Topic Modeling

Topic Modeling

Document 1

Document 2

Document 3

⋮

Document N

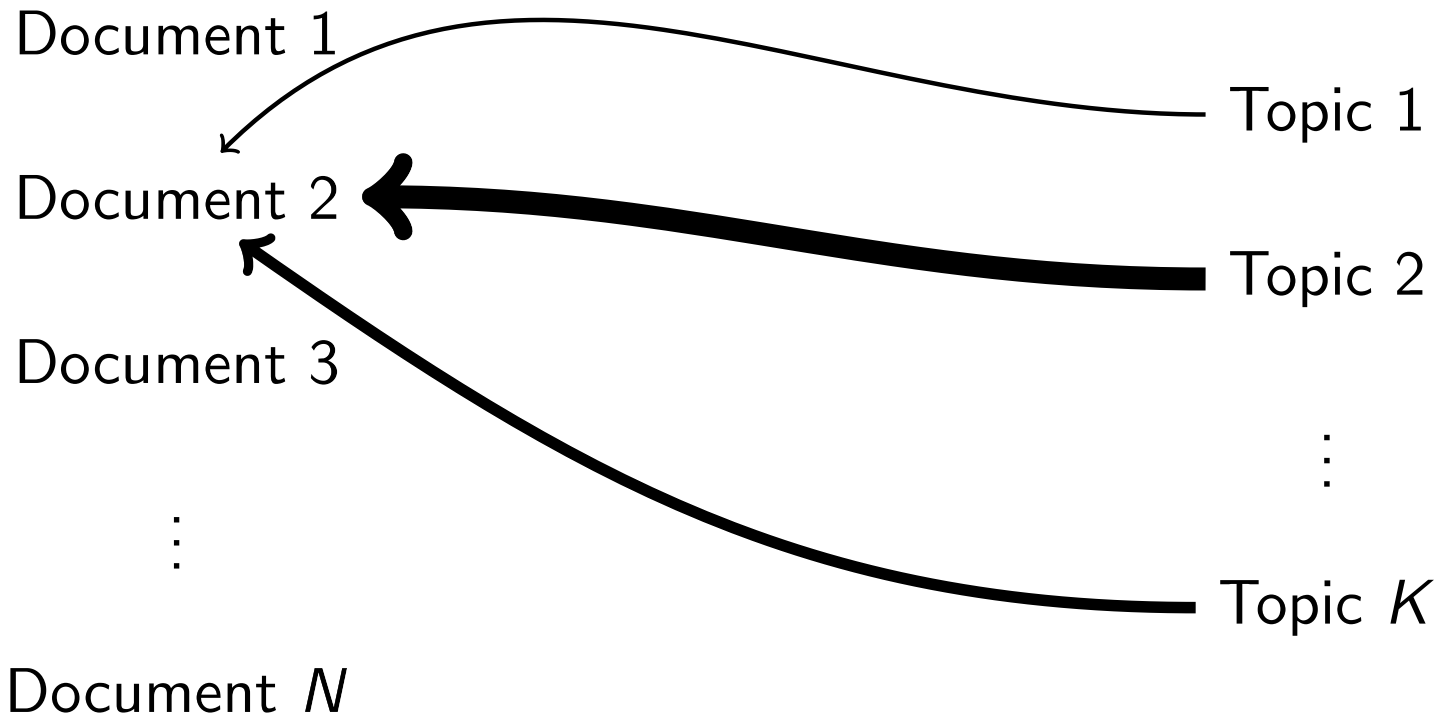
Topic 1

Topic 2

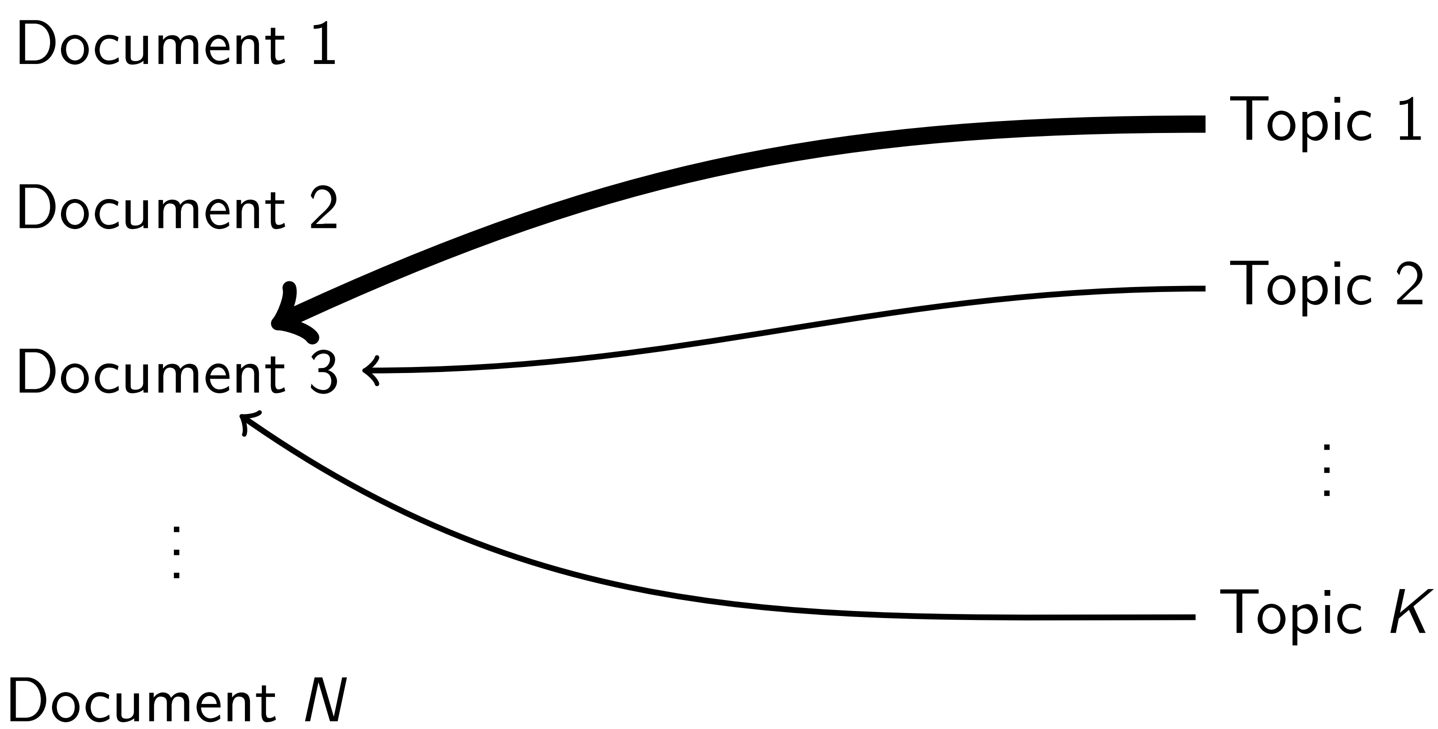
⋮

Topic K

Topic Modeling



Topic Modeling



DGP: intuition

DGP: intuition

Documents exhibit different topics,

DGP: intuition

Documents exhibit different topics, and in different proportions.

DGP: intuition

Documents exhibit different topics, and in different proportions.

E.g. a speech by the Finance minister might be 50% drawn from the `trade` topic, 40% from the `spending` topic, 9.9% from the `taxation` topic, 0.1% from the `health` topic.

DGP: intuition

Documents exhibit different topics, and in different proportions.

E.g. a speech by the Finance minister might be 50% drawn from the `trade` topic, 40% from the `spending` topic, 9.9% from the `taxation` topic, 0.1% from the `health` topic.

Think of a topic as a distribution over a fixed vocabulary.

DGP: intuition

Documents exhibit different topics, and in different proportions.

E.g. a speech by the Finance minister might be 50% drawn from the `trade` topic, 40% from the `spending` topic, 9.9% from the `taxation` topic, 0.1% from the `health` topic.

Think of a topic as a distribution over a fixed vocabulary.

E.g. the trade topic will have words like `import` and `tariff` with high probability.

DGP: intuition

Documents exhibit different topics, and in different proportions.

E.g. a speech by the Finance minister might be 50% drawn from the `trade` topic, 40% from the `spending` topic, 9.9% from the `taxation` topic, 0.1% from the `health` topic.

Think of a topic as a distribution over a fixed vocabulary.

E.g. the trade topic will have words like `import` and `tariff` with high probability.

Technically we assume the topics are generated first,

DGP: intuition

Documents exhibit different topics, and in different proportions.

E.g. a speech by the Finance minister might be 50% drawn from the `trade` topic, 40% from the `spending` topic, 9.9% from the `taxation` topic, 0.1% from the `health` topic.

Think of a topic as a distribution over a fixed vocabulary.

E.g. the trade topic will have words like `import` and `tariff` with high probability.

Technically we assume the topics are generated first, and the documents are generated second (from those topics).

DGP: intuition

Documents exhibit different topics, and in different proportions.

E.g. a speech by the Finance minister might be 50% drawn from the $\boxed{\texttt{trade}}$ topic, 40% from the $\boxed{\texttt{spending}}$ topic, 9.9% from the $\boxed{\texttt{taxation}}$ topic, 0.1% from the $\boxed{\texttt{health}}$ topic.

Think of a topic as a distribution over a fixed vocabulary.

E.g. the $\boxed{\text{trade}}$ topic will have words like `import` and `tariff` with high probability.

Technically we assume the topics are generated first, and the documents are generated second (from those topics).

Now, where do the words in the documents come from?

Intuition: Generating Words

Intuition: Generating Words

For each document. . .

Intuition: Generating Words

For each document...

1. Randomly choose a distribution over topics.

Intuition: Generating Words

For each document...

1. Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.

Intuition: Generating Words

For each document...

1. Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.

2. Then, for every word in the document...

Intuition: Generating Words

For each document...

1. Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.
2. Then, for every word in the document...
   1. Randomly choose a topic from the distribution over topics from step 1.

Intuition: Generating Words

For each document...

1. Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.

2. Then, for every word in the document...
   1. Randomly choose a topic from the distribution over topics from step 1.

   2. Randomly choose a word from the distribution over the vocabulary that the topic implies.

First Part

First Part

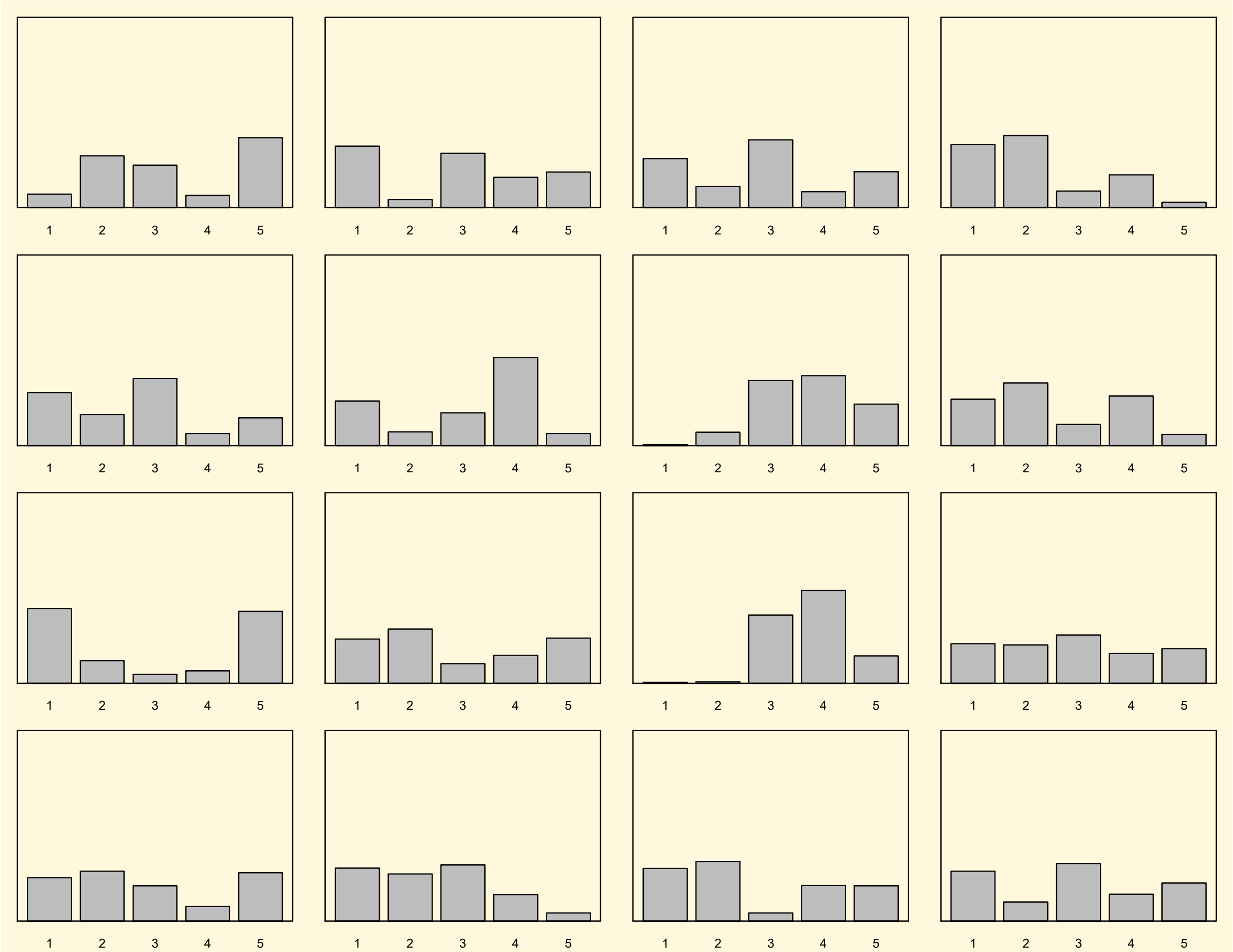
Randomly choose a distribution over topics.

First Part

Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.

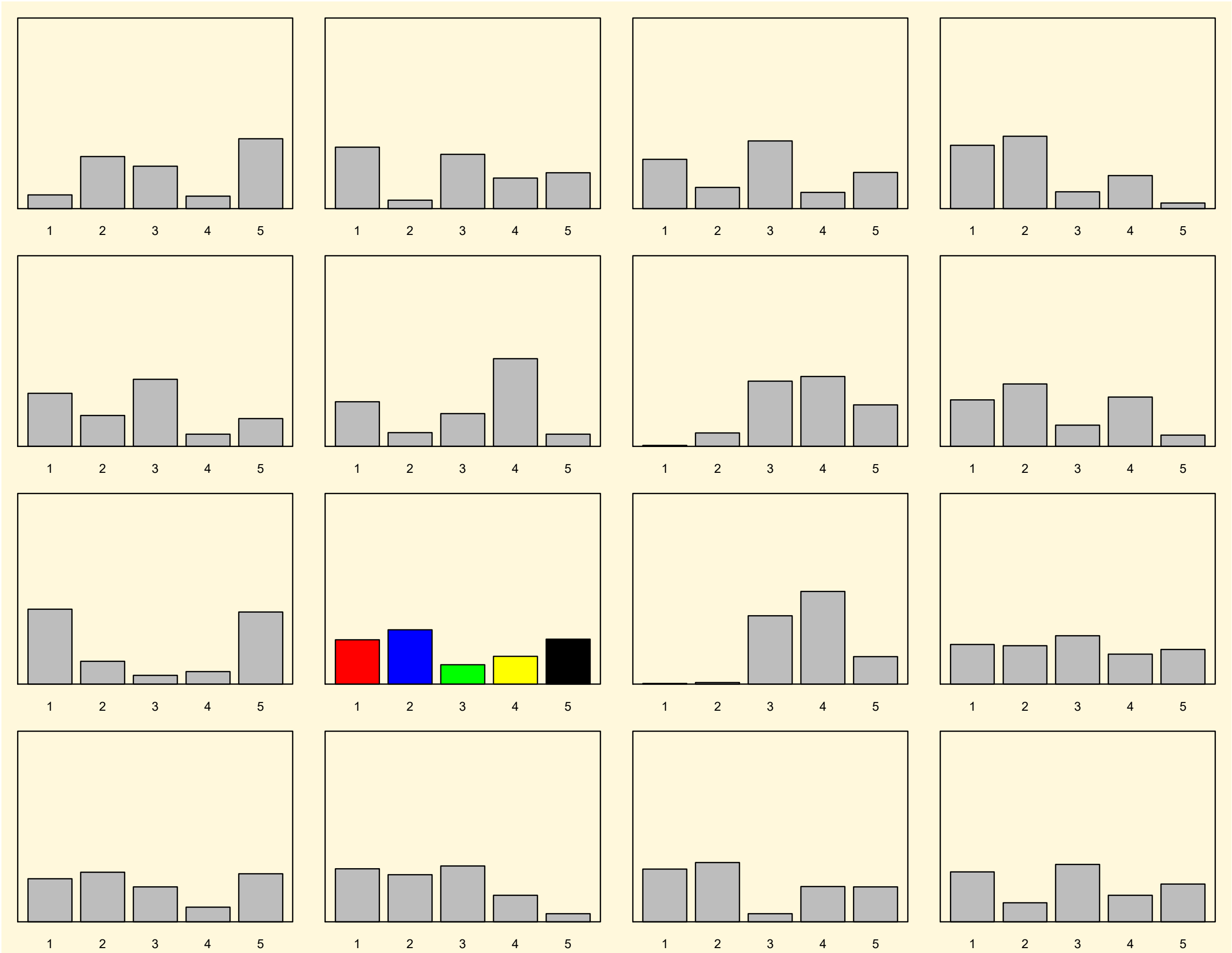
First Part

Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.



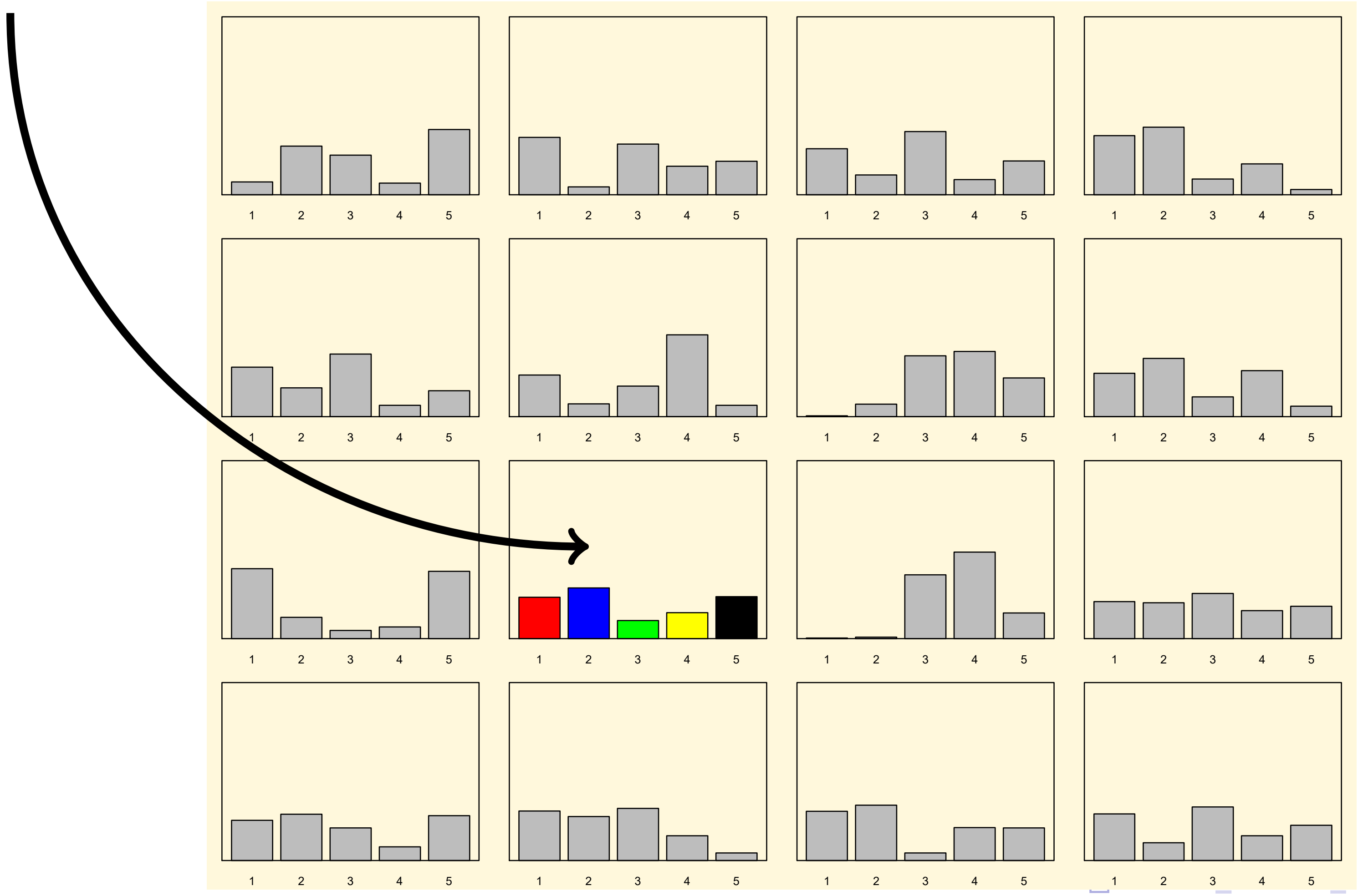
First Part

Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.



First Part

Randomly choose a distribution over topics. That is, choose one of many multinomial distributions, each which mixes the topics in different proportions.



Second Part

Second Part

Then, for every word in the document...

Second Part

Then, for every word in the document...

1. Randomly choose a topic from the distribution over topics from step 1.

Second Part

Then, for every word in the document. . .

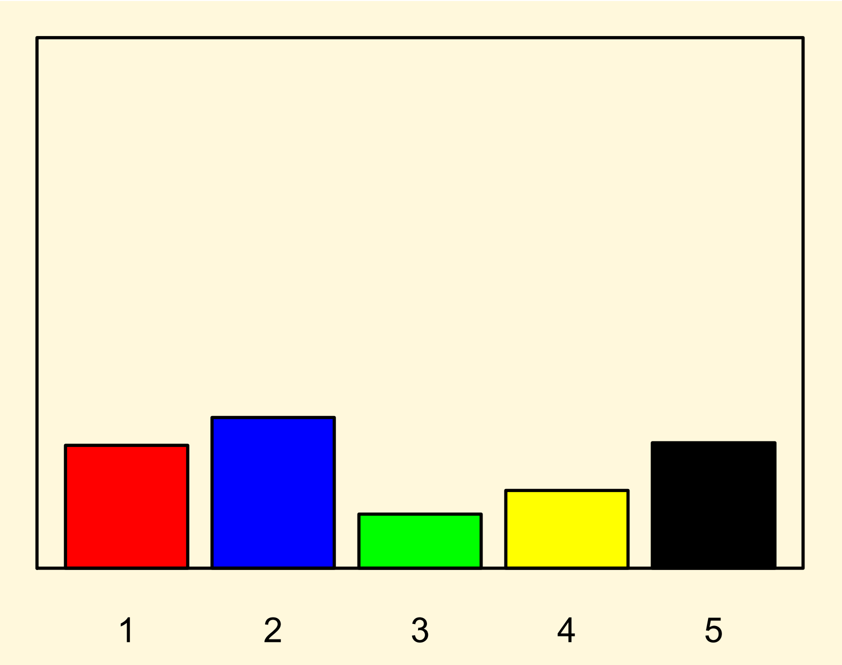
1. Randomly choose a topic from the distribution over topics from step 1.
2. Randomly choose a word from the distribution over the vocabulary that the topic implies.

Second Part

Then, for every word in the document...

1. Randomly choose a topic from the distribution over topics from step 1.

2. Randomly choose a word from the distribution over the vocabulary that the topic implies.

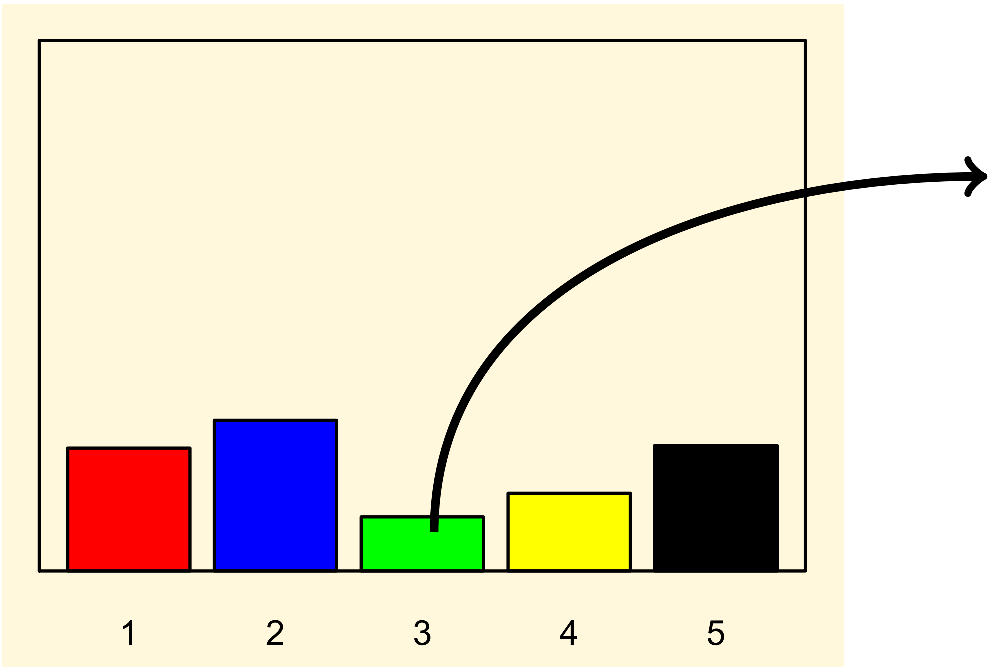


Second Part

Then, for every word in the document. . .

1. Randomly choose a topic from the distribution over topics from step 1.

2. Randomly choose a word from the distribution over the vocabulary that the topic implies.

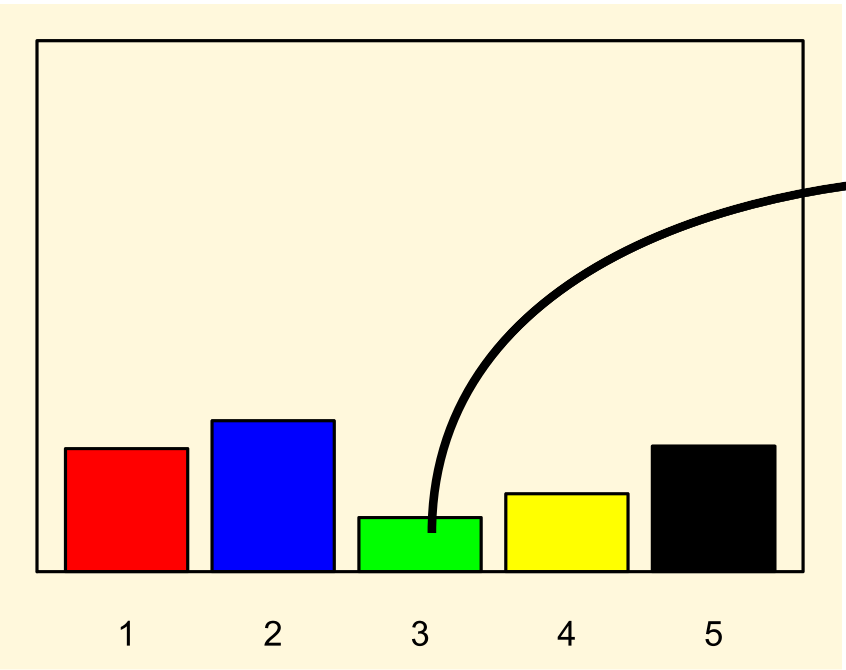


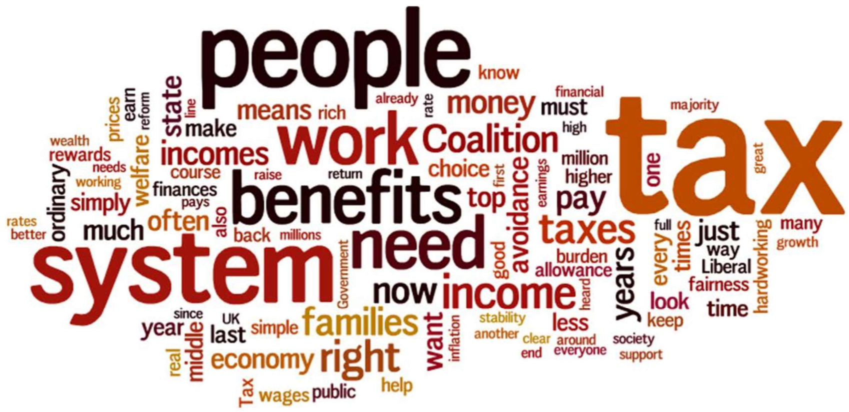
Second Part

Then, for every word in the document...

1. Randomly choose a topic from the distribution over topics from step 1.

2. Randomly choose a word from the distribution over the vocabulary that the topic implies.



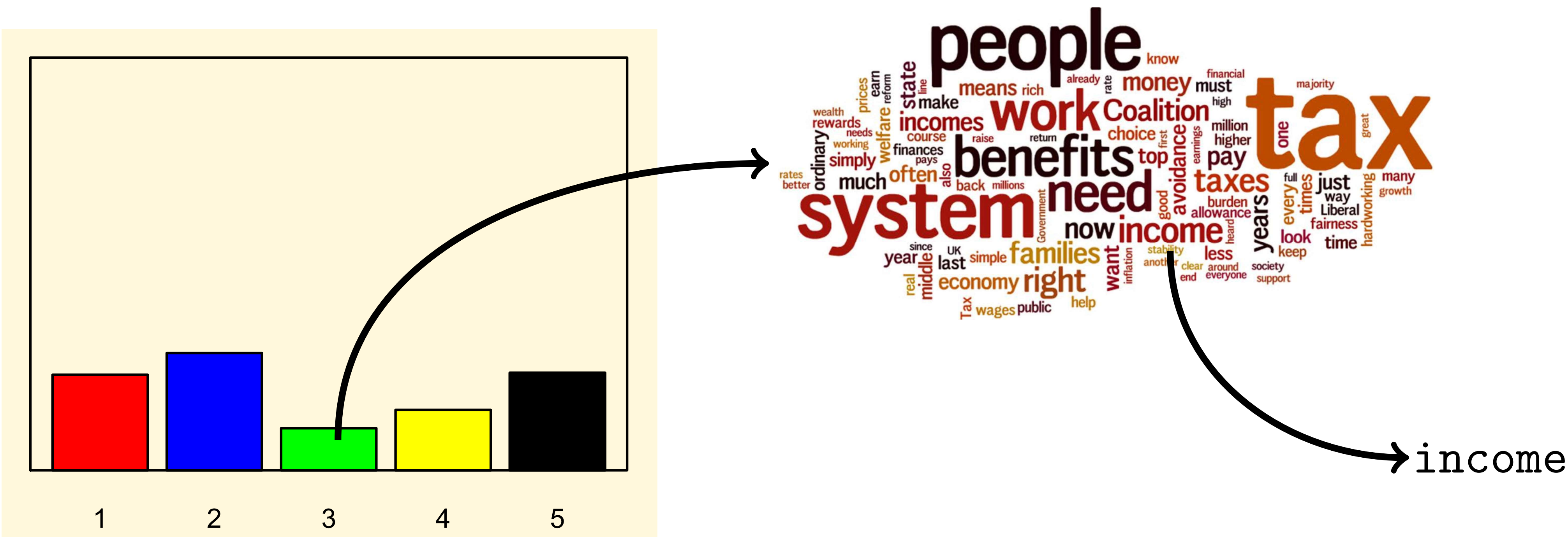


Second Part

Then, for every word in the document...

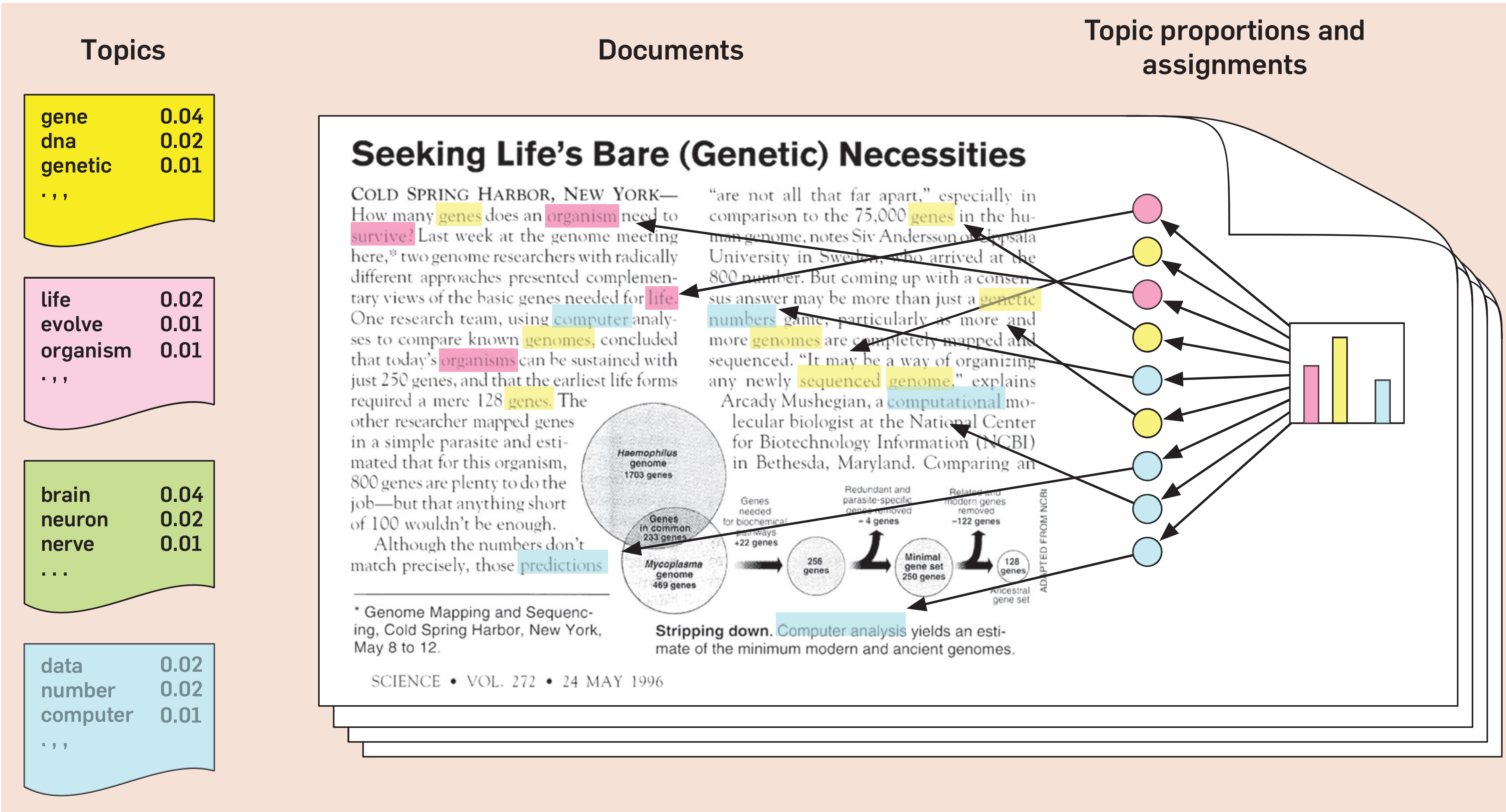
1. Randomly choose a topic from the distribution over topics from step 1.

2. Randomly choose a word from the distribution over the vocabulary that the topic implies.



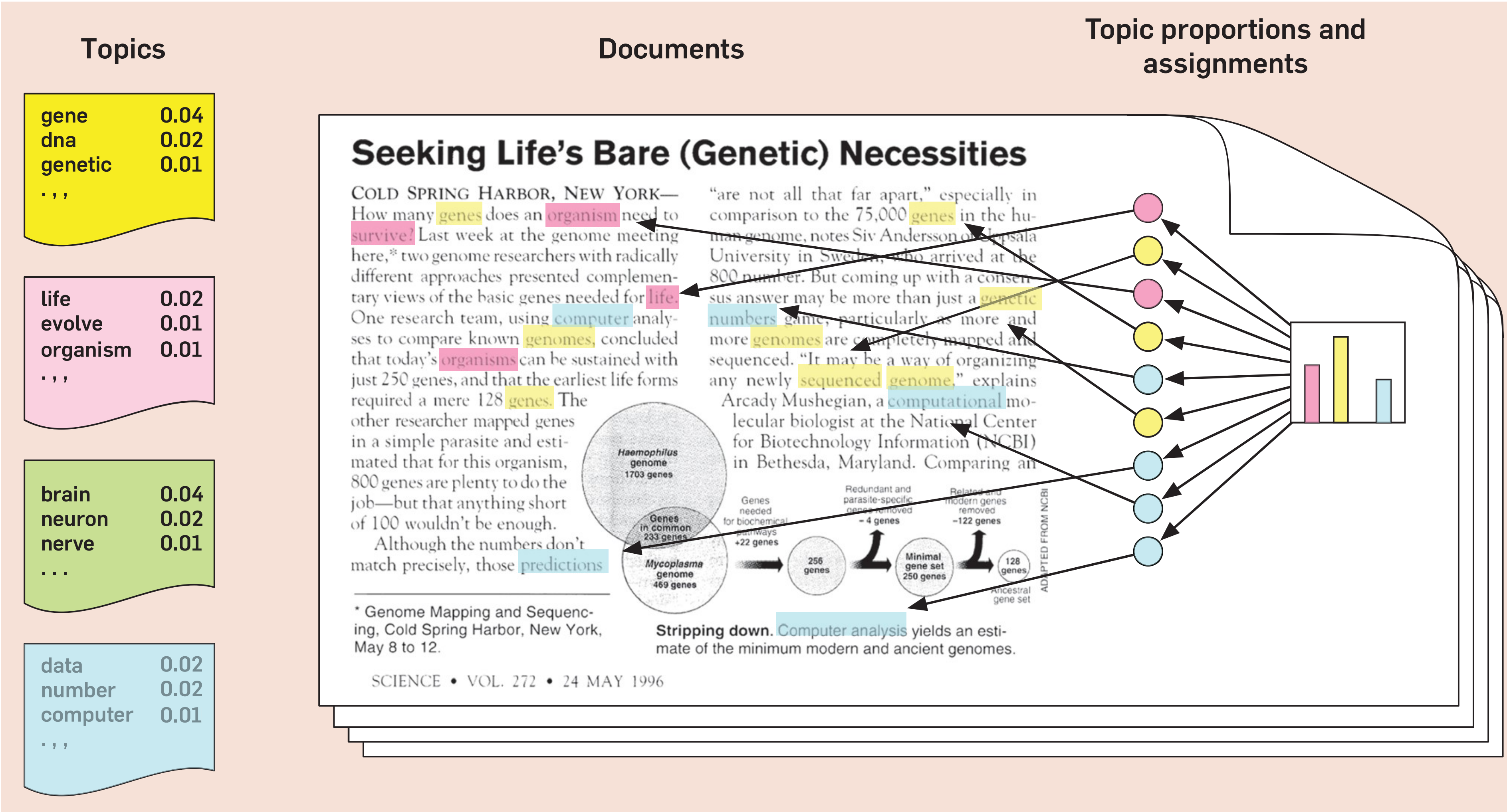
Topic Modeling a Document (Blei, 2012)

Topic Modeling a Document (Blei, 2012)



Note that all documents share same set of topics:

Topic Modeling a Document (Blei, 2012)



Note that all documents share same set of topics: but some (e.g. neuro) may be (basically) absent in a given document.

Notes

Notes

Some of our variables—the documents which contain the words—are observable.

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are $\boxed{\text{latent}}$.

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are $\boxed{\text{latent}}$.

We need a distribution from which to draw the per-document topic distribution. We use a $\boxed{\text{Dirichlet}}$ distribution as a prior for that.

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are latent.

We need a distribution from which to draw the per-document topic distribution. We use a Dirichlet distribution as a prior for that.

And Dirichlet is used for the allocation of the words in the documents to different topics:

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are $\boxed{\text{latent}}$.

We need a distribution from which to draw the per-document topic distribution. We use a $\boxed{\text{Dirichlet}}$ distribution as a prior for that.

And Dirichlet is used for the $\boxed{\text{allocation}}$ of the words in the documents to different topics: it is used as a prior over the distribution of words (which define the topics).

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are $\boxed{\text{latent}}$.

We need a distribution from which to draw the per-document topic distribution. We use a $\boxed{\text{Dirichlet}}$ distribution as a prior for that.

And Dirichlet is used for the $\boxed{\text{allocation}}$ of the words in the documents to different topics: it is used as a prior over the distribution of words (which define the topics).

$\rightarrow$ Latent

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are $\boxed{\text{latent}}$.

We need a distribution from which to draw the per-document topic distribution. We use a $\boxed{\text{Dirichlet}}$ distribution as a prior for that.

And Dirichlet is used for the $\boxed{\text{allocation}}$ of the words in the documents to different topics: it is used as a prior over the distribution of words (which define the topics).

$\rightarrow$ Latent Dirichlet

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are $\boxed{\text{latent}}$.

We need a distribution from which to draw the per-document topic distribution. We use a $\boxed{\text{Dirichlet}}$ distribution as a prior for that.

And Dirichlet is used for the $\boxed{\text{allocation}}$ of the words in the documents to different topics: it is used as a prior over the distribution of words (which define the topics).

$\rightarrow$ Latent Dirichlet Allocation.

Notes

Some of our variables—the documents which contain the words—are observable. But, topic structure—topics themselves, per-document topic distributions, per-document per-word topic assignments—are latent.

We need a distribution from which to draw the per-document topic distribution. We use a Dirichlet distribution as a prior for that.

And Dirichlet is used for the allocation of the words in the documents to different topics: it is used as a prior over the distribution of words (which define the topics).

$\rightarrow$ Latent Dirichlet Allocation. **LDA**.

Results

For a user-selected k, a typical implementation of LDA will return...

Results

For a user-selected k, a typical implementation of LDA will return...

The word distribution for each topic.

Results

For a user-selected k, a typical implementation of LDA will return...

The word distribution for each topic.

The topic distribution for each document.

Results

For a user-selected k, a typical implementation of LDA will return...

The word distribution for each topic.

The topic distribution for each document.

Some implementations allow you to estimate e.g. α, in which case this is also returned.

Results

For a user-selected k, a typical implementation of LDA will return...

The word distribution for each topic.

The topic distribution for each document.

Some implementations allow you to estimate e.g. α, in which case this is also returned. And perhaps some kind of fit statistic(s).

A Manifesto Example

A Manifesto Example

69 UK manifestos.

A Manifesto Example

69 UK manifestos. Some preprocessing.

A Manifesto Example

69 UK manifestos. Some preprocessing. Used `topicmodels` to fit five topics.

A Manifesto Example

69 UK manifestos. Some preprocessing. Used `topicmodels` to fit five topics. Has Gibbs sampling and variational options.

A Manifesto Example

69 UK manifestos. Some preprocessing. Used `topicmodels` to fit five topics. Has Gibbs sampling and variational options.

The (some selected) word distributions for each topic.

A Manifesto Example

69 UK manifestos. Some preprocessing. Used `topicmodels` to fit five topics. Has Gibbs sampling and variational options.

The (some selected) word distributions for each topic. Sum down the columns is one.

A Manifesto Example

69 UK manifestos. Some preprocessing. Used `topicmodels` to fit five topics. Has Gibbs sampling and variational options.

The (some selected) word distributions for each topic. Sum down the columns is one.

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
conservative	0.00188	0.00088	0.00185	0.00221	0.00168
party	0.00145	0.00067	0.00066	0.00577	0.00093
general	0.00073	0.00033	0.00018	0.00192	0.00040
election	0.00079	0.00053	0.00022	0.00235	0.00076
manifesto	0.00059	0.00078	0.00032	0.00099	0.00048
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Continued. . .

Continued...

'Top' 6 most frequent words in each topic:

Continued. . .

'Top' 6 most frequent words in each topic: might help interpretation (!)

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	people	new	[markup]	new	must
2	local	government	people	labour	government
3	government	people	new	government	labour
4	new	continue	work	people	shall
5	tax	can	[markup]	shall	can
6	liberal	conservative	support	britain	policy

Continued...

'Top' 6 most frequent words in each topic: might help interpretation (!)

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	people	new	[markup]	new	must
2	local	government	people	labour	government
3	government	people	new	government	labour
4	new	continue	work	people	shall
5	tax	can	[markup]	shall	can
6	liberal	conservative	support	britain	policy

Up to analyst to label the topics!

Continued...

'Top' 6 most frequent words in each topic: might help interpretation (!)

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	people	new	[markup]	new	must
2	local	government	people	labour	government
3	government	people	new	government	labour
4	new	continue	work	people	shall
5	tax	can	[markup]	shall	can
6	liberal	conservative	support	britain	policy

Up to analyst to label the topics!

Meaningless 'junk' topics not unusual:

Continued. . .

'Top' 6 most frequent words in each topic: might help interpretation (!)

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	people	new	[markup]	new	must
2	local	government	people	labour	government
3	government	people	new	government	labour
4	new	continue	work	people	shall
5	tax	can	[markup]	shall	can
6	liberal	conservative	support	britain	policy

Up to analyst to label the topics!

Meaningless 'junk' topics not unusual: debate as to whether one has to interpret every topic.

Continued

Continued

The topic distribution for each document. . .

Continued

The topic distribution for each document. . .

Continued

The topic distribution for each document...

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
doc 1	0.00009	0.00009	0.00009	0.00009	0.99965
doc 2	0.00011	0.00011	0.00011	0.00011	0.99954
doc 3	0.00010	0.00010	0.00010	0.00010	0.99959
doc 4	0.00006	0.00006	0.00006	0.00006	0.99978
doc 5	0.00002	0.00002	0.00002	0.00002	0.99991
doc 6	0.00019	0.00019	0.00019	0.00019	0.99924
⋮	⋮	⋮	⋮	⋮	⋮

Continued

The topic distribution for each document. . .

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
doc 1	0.00009	0.00009	0.00009	0.00009	0.99965
doc 2	0.00011	0.00011	0.00011	0.00011	0.99954
doc 3	0.00010	0.00010	0.00010	0.00010	0.99959
doc 4	0.00006	0.00006	0.00006	0.00006	0.99978
doc 5	0.00002	0.00002	0.00002	0.00002	0.99991
doc 6	0.00019	0.00019	0.00019	0.00019	0.99924
⋮	⋮	⋮	⋮	⋮	⋮

Practical Notes I

Practical Notes I

Texts are usually preprocessed:

Practical Notes I

Texts are usually preprocessed: stop words removed,

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed.

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed. Punctuation often removed.

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed. Punctuation often removed. Stemming seems less common.

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed. Punctuation often removed. Stemming seems less common.

In most social science examples, the number of topics, K, is not picked automatically.

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed. Punctuation often removed. Stemming seems less common.

In most social science examples, the number of topics, K, is not picked automatically. Analysts select various Ks and check that their results are 'robust'. But see over.

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed. Punctuation often removed. Stemming seems less common.

In most social science examples, the number of topics, K, is not picked automatically. Analysts select various Ks and check that their results are 'robust'. But see over.

As with all unsupervised learning,

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed. Punctuation often removed. Stemming seems less common.

In most social science examples, the number of topics, K, is not picked automatically. Analysts select various Ks and check that their results are 'robust'. But see over.

As with all unsupervised learning, interpretation is non-trivial, and requires a lot of validation. Rant: 'just-so' stories abound. Lazy analysts conclude whatever they want.

Practical Notes I

Texts are usually preprocessed: stop words removed, (very) rare tokens removed. Punctuation often removed. Stemming seems less common.

In most social science examples, the number of topics, K, is not picked automatically. Analysts select various Ks and check that their results are 'robust'. But see over.

As with all unsupervised learning, interpretation is non-trivial, and requires a lot of validation. Rant: 'just-so' stories abound. Lazy analysts conclude whatever they want.

Practical Notes II: Picking k

Practical Notes II: Picking k

Crudely: in social science,

Practical Notes II: Picking k

Crudely: in social science, researchers fit ‘enough’ topics until they see what they think they should.

Practical Notes II: Picking k

Crudely: in social science, researchers fit 'enough' topics until they see what they think they should. E.g. a certain topic—like `finance` suddenly peels off—so stop there.

Practical Notes II: Picking k

Crudely: in social science, researchers fit 'enough' topics until they see what they think they should. E.g. a certain topic—like `finance` suddenly peels off—so stop there.

$\rightarrow$ Check findings are robust in the neighborhood: if best model has $k = 35$, check $k = 30 - 40$ yields similar inferences.

Practical Notes II: Picking k

Crudely: in social science, researchers fit 'enough' topics until they see what they think they should. E.g. a certain topic—like `finance` suddenly peels off—so stop there.

$\rightarrow$ Check findings are robust in the neighborhood: if best model has $k = 35$, check $k = 30 - 40$ yields similar inferences.

NB: social scientists typically fit far fewer topics than CS, even to same data.

Practical Notes II: Picking k

Crudely: in social science, researchers fit 'enough' topics until they see what they think they should. E.g. a certain topic—like `finance` suddenly peels off—so stop there.

$\rightarrow$ Check findings are robust in the neighborhood: if best model has $k = 35$, check $k = 30 - 40$ yields similar inferences.

NB: social scientists typically fit far fewer topics than CS, even to same data.

Picking k, continued. . .

Picking k, continued...

CS: split into training and test sets.

Pork to Policy (Catalinac, 2016)

Pork to Policy (Catalinac, 2016)

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case:

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy.

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area.

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

1. Rise of China?

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

1. Rise of China? Need to focus on security.

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

1. Rise of China? Need to focus on security.

vs.

2. Change in Electoral System?

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

1. Rise of China? Need to focus on security.

vs.

2. Change in Electoral System? Moved from promising pork to having to deliver policy as part of Westminster-style polity.

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

1. Rise of China? Need to focus on security.

vs.

2. Change in Electoral System? Moved from promising pork to having to deliver policy as part of Westminster-style polity.

To decide, we need data source that covers all lower house legislators

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

1. Rise of China? Need to focus on security.

vs.

2. Change in Electoral System? Moved from promising pork to having to deliver policy as part of Westminster-style polity.

To decide, we need data source that covers all lower house legislators where they set out their policy priorities over time.

Pork to Policy (Catalinac, 2016)

Japan is a curious IR case: wealthy post-war not very interested in foreign policy. Recent times have seen a (re-)emergence in this area. Why?

1. Rise of China? Need to focus on security.

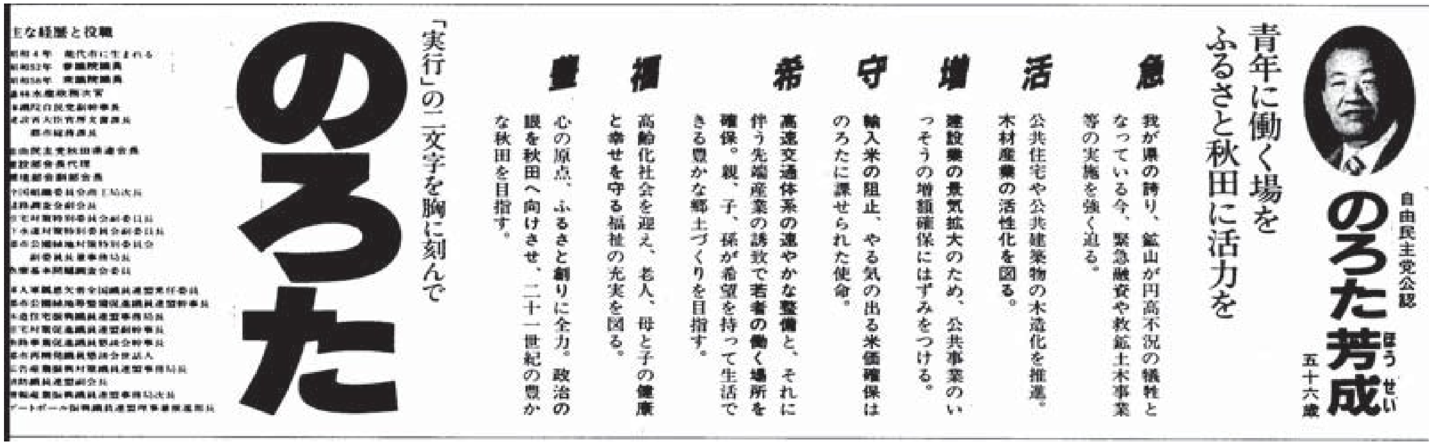
vs.

2. Change in Electoral System? Moved from promising pork to having to deliver policy as part of Westminster-style polity.

To decide, we need data source that covers all lower house legislators where they set out their policy priorities over time. See if/when they shift priorities.

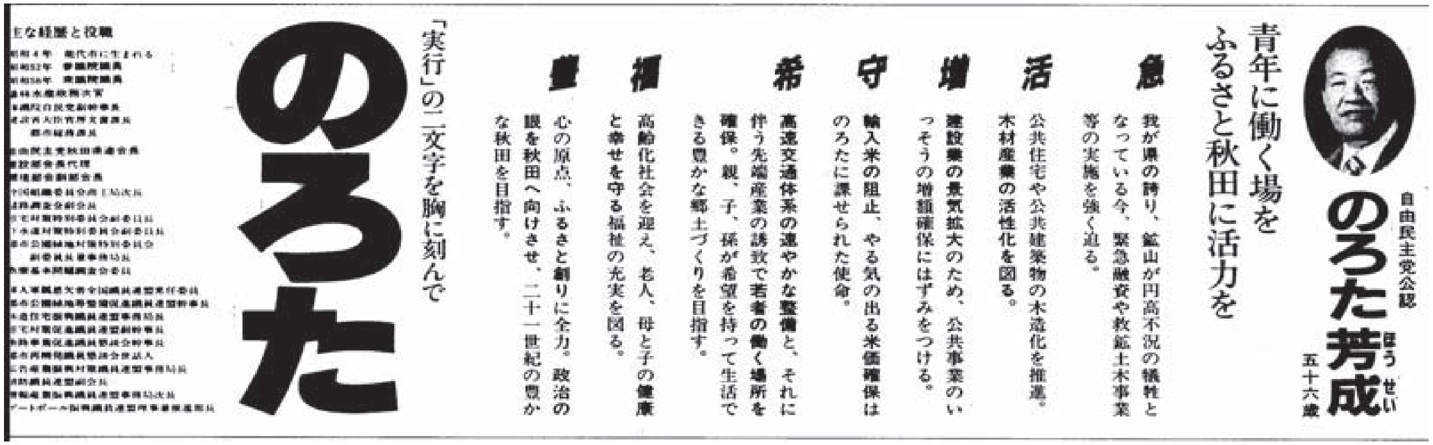
Manifestos

Manifestos

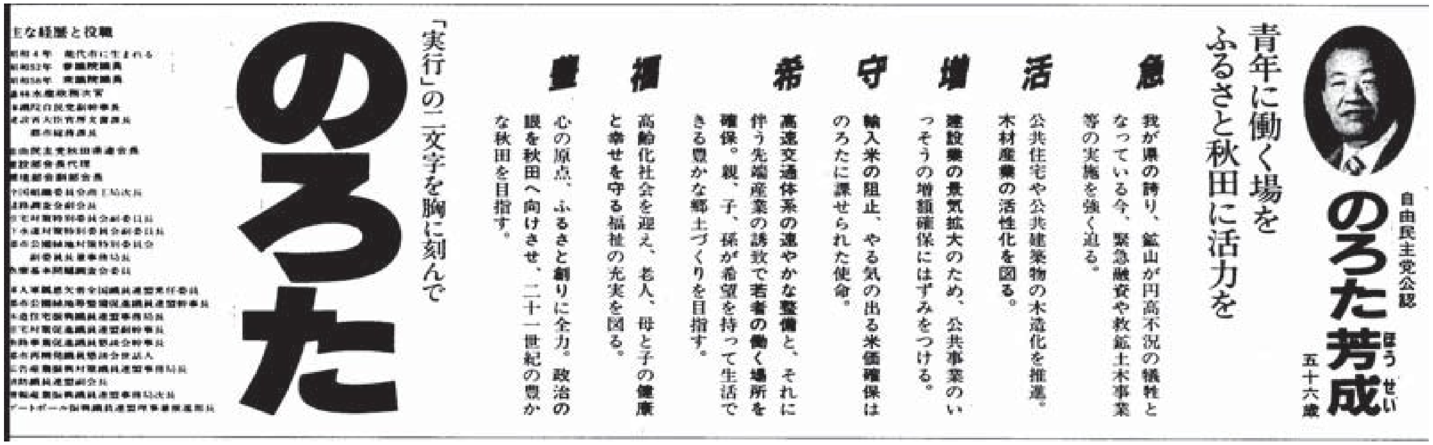


自由民主党公認
のろた芳成
ほうせい
五十六歳
青年に働く場を
ふるさと秋田に活力を
急
我が県の誇り、鉱山が円高不況の犠牲となっている今、緊急融資や救鉱土木事業等の実施を強く迫る。
活
公共住宅や公共建築物の木造化を推進。木材産業の活性化を図る。
増
建設業の景気拡大のため、公共事業のいっそうの増額確保にはずみをつける。
守
輸入米の阻止、やる気の出る米価確保はのろたに課せられた使命。
希
高速交通体系の速やかな整備と、それに伴う先端産業の誘致で若者の働く場所を確保。親、子、孫が希望を持って生活できる豊かな郷土づくりを目指す。
福
高齢化社会を迎え、老人、母と子の健康と幸せを守る福祉の充実を図る。
心の原点、ふるさと創りに全力。政治の眼を秋田へ向けさせ、二十一世紀の豊かな秋田を目指す。
「実行」の二文字を胸に刻んで
のろた

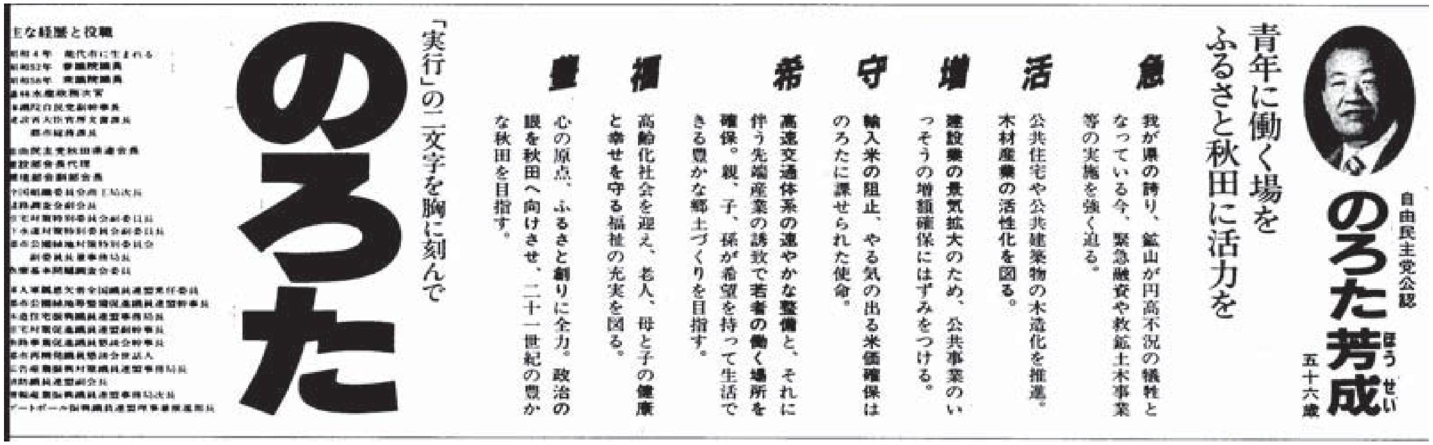
Manifestos



7,497.



7,497. 1986–2009.



自由民主党公認

のろた芳成（ほうせい）

五十六歳

青年に働く場を
ふるさと秋田に活力を

急 我が県の誇り、鉱山が円高不況の犠牲となっている今、緊急融資や救鉱土木事業等の実施を強く迫る。

活 公共住宅や公共建築物の木造化を推進。木材産業の活性化を図る。

増 建設業の景気拡大のため、公共事業のいっそうの増額確保にはずみをつける。

守 輸入米の阻止、やる気の出る米価確保はのろたに課せられた使命。

希 高速交通体系の速やかな整備と、それに伴う先端産業の誘致で若者の働く場所を確保。親、子、孫が希望を持って生活できる豊かな郷土づくりを目指す。

福 高齢化社会を迎え、老人、母と子の健康と幸せを守る福祉の充実を図る。

豊 心の原点、ふるさと創りに全力。政治の眼を秋田へ向けさせ、二十一世紀の豊かな秋田を目指す。

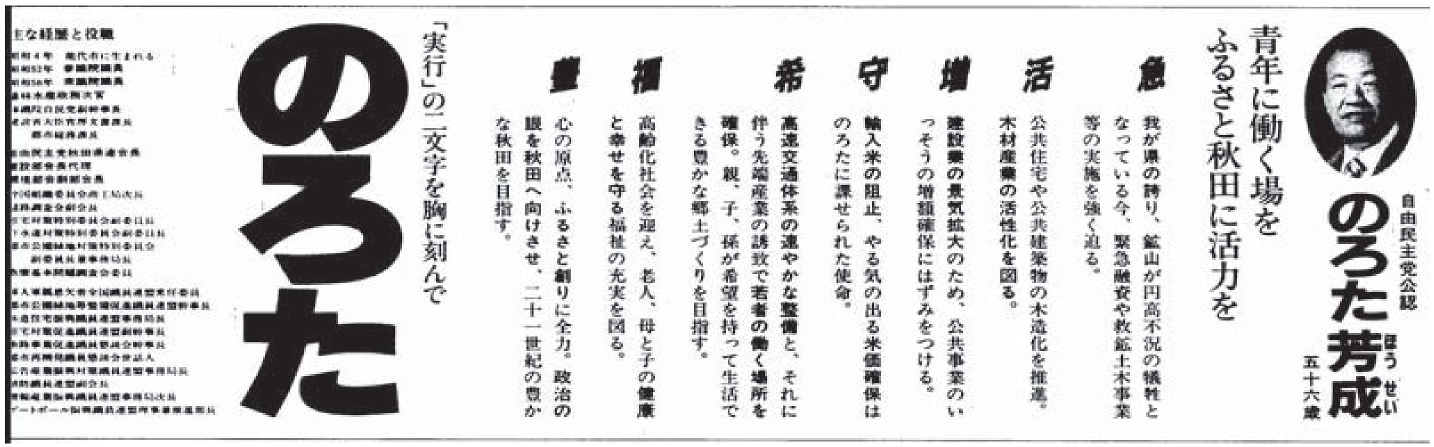
「実行」の二文字を胸に刻んで

のろた

主な経歴と役職

7,497. 1986–2009. Standardized form.

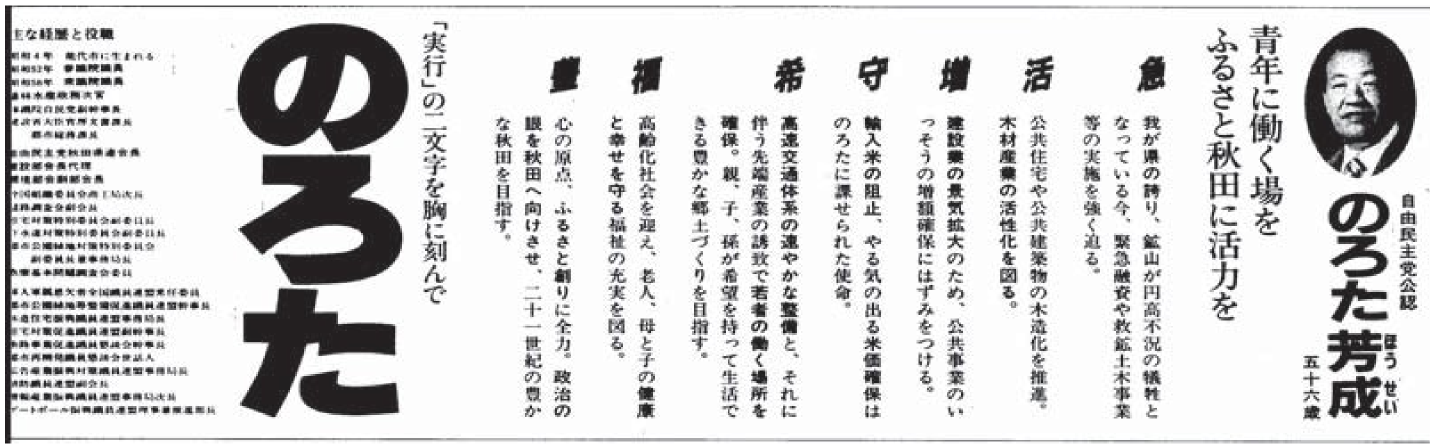
Manifestos



7,497. 1986–2009. Standardized form.

"... instructed to write whatever they want in the form and return it before 5 PM of the first day of the campaign. At least two days before the election, local electoral commissions are required to distribute the forms of all candidates running in the district to all registered voters"

Manifestos



自由民主党公認

のろた芳成（ほうせい）

五十六歳

青年に働く場を
ふるさと秋田に活力を

急

我が県の誇り、鉱山が円高不況の犠牲となっている今、緊急融資や救鉱土木事業等の実施を強く迫る。

活

公共住宅や公共建築物の木造化を推進。木材産業の活性化を図る。

増

建設業の景気拡大のため、公共事業のいっそうの増額確保にはずみをつける。

守

輸入米の阻止、やる気の出る米価確保はのろたに課せられた使命。

希

高速交通体系の速やかな整備と、それに伴う先端産業の誘致で若者の働く場所を確保。親、子、孫が希望を持って生活できる豊かな郷土づくりを目指す。

福

高齢化社会を迎え、老人、母と子の健康と幸せを守る福祉の充実を図る。

豊

心の原点、ふるさと創りに全力。政治の眼を秋田へ向けさせ、二十一世紀の豊かな秋田を目指す。

「実行」の二文字を胸に刻んで

のろた

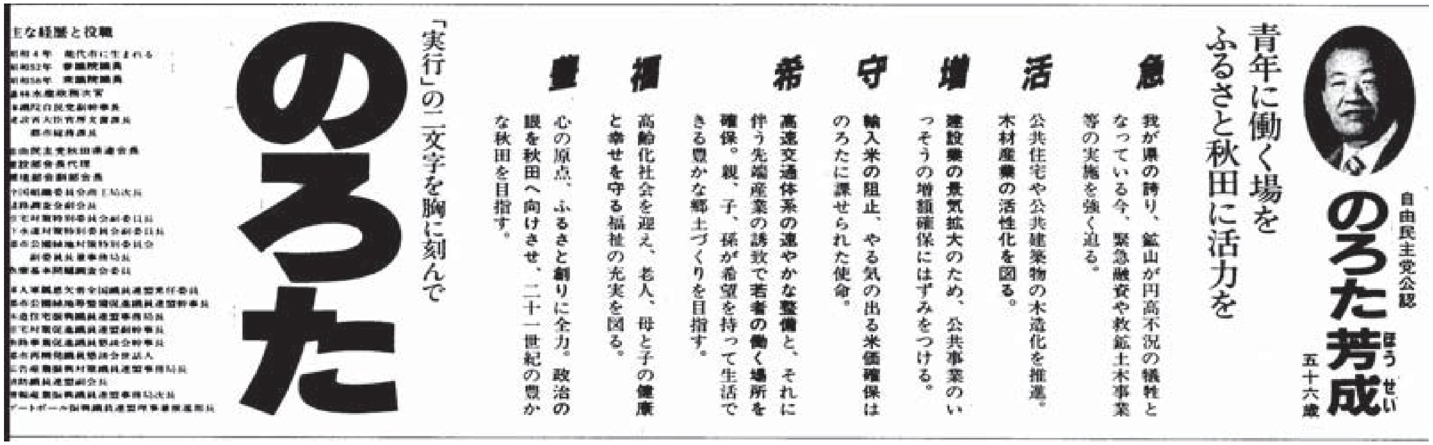
主な経歴と役職

7,497. 1986–2009. Standardized form.

"...instructed to write whatever they want in the form and return it before 5 PM of the first day of the campaign. At least two days before the election, local electoral commissions are required to distribute the forms of all candidates running in the district to all registered voters"

Manifestos were hand transcribed from microfilm.

Manifestos



自由民主党公認
のろた芳成（ほうせい）
五十六歳

青年に働く場を
ふるさと秋田に活力を

急　我が県の誇り、鉱山が円高不況の犠牲となっている今、緊急融資や救鉱土木事業等の実施を強く迫る。

活　公共住宅や公共建築物の木造化を推進。木材産業の活性化を図る。

増　建設業の景気拡大のため、公共事業のいっそうの増額確保にはずみをつける。

守　輸入米の阻止、やる気の出る米価確保はのろたに課せられた使命。

希　高速交通体系の速やかな整備と、それに伴う先端産業の誘致で若者の働く場所を確保。親、子、孫が希望を持って生活できる豊かな郷土づくりを目指す。

福　高齢化社会を迎え、老人、母と子の健康と幸せを守る福祉の充実を図る。

豊　心の原点、ふるさと創りに全力。政治の眼を秋田へ向けさせ、二十一世紀の豊かな秋田を目指す。

「実行」の二文字を胸に刻んで

のろた

主な経歴と役職

7,497. 1986–2009. Standardized form.

"...instructed to write whatever they want in the form and return it before 5 PM of the first day of the campaign. At least two days before the election, local electoral commissions are required to distribute the forms of all candidates running in the district to all registered voters"

Manifestos were hand transcribed from microfilm. Japanese install of Windows/R used to fit LDA.

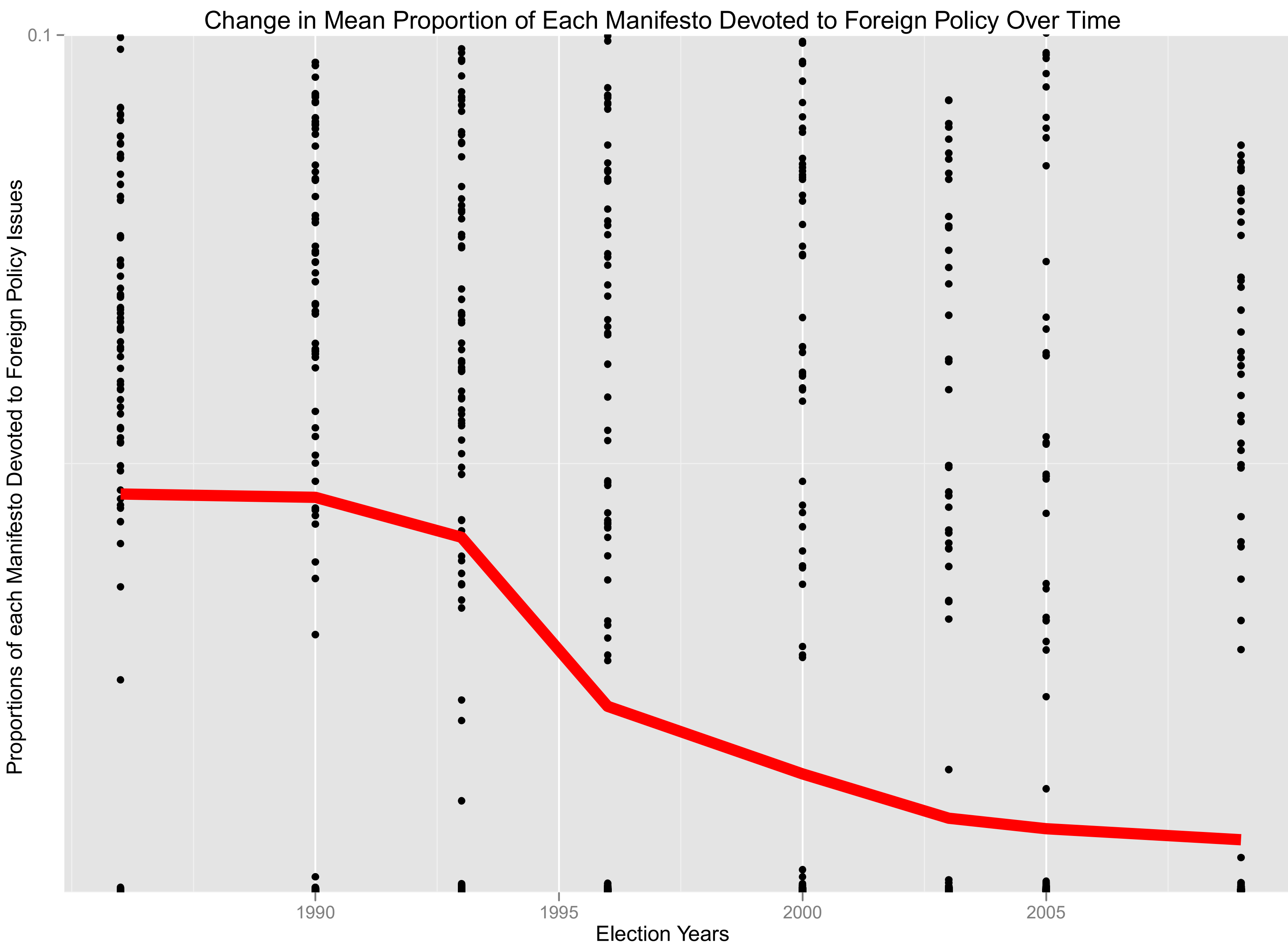
Topic Distribution over Words

Topic Distribution over Words

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
1	改革	年金	推進	区	政治	日本
2	郵政	円	整備	政策	改革	国
3	民営	廃止	図る	地域	国民	外交
4	小泉	改革	つとめる	まち	企業	国家
5	構造	兆	社会	鹿児島	自民党	社会
6	政府	実現	対策	全力	日本	国民
7	官	無駄	振興	選挙	共産党	保障
8	推進	日本	充実	国政	献金	安全
9	民	増税	促進	作り	金権	地域
10	自民党	削減	安定	横浜	党	拉致
11	日本	一元化	確立	対策	選挙	経済
12	制度	政権	企業	中小	禁止	守る
13	民間	子供	実現	発電	憲法	問題
14	年金	地域	中小	推進	腐敗	北朝鮮
15	実現	ひと	育成	エネルギー	団体	教育
16	進める	サラリーマン	制度	企業	区	責任
17	断行	制度	政治	声	ソ連	力
18	地方	議員	地域	実現	守る	創る
19	止める	金	福祉	活性	平和	安心
20	保障	民主党	事業	自民党	円	目指す
21	財政	年間	改革	地方	反対	誤り
22	作る	一掃	確保	尽くす	真	憲法
23	賛成	郵政	強化	商店	是正	可能
24	社会	道路	教育	いかす	一掃	道
25	国民	交代	施設	全国	悪政	未来
26	公務員	社会保険庁	生活	政党	抜本	ひと
27	力	月額	支援	ひと	定数	再生
28	経済	手当	環境	支援	政党	将来
29	国	談合	発展	経済	金丸	解決
30	安心	支援	施策	福祉	改革	基本

Change in proportion of 'Pork' Topic

Change in proportion of ‘Pork’ Topic



Change in proportion of 'Foreign Policy' Topic

Change in proportion of ‘Foreign Policy’ Topic