this „Turtleneck“ won't show up at a party together!

- Instance: Ein Teil der Frage enthält die übergeordnete Kategorie der Lösung
 - Bsp: This >>spread<< of hazelnut, nougat and milk is also known as „the brown gold“!
 - Antwort: Nutella (Übergeordnete Kategorie: Spread-Brottaufstrich)
- Relation: Aus der Frage und den Hinweisen ergibt sich eine (in)direkte Verbindung
 - Bsp: Even if they wanted to, Lucy Koh, Kwon Oh Hyun and this „Turtleneck“ won't show up at a party together!
 - Antwort: Steve Jobs

- Instance: Ein Teil der Frage enthält die übergeordnete Kategorie der Lösung
 - Bsp: This >>spread<< of hazelnut, nougat and milk is also known as „the brown gold“!
 - Antwort: Nutella (Übergeordnete Kategorie: Spread-Brottaufstrich)
- Relation: Aus der Frage und den Hinweisen ergibt sich eine (in)direkte Verbindung
 - Bsp: Even if they wanted to, Lucy Koh, Kwon Oh Hyun and this „Turtleneck“ won't show up at a party together!
 - Antwort: Steve Jobs
 - Verbindung: Kwon Oh Hyun(Samsung CEO)<-> Lucy Koh(Richterin im Patentprozess Samsung gegen Apple)<->Apple<->Turtleneck<->Steve Jobs

- Lassen sich in 3 Kategorien zusammenfassen

- Lassen sich in 3 Kategorien zusammenfassen
 - Before & After und Rhyme-Time
 - Math
 - Anagrams und Hidden Words

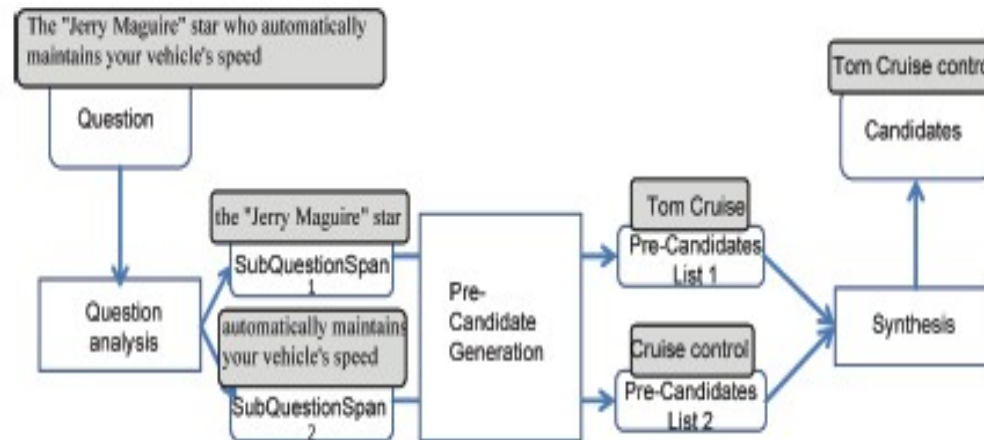
- Before and After

- Before and After
 - letzte n Worte der ersten Teillösung == erste n Worte der zweiten Teillösung
 - Bsp: The „Jerry Maguire“ star who automatically maintains your vehicles speed.

- Before and After
 - letzte n Worte der ersten Teillösung == erste n Worte der zweiten Teillösung
 - Bsp: The „Jerry Maguire“ star who automatically maintains your vehicles speed.
 - Antwort: “Tom Cruise control”

- Before and After
 - letzte n Worte der ersten Teillösung == erste n Worte der zweiten Teillösung
 - Bsp: The „Jerry Maguire“ star who automatically maintains your vehicles speed.
 - Antwort: „Tom Cruise control“

- Auswertung:



- Rhyme-Time

- Rhyme-Time
 - Hier können Neologismen entstehen

- Rhyme-Time
 - Hier können Neologismen entstehen
 - Bsp: Kategorie Authors Rhyme Time, Frage: Ian's scandinavian rodents!

- Rhyme-Time
 - Hier können Neologismen entstehen
 - Bsp: Kategorie Authors Rhyme Time, Frage: Ian's scandinavian rodents!
 - Antwort: Flemming's lemmings
- Auswertung:
 - Einteilung in SubQuestions
 - Ranking nach phonetischem Score
 - Carnegie Mellon Pronouncing Dictionary
 - sonst wie bei B&A

- Math

- Math
 - 3 Hauptkategorien
 - Rechnen mit Zahlen(Math Class)
 - Geometrie
 - Textaufgaben(You do the math)

- Math
 - 3 Hauptkategorien
 - Rechnen mit Zahlen(Math Class)
 - Geometrie
 - Textaufgaben(You do the math)
- Auswertung:
 - Math-Class:Definite Clause Grammar
 - Bsp: This number equals one twelfth of twentyfour!

- Math
 - 3 Hauptkategorien
 - Rechnen mit Zahlen(Math Class)
 - Geometrie
 - Textaufgaben(You do the math)
- Auswertung:
 - Math-Class:Definite Clause Grammar
 - Bsp: This number equals one twelfth of twentyfour!
 - Geometrie: Trigger-Wörter
 - Bsp:This is the >>diameter<< of a >>circle<< times pi!

- Math
 - 3 Hauptkategorien
 - Rechnen mit Zahlen(Math Class)
 - Geometrie
 - Textaufgaben(You do the math)
- Auswertung:
 - Math-Class:Definite Clause Grammar
 - Bsp: This number equals one twelfth of twentyfour!
 - Geometrie: Trigger-Wörter
 - Bsp:This is the >>diameter<< of a >>circle<< times pi!
 - Textaufgaben: Trigger durch „number of“
 - Bsp:The >>number of<< tires on a road-train multiplied with the >>number of<< of engines on a Boeing 787!

- Anagrams und Hidden Words

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt, aber nicht zwangsläufig

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt, aber nicht zwangsläufig
 - Maschinelles lernen

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt, aber nicht zwangsläufig
 - Maschinelles lernen
 - Bsp: Kategorie Anagrams, Frage: It's the only kitchen appliance Rosetta knows how to use.

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt, aber nicht zwangsläufig
 - Maschinelles lernen
 - Bsp: Kategorie Anagrams, Frage: It's the only kitchen appliance Rosetta knows how to use.
 - Antwort: "Toaster" (Rosetta-Anagram)

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt, aber nicht zwangsläufig
 - Maschinelles lernen
 - Bsp: Kategorie Anagrams, Frage: It's the only kitchen appliance Rosetta knows how to use.
 - Antwort: "Toaster" (Rosetta-Anagram)
- Auswertung:
 - Einteilung in Kategorien(kitchen appliance-Küchengeräte)

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt, aber nicht zwangsläufig
 - Maschinelles lernen
 - Bsp: Kategorie Anagrams, Frage: It's the only kitchen appliance Rosetta knows how to use.
 - Antwort: "Toaster" (Rosetta-Anagram)
- Auswertung:
 - Einteilung in Kategorien(kitchen appliance-Küchengeräte)
 - Anagramme oder versteckte Wörter generieren

- Anagrams und Hidden Words
 - Werden beide meistens in der Fragestellung genannt, aber nicht zwangsläufig
 - Maschinelles lernen
 - Bsp: Kategorie Anagrams, Frage: It's the only kitchen appliance Rosetta knows how to use.
 - Antwort: "Toaster" (Rosetta-Anagram)
- Auswertung:
 - Einteilung in Kategorien(kitchen appliance-Küchengeräte)
 - Anagramme oder versteckte Wörter generieren
 - Gegenprüfung gegen die Kategorien und Ranking

- 1. Motivation
 - Was sind Special Questions überhaupt?
- 2. Fragetypen und ihre Auswertung
 - Fill in the Blank
 - Multiple Choice
 - Puzzle
- 3. Verbesserungen

- zuerst einmal 2 Statistiken
 - Auf 3571 Fragen:

	<i>Puzzle</i>	<i>FITB</i>	<i>Multiple-Choice</i>
<i>Number detected</i>	43	142	20
<i>Number in gold standard</i>	80	135	19
<i>Number correct</i>	42	96	13
<i>Recognition precision</i>	0.977	0.676	0.650
<i>Recognition recall</i>	0.525	0.711	0.684
<i>F₁</i>	0.683	0.693	0.667
<i>Frequency in gold standard</i>	0.023	0.038	0.005
<i>Answering accuracy</i>	0.512	0.803	0.6
<i>Answering Precision@70</i>	0.645	0.970	0.571

- Problemfälle: Before and After und Anagramme

<i>Category of failure</i>	<i>Detail</i>	<i>Percentage of questions</i>
Question analysis	Parse problem	4%
	Incorrect decomposition	11%
Search	No recall in one search	19%
	No recall in both searches	5%
Algorithm deficiency	Question does not cleanly break into left and right parts	3%
	Pre-candidates do not overlap on word boundaries	2%
Scoring	Correct answer not in first place	5%
Other		5%

- Problemfälle: Before and After und Anagramme
 - Fragestellungen wurden teilweise nicht erkannt

- Problemfälle: Before and After und Anagramme
 - Fragestellungen wurden teilweise nicht erkannt
 - Die Lösung des Problems: Lernen durch aufgedeckte Lösungen

- Problemfälle: Before and After und Anagramme
 - Fragestellungen wurden teilweise nicht erkannt
 - Die Lösung des Problems: Lernen durch aufgedeckte Lösungen
- Bsp: Kategorie: Whirled Capitals, Frage: Non Old.
 - Antwort: London

- Problemfälle: Before and After und Anagramme
 - Fragestellungen wurden teilweise nicht erkannt
 - Die Lösung des Problems: Lernen durch aufgedeckte Lösungen
- Bsp: Kategorie: Whirled Capitals, Frage: Non Old.
 - Antwort: London
 - Watson konnte den LAT nicht erkennen

- Problemfälle: Before and After und Anagramme
 - Fragestellungen wurden teilweise nicht erkannt
 - Die Lösung des Problems: Lernen durch aufgedeckte Lösungen
- Bsp: Kategorie: Whirled Capitals, Frage: Non Old.
 - Antwort: London
 - Watson konnte den LAT nicht erkennen
 - Lösungen aus der Runde wurden nacheinander aufgedeckt

- Problemfälle: Before and After und Anagramme
 - Fragestellungen wurden teilweise nicht erkannt
 - Die Lösung des Problems: Lernen durch aufgedeckte Lösungen
- Bsp: Kategorie: Whirled Capitals, Frage: Non Old.
 - Antwort: London
 - Watson konnte den LAT nicht erkennen
 - Lösungen aus der Runde wurden nacheinander aufgedeckt
 - Watson konnte sich LAT zusammenbasteln
 - Resultat:

<i>Category</i>	<i>Question</i>	<i>Answer</i>	<i>Probability</i>	<i>Triggered</i>
I-I-I-I	It's the inflammation of the gums	Gingivitis	.0062	No
	On Dec. 10, 1817 it entered the Union as the 20th state	Mississippi	.0089	No
	In this early '90s video game by Sid Meier, you built "an empire to stand the test of time"	Civilization	.5917	No
	It's the process of converting data into a form that can be used by computers	Digitization	.9957	Yes
	Voters who "take" this can propose their own constitutional amendment	Initiative	.9999	Yes
WHIRLED CAPITALS	Non old	London	.00017	No
	In vane	Vienna	.9453	No
	I rasp	Paris	.9999	Yes
	Ape rug	Prague	.9999	Yes
	In silk, eh?	Helsinki	.9999	Yes

- Watson löst die Teilfragen nicht mehr parallel, sondern synthetisiert eigene Fragen.
- Watson lernt Humor.
- Watson lernt eventuell, Reime auch zu vernuscheln.
- Watson lernt „um die Ecke denken“!

„I'm afraid i can't let you do that,
Dave“-HAL

Vielen Dank!