

Content-Driven  
Author Reputation and Text Trust  
for the Wikipedia

Luca de Alfaro

UC Santa Cruz

Joint work with

Bo Adler, Ian Pye, Caitlin Sadowski (UCSC)

Wikimania, August 2007

# Author Reputation and Text Trust

---

## Author Reputation:

- **Goal:** Encourage authors to provide lasting contributions.

# Author Reputation and Text Trust

---

## Author Reputation:

- **Goal:** Encourage authors to provide lasting contributions.

## Text Trust:

- **Goal:** provide a measure of the reliability of the text.
- **Method:** computed from the reputation of the authors who create and revise the text.

occasionally recruited from outside the Folketing.

Since 27 November 2001, the economist Anders Fjogh Rasmussen has been Prime Minister to Denmark.

As known in other parliamentary systems of government,

# Reputation: Our guiding principles

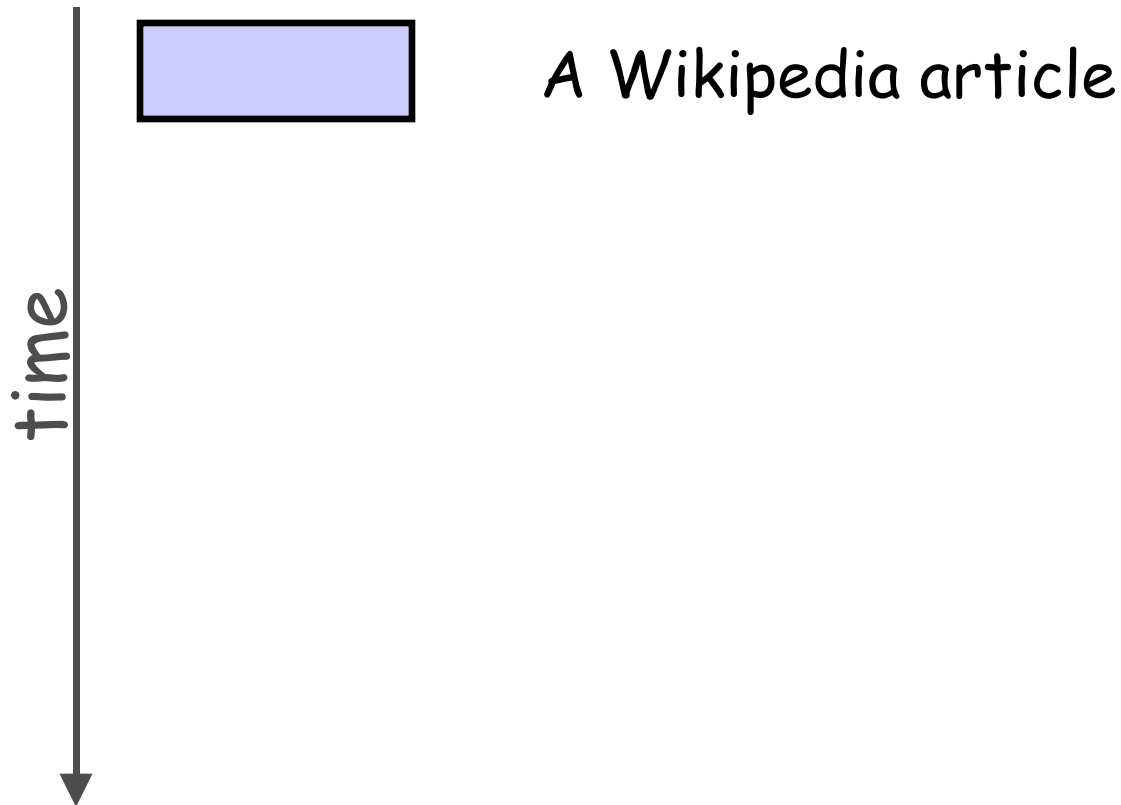
---

- **Do not alter the Wikipedia user experience**
  - Compute reputation from content evolution, rather than user-to-user comments.
- **Be welcoming to all users**
  - Never publicly display user reputation values. Authors know only their own reputation.
- **Be objective**
  - Rely on content evolution rather than comments.
  - Quantitatively evaluate how well it works.

# Content-driven reputation

---

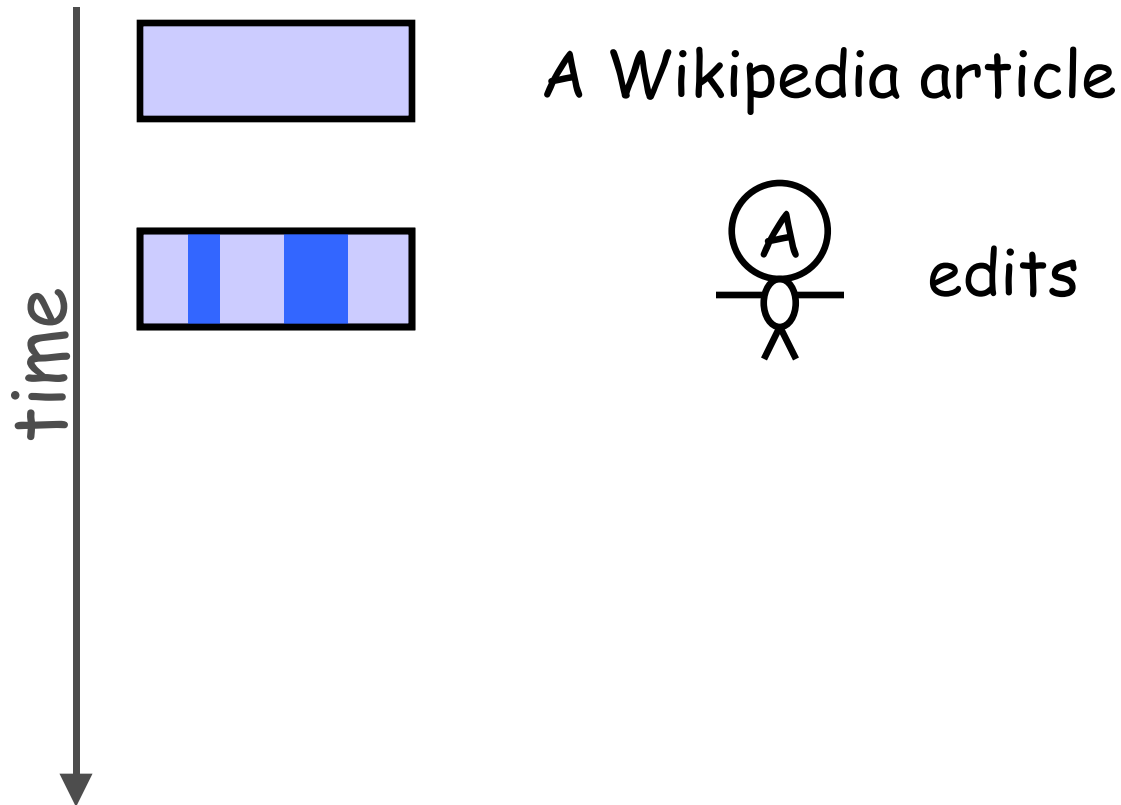
- Authors of long-lived contributions gain reputation
- Authors of reverted contributions lose reputation



# Content-driven reputation

---

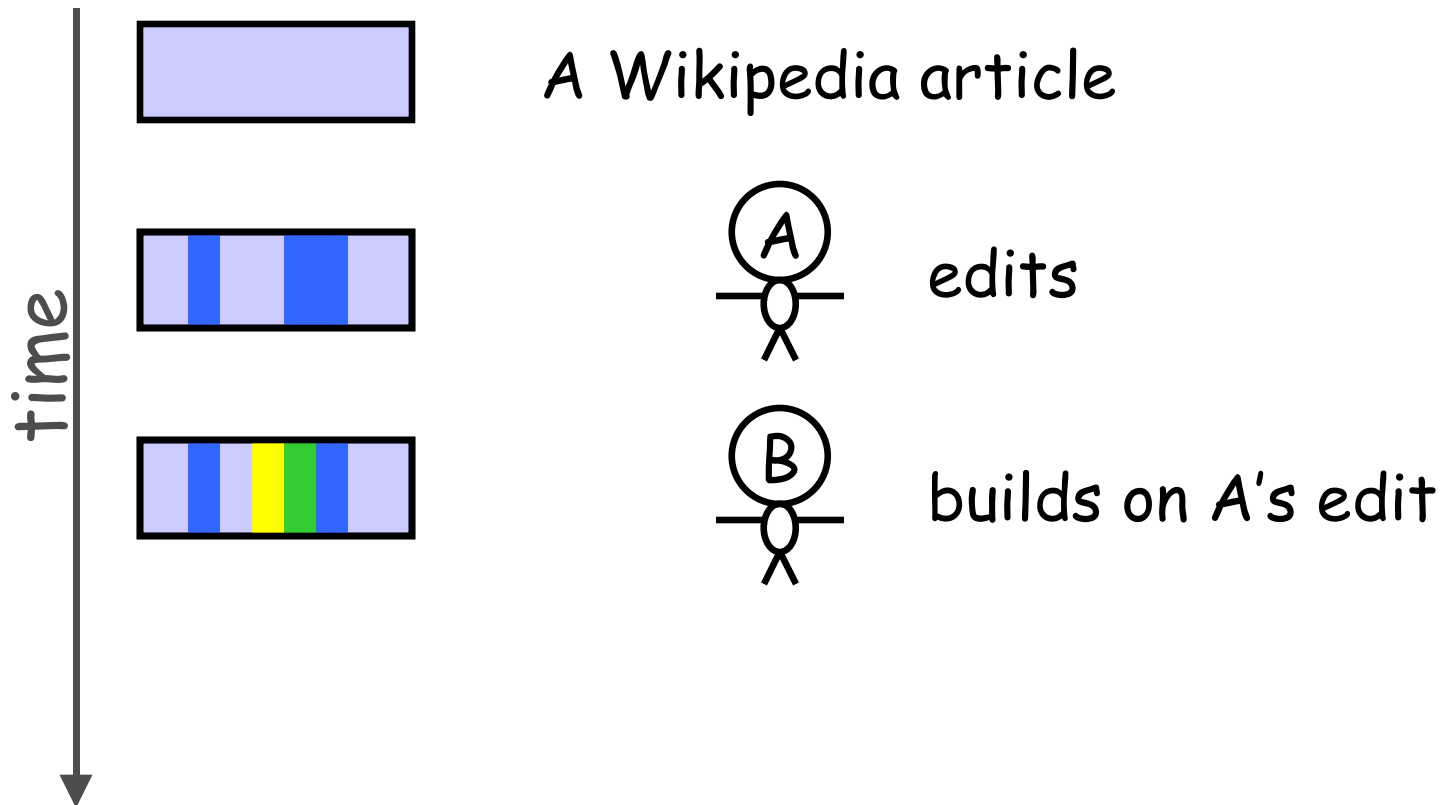
- Authors of long-lived contributions gain reputation
- Authors of reverted contributions lose reputation



# Content-driven reputation

---

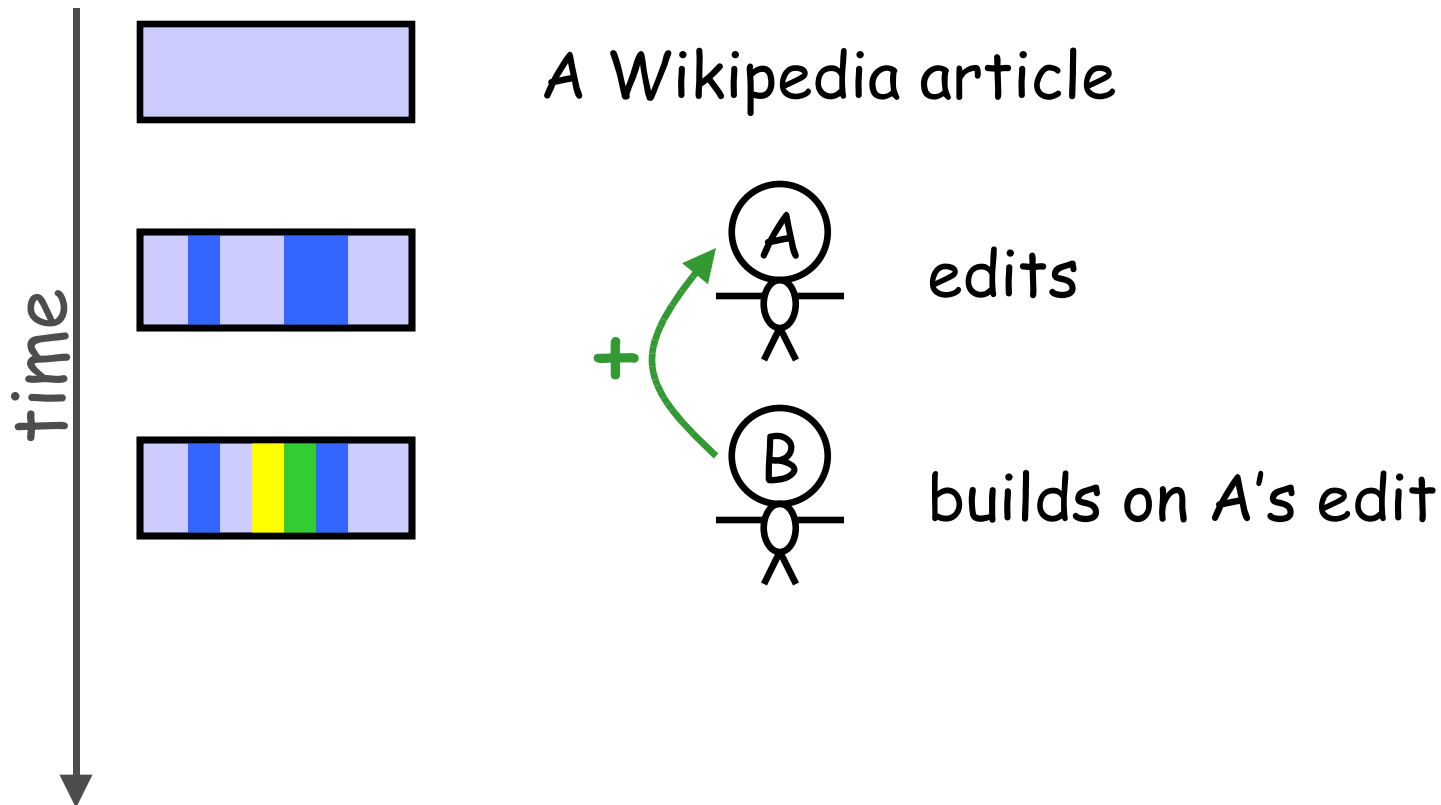
- Authors of long-lived contributions gain reputation
- Authors of reverted contributions lose reputation



# Content-driven reputation

---

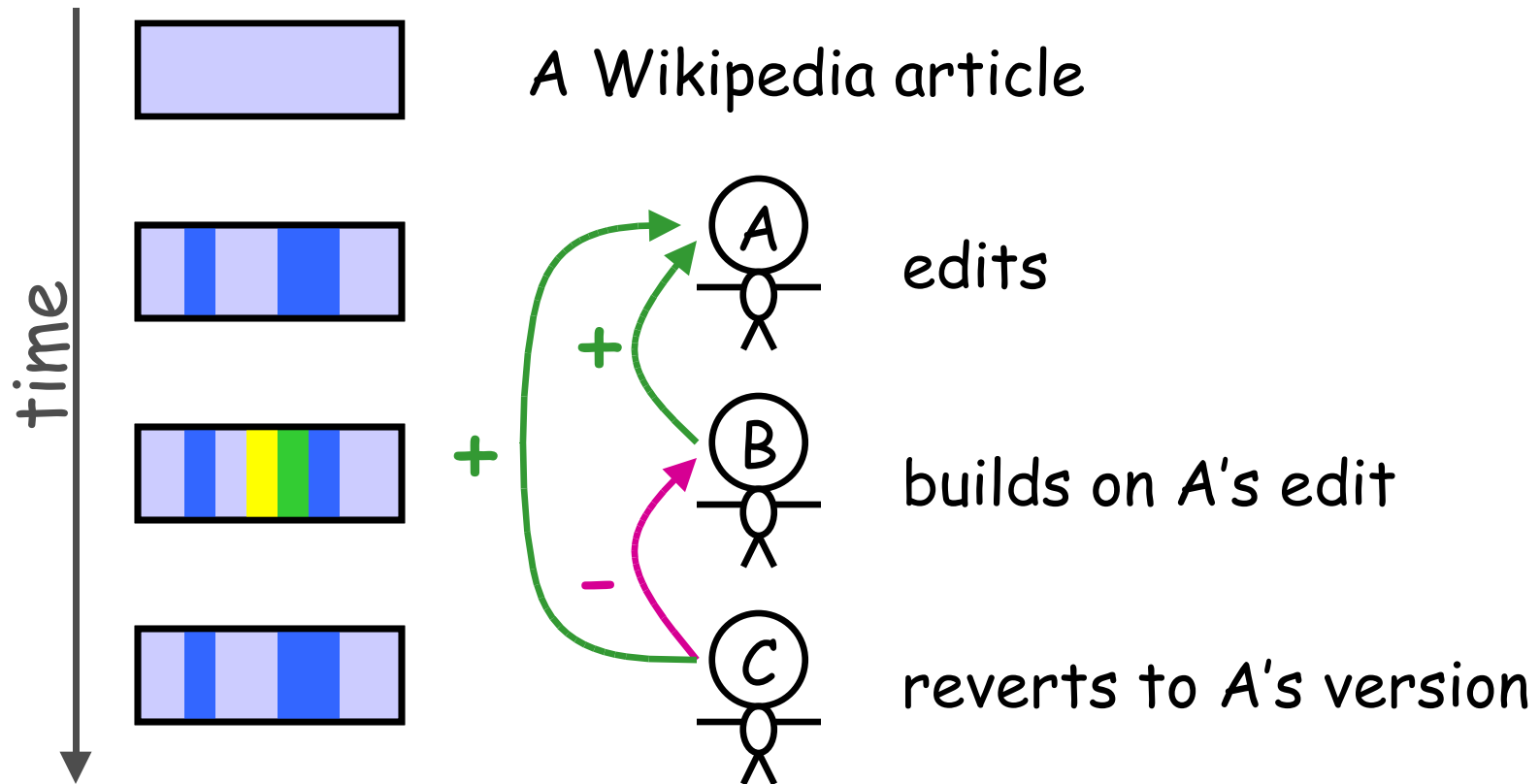
- Authors of long-lived contributions gain reputation
- Authors of reverted contributions lose reputation





# Content-driven reputation

- Authors of long-lived contributions gain reputation
- Authors of reverted contributions lose reputation



# Content-driven reputation mitigates reputation wars

---

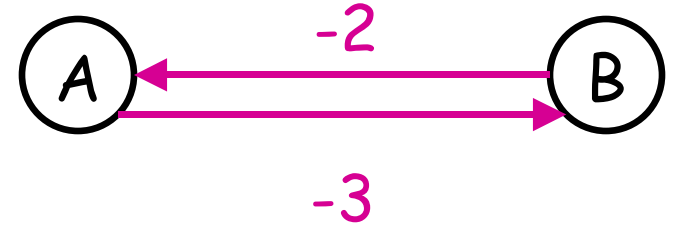
Wars in user-driven reputation:



# Content-driven reputation mitigates reputation wars

---

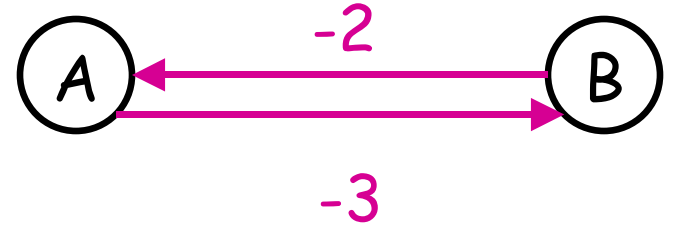
Wars in user-driven reputation:



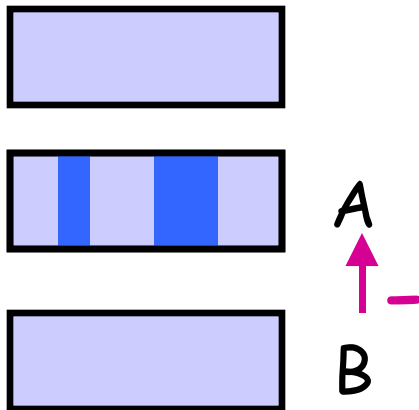
# Content-driven reputation mitigates reputation wars

---

Wars in user-driven reputation:



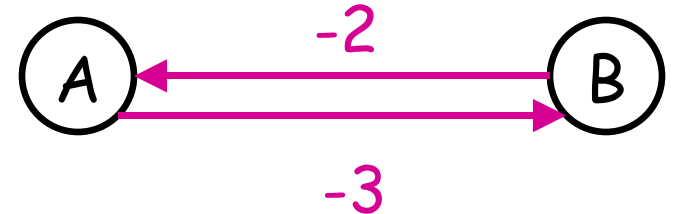
Wars in content-driven reputation:



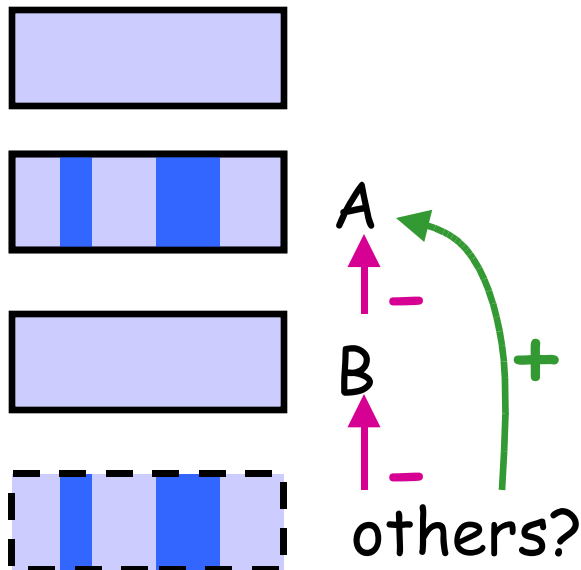
- B can badmouth A by undoing her work

# Content-driven reputation mitigates reputation wars

Wars in user-driven reputation:



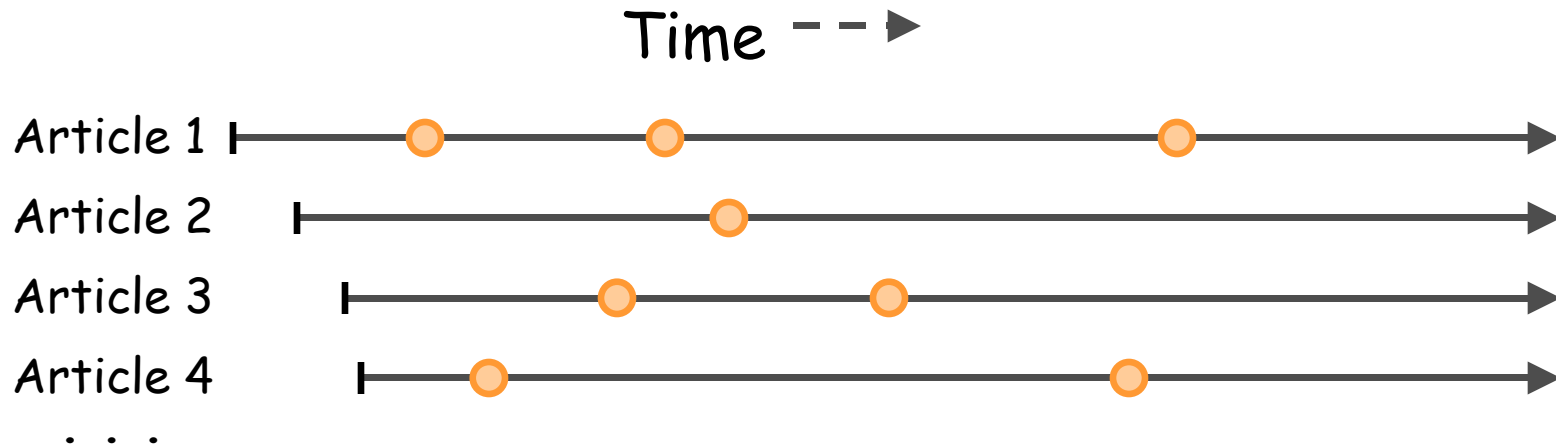
Wars in content-driven reputation:



- B can badmouth A by undoing her work
- But this is risky: if others then re-instate A's work, it is B's reputation that suffers.

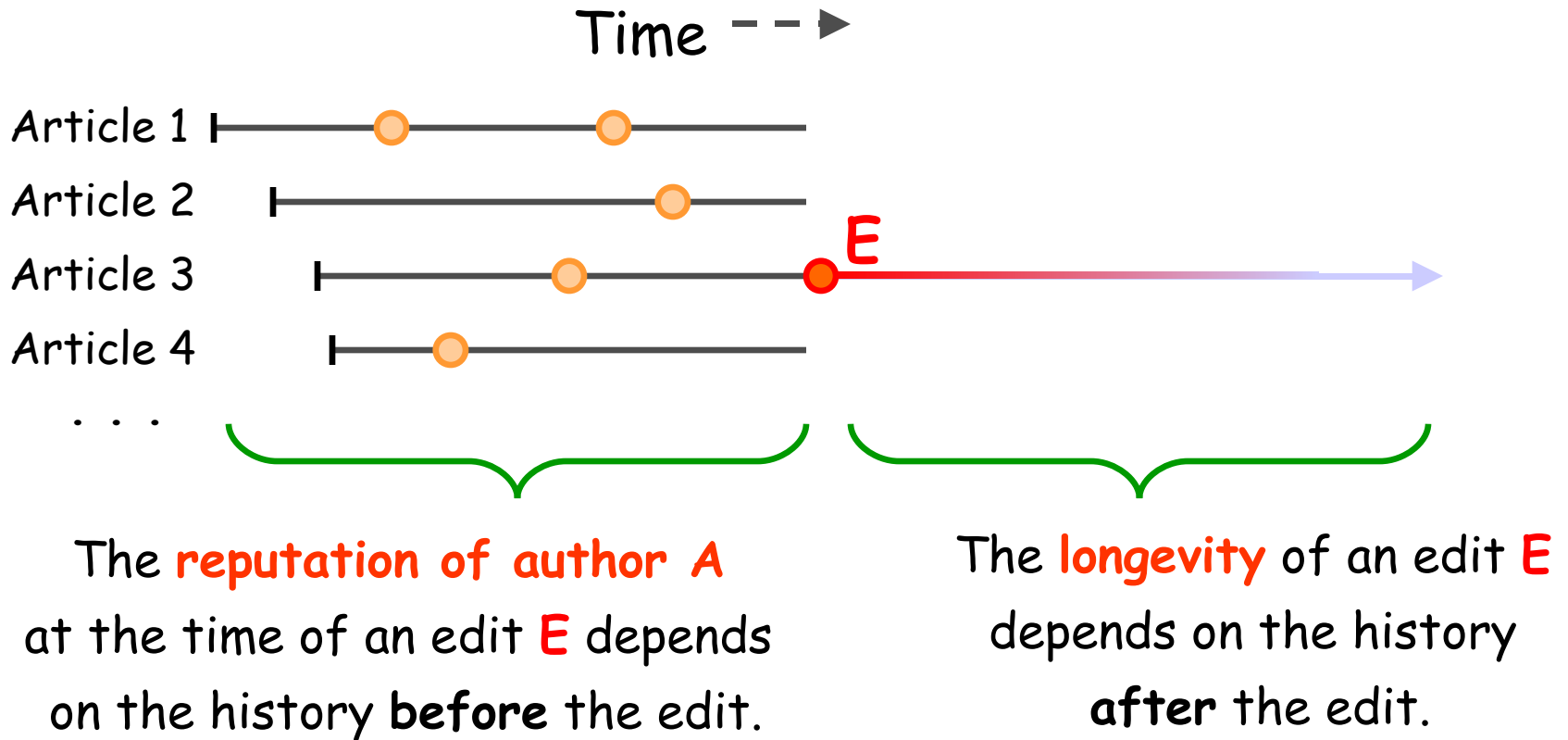
# Validation: Does our reputation have predictive value?

---



○ = edits by user **A**

# Validation: Does our reputation have predictive value?



Can we show a correlation between **author reputation** and **edit longevity** ?

# Building a content-driven reputation system for Wikipedia

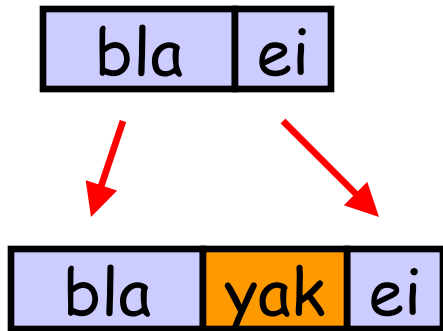
This is a summary; for details see:

B.T. Adler, L. de Alfaro. [A Content Driven Reputation System for the Wikipedia](#). In Proc. of WWW 2007.



# What is a "contribution"?

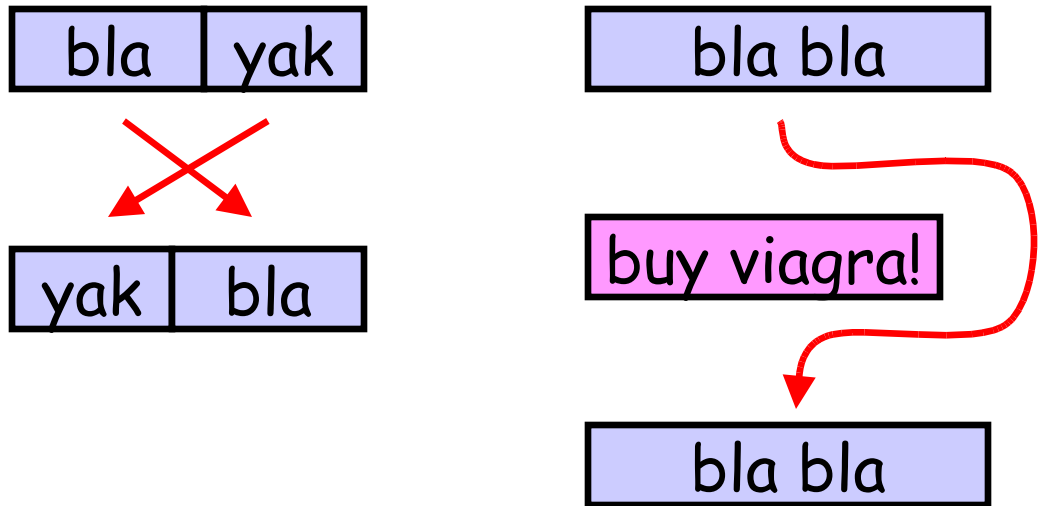
## Text



We measure how long the added text survives.

Based on text tracking.

## Edit



We measure how long the "edit" (reorganization) survives.

Based on edit distance.

# Text

---

version 9

bla	bla	wuga	boink
5	8	9	6

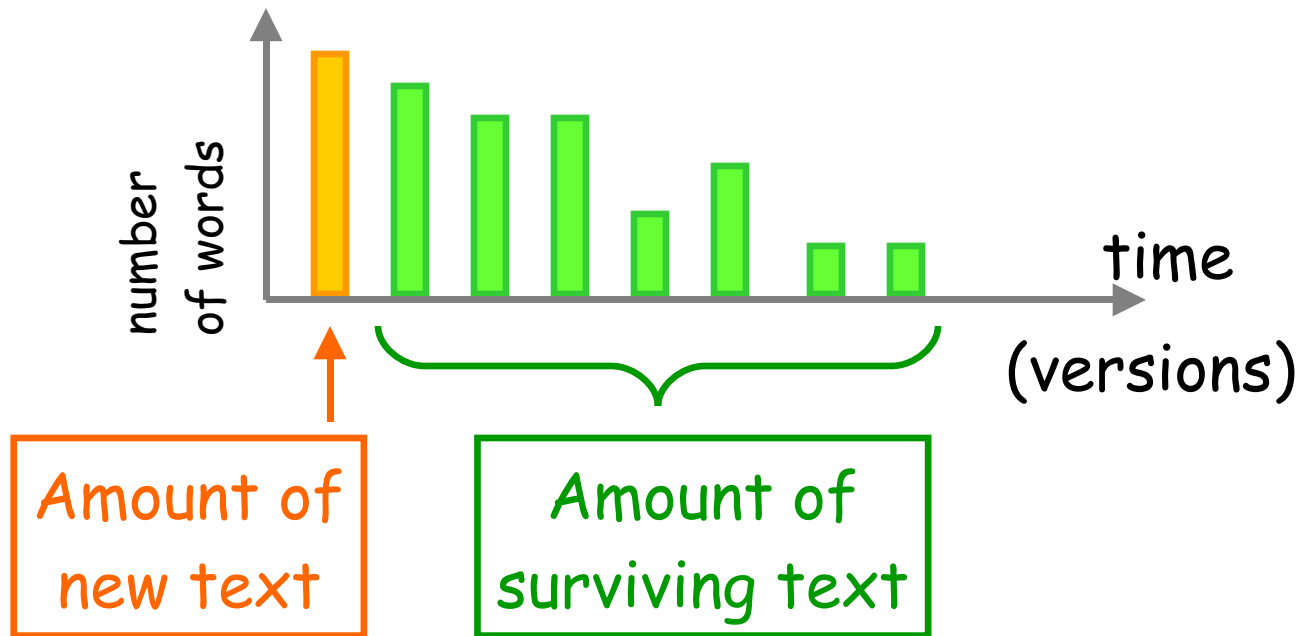
version 10

bla	bla	wuga	wuga	wuga	boink
5	8	10	10	9	6

We label each word with the version where it was introduced. This enables us to keep track of how long it lives.

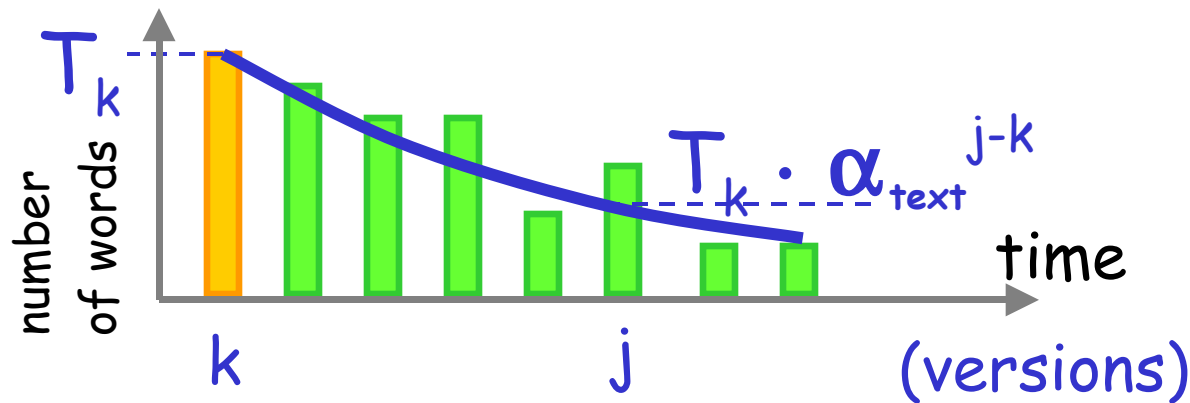
# Text: the destiny of a contribution

---



The life of the text introduced at a revision.

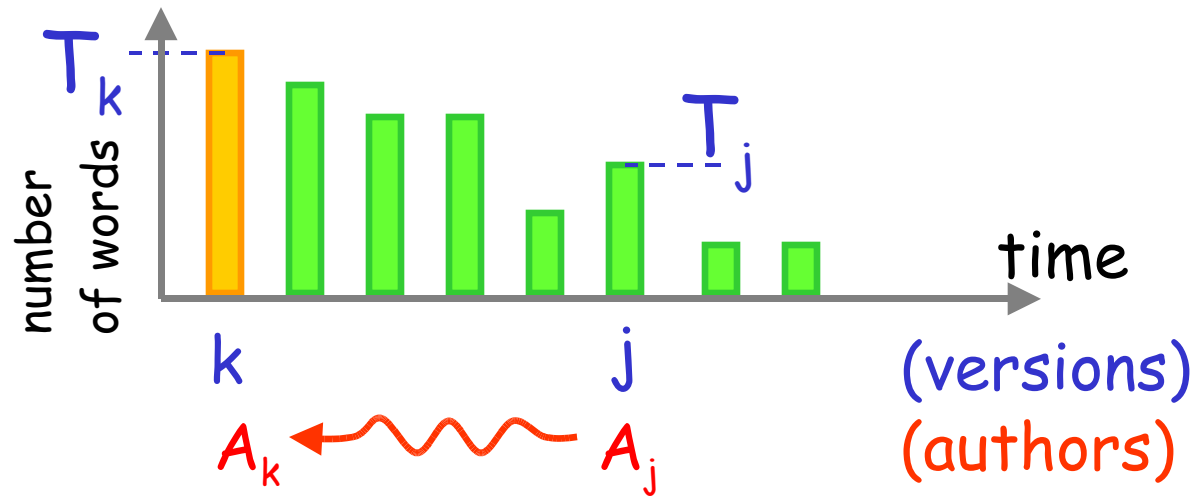
# Text: Longevity



- Text longevity: the  $\alpha_{\text{text}} \in [0,1]$  that yields the best geometrical approximation for the amount of residual text.
- Short-lived text:  $\alpha_{\text{text}} < 0.2$  (at most 20% of the text makes it from one version to the next).

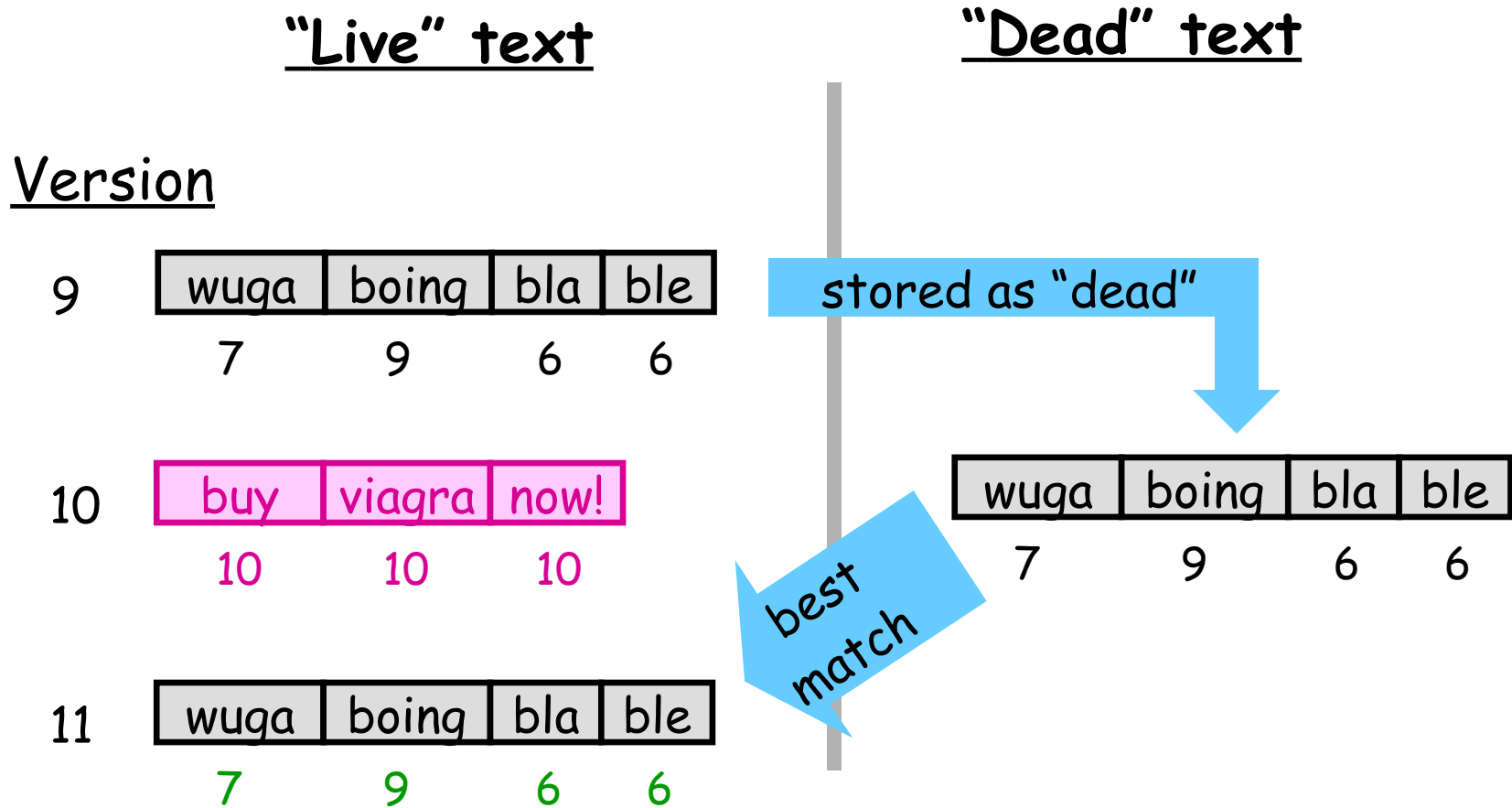
# Text: Reputation update

---



As a consequence of edit  $j$ , we increase the reputation of  $A_k$  by an amount proportional to  $T_j$  and to the reputation of  $A_j$

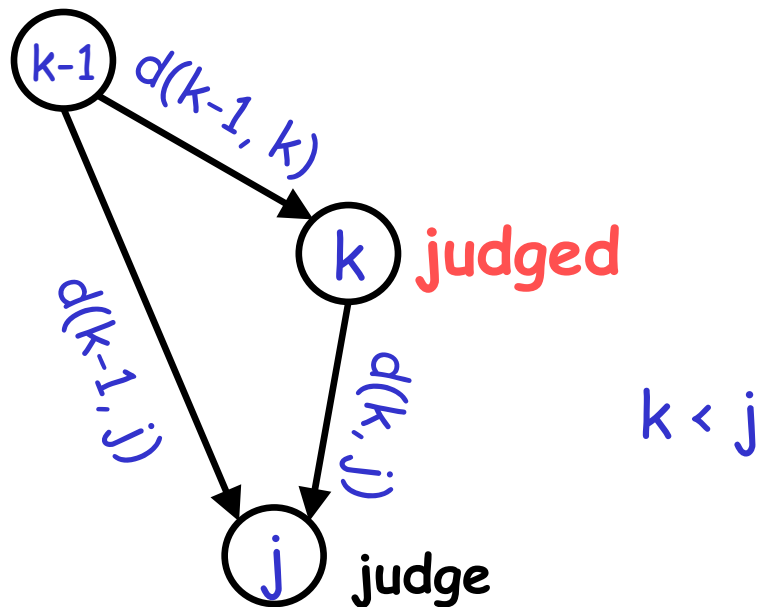
# Measuring surviving text



We track authorship of deleted text, and we match the text of new versions both with live and with dead text.

# Edit

---

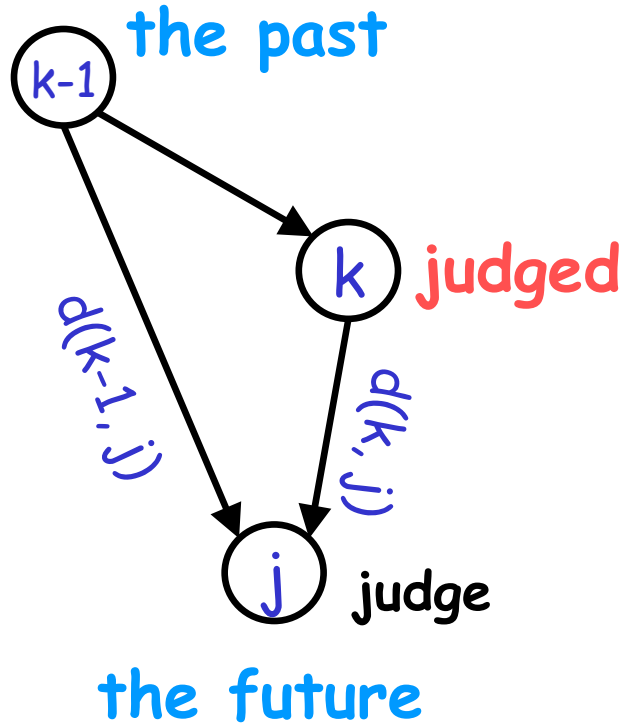


We compute the edit distance between versions  $k-1$ ,  $k$ , and  $j$ , with  $k < j$

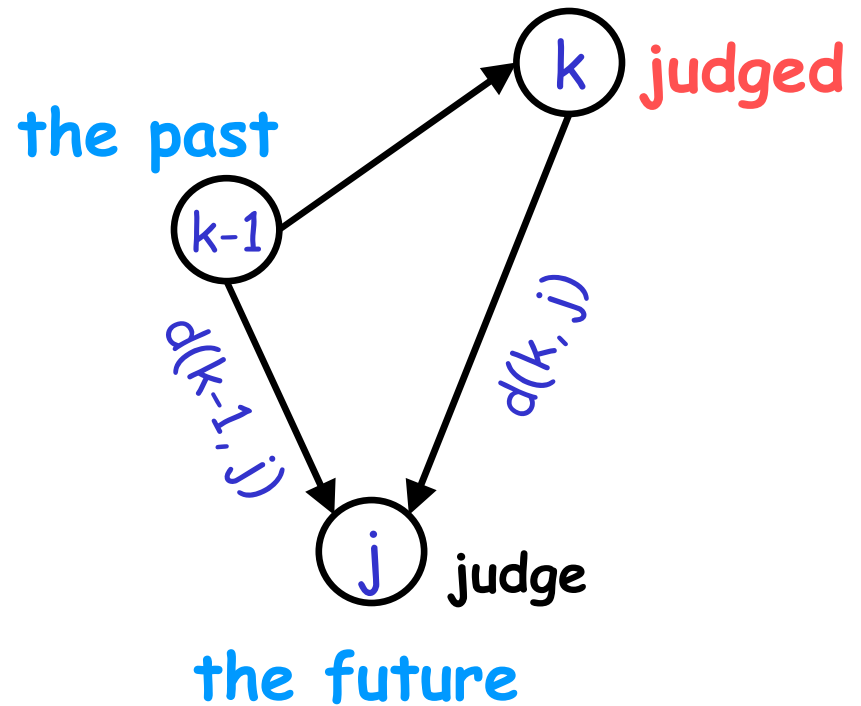
(see paper for details on the distance)

# Edit: good or bad?

---



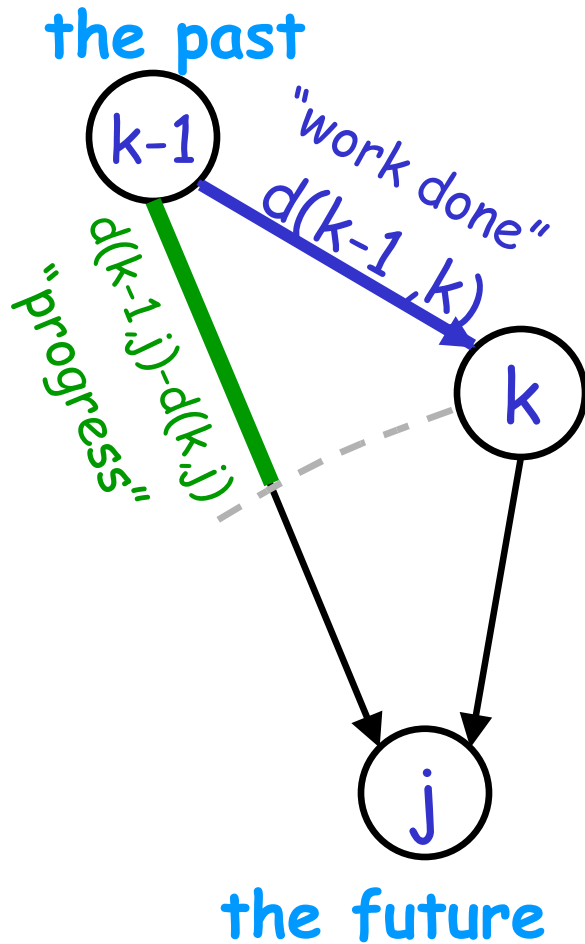
k is **good**:  $d(k-1, j) > d(k, j)$   
"k went towards the future"



k is **bad**:  $d(k-1, j) < d(k, j)$   
"k went against the future"



# Edit: Longevity



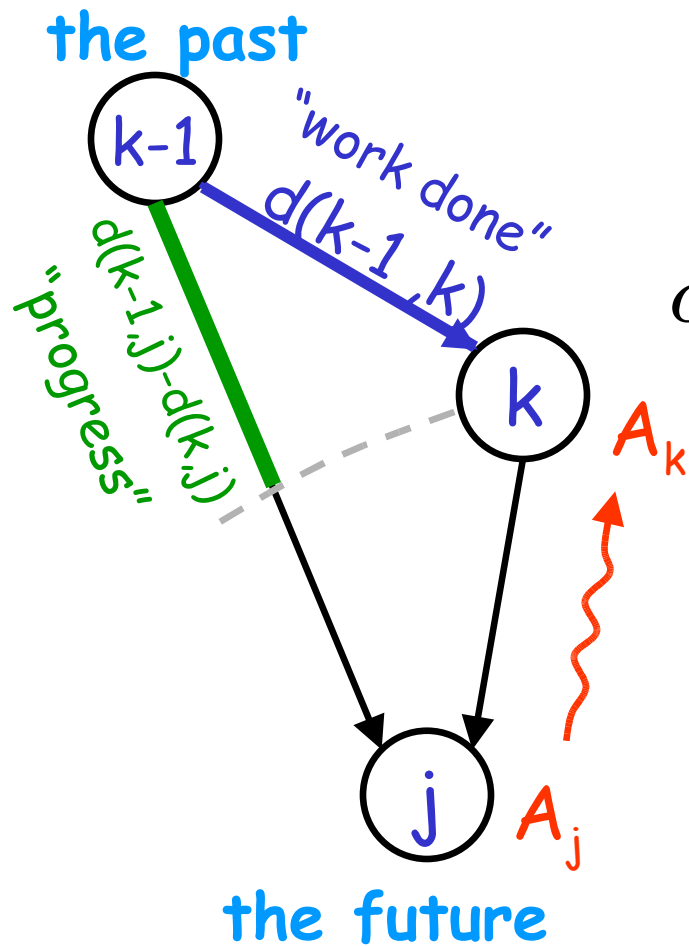
## Edit Longevity:

$$\alpha_{edit} = \frac{d(k-1, j) - d(k, j)}{d(k-1, k)}$$

The fraction of change that is in the same direction of the future.

- $\alpha_{edit} \simeq 1$ : k is a good edit
- $\alpha_{edit} \simeq -1$ : k is reverted

# Edit: Updating reputation



$\alpha_{edit} =$

Edit Longevity:

$$\frac{d(k-1, j) - d(k, j)}{d(k-1, k)}$$

Reputation update:

The reputation of  $A_k$

- increases if  $\alpha_{edit} > 0$ ,
- decreases if  $\alpha_{edit} < 0$ .

(see paper for details)

# Data Sets

---

- English till Feb 07 1,988,627 pages, 40,455,416 versions
- French till Feb 07 452,577 pages, 5,643,636 versions
- Italian till May 07 301,584 pages, 3,129,453 versions

The entire Wikipedias, with the whole history, not just a sample (we wanted to compute the reputation using all edits of each user).

# Results: English Wikipedia, in detail

% of edits below a given longevity →

	Bin	%_data	$\tau < 0.8$	$\tau < 0.4$	$\tau < 0.0$	$\tau < -0.4$	$\tau < -0.8$
$\log(1 + \text{reputation})$ ↓	0	16.922	93.11	91.65	89.15	83.76	73.53
	1	1.191	77.24	69.83	65.60	61.11	56.00
	2	1.335	69.53	57.08	49.79	45.71	41.25
	3	1.627	38.00	28.61	20.23	16.16	13.62
	4	2.780	32.84	22.31	13.32	9.57	8.04
	5	4.408	41.70	15.76	5.90	3.80	2.57
	6	6.698	29.40	16.74	7.54	4.35	3.12
	7	8.281	32.04	15.16	5.44	2.25	1.40
	8	12.233	34.06	16.64	6.78	3.79	2.73
	9	44.524	32.55	15.51	5.05	1.88	1.14

# Results: English Wikipedia, in detail

% of edits below a given longevity

	Bin	% data	$l < 0.8$	$l < 0.4$	$l < 0.0$	$l < -0.4$	$l < -0.8$
low rep	0	16.922	93.11	91.65	89.15	83.76	73.53
	1	1.191	77.24	69.83	65.60	61.11	56.00
log(1 + reputation) ↓	2	1.335	69.53	57.08	49.79	45.71	41.25
	3	1.627	38.00	28.61	20.23	16.16	13.62
	4	2.780	32.84	22.31	13.32	9.57	8.04
	5	4.408	41.70	15.76	5.90	3.80	2.57
	6	6.698	29.40	16.74	7.54	4.35	3.12
	7	8.281	32.04	15.16	5.44	2.25	1.40
	8	12.233	34.06	16.64	6.78	3.79	2.73
	9	44.524	32.55	15.51	5.05	1.88	1.14

Short-Lived

# Predictive power of low reputation

**Low-reputation:** Lower 20% of range

**Short-lived edits**

$$\alpha_{\text{edit}} \leq -0.8$$

(almost entirely undone)

**Short-lived text**

$$\alpha_{\text{text}} \leq 0.2$$

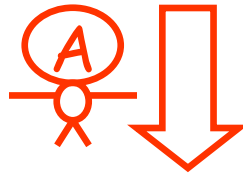
(less than 20%  
survives each revision)

Wikipedia	Short-lived edits		Short-lived text	
	Precision	Recall	Precision	Recall
English	74.2 %	84.5 %	62.0 %	83.0 %
French	38.4 %	41.2 %	17.3 %	48.1 %
Italian	21.7 %	30.4 %	6.2 %	21.4 %

# Text trust

---

Yadda yadda wuga wuga | bla bla bla bing bong



Yadda yadda wuga wuga | yak yak yuk | bla bla bla bing bong

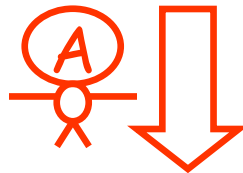
Old text is colored according to the reputation of its original author, and of all subsequent revisors (including A).

New text is colored according to the reputation of A

# Text trust

---

Yadda yadda wuga wuga | bla bla bla bing bong

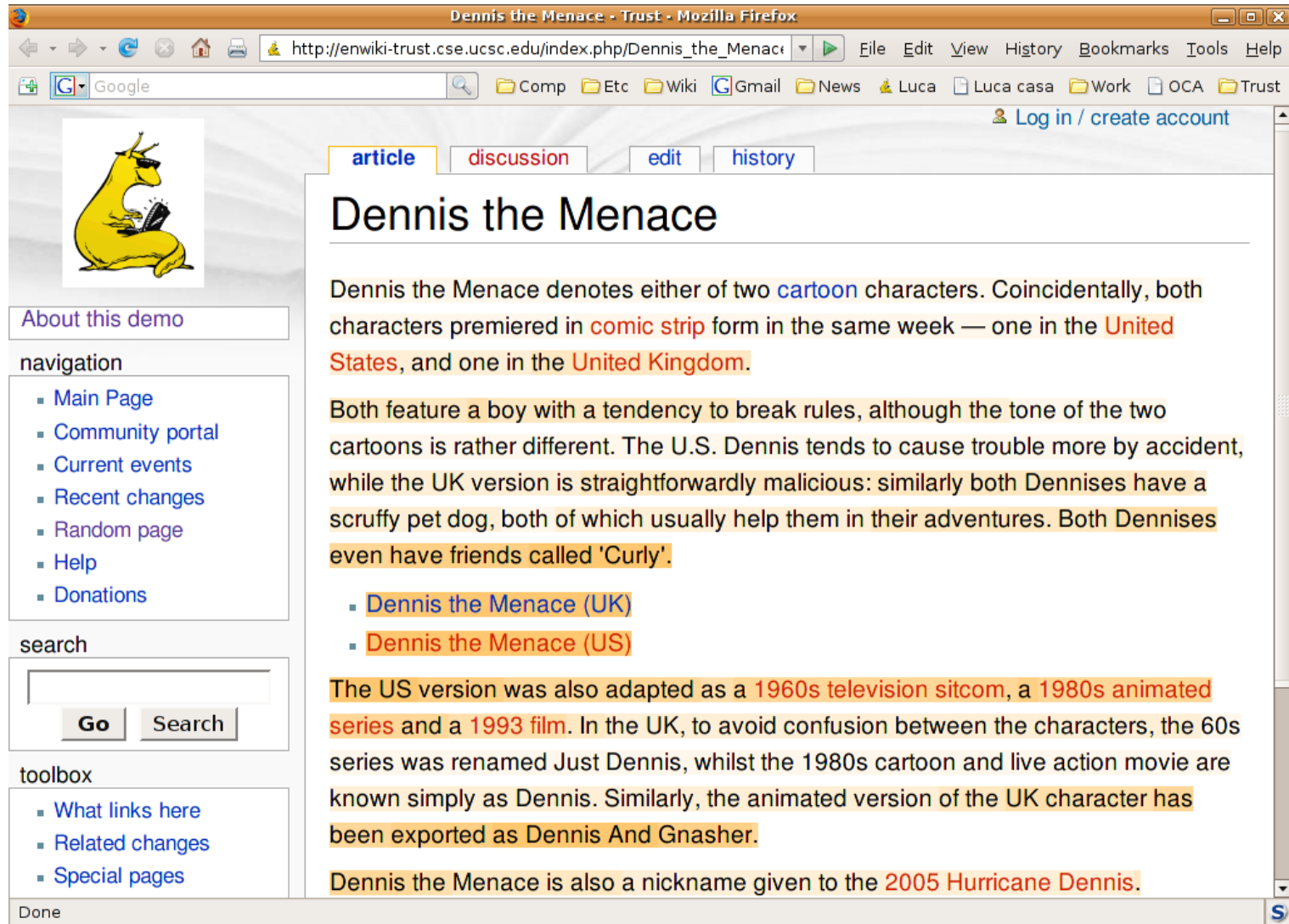


Yadda yadda wuga wuga | yak yak yuk | bla bla bla bing bong

- On the English Wikipedia, we should be able to spot untrusted content with over 80% recall and 60% precision!
  - In fact, we do even better than this, as new content is always flagged lower trust (see next).



# Demo: <http://trust.cse.ucsc.edu/>



The screenshot shows a Mozilla Firefox browser window with the title "Dennis the Menace - Trust - Mozilla Firefox". The address bar contains the URL "http://enwiki-trust.cse.ucsc.edu/index.php/Dennis\_the\_Menace". The browser's menu bar includes "File", "Edit", "View", "History", "Bookmarks", "Tools", and "Help". The toolbar shows icons for "Google", "Comp", "Etc", "Wiki", "Gmail", "News", "Luca", "Luca casa", "Work", "OCA", and "Trust". A "Log in / create account" link is visible in the top right corner.

The main content area displays the article "Dennis the Menace" with tabs for "article", "discussion", "edit", and "history". The article text is as follows:

Dennis the Menace denotes either of two **cartoon** characters. Coincidentally, both characters premiered in **comic strip** form in the same week — one in the **United States**, and one in the **United Kingdom**.

Both feature a boy with a tendency to break rules, although the tone of the two cartoons is rather different. The U.S. Dennis tends to cause trouble more by accident, while the UK version is straightforwardly malicious: similarly both Dennises have a scruffy pet dog, both of which usually help them in their adventures. Both Dennises even have friends called 'Curly'.

- [Dennis the Menace \(UK\)](#)
- [Dennis the Menace \(US\)](#)

The US version was also adapted as a **1960s television sitcom**, a **1980s animated series** and a **1993 film**. In the UK, to avoid confusion between the characters, the 60s series was renamed Just Dennis, whilst the 1980s cartoon and live action movie are known simply as Dennis. Similarly, the animated version of the UK character has been exported as **Dennis And Gnasher**.

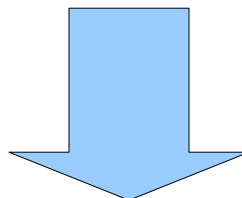
Dennis the Menace is also a nickname given to the **2005 Hurricane Dennis**.

The left sidebar contains a navigation menu with links to "Main Page", "Community portal", "Current events", "Recent changes", "Random page", "Help", and "Donations". Below the navigation menu is a search box with "Go" and "Search" buttons. At the bottom of the sidebar is a "toolbox" with links to "What links here", "Related changes", and "Special pages".

# Text trust: How is "Fogh" spelled?

---

occasionally recruited from outside the Folketing.  
Since 27 November 2001, the economist Anders Fogh Rasmussen has been Prime Minister to Denmark.  
As known in other parliamentary systems of government,



occasionally recruited from outside the Folketing.  
Since 27 November 2001, the economist Anders Fjogh Rasmussen has been Prime Minister to Denmark.  
As known in other parliamentary systems of government,

# Text Trust: more examples from the demo

---

## Consul

In 63 BC, Cicero was elected consul. Nice! His most significant accomplishment during his year in office was the suppression of the supposed Catiline conspiracy, generally

ensures a focus on public-sector efficiency as well as devolved responsibilities of local government on regional and municipal levels. 897678678

# Text Trust: Details

---

## Trust depends on:

- **Authorship:** Author lends 50% of their reputation to the text they create.
  - Thus, even text from high-rep authors is only medium-rep when added: high trust is achieved only via multiple reviews, never via a single author.
- **Revision:** When an author of reputation  $r$  preserves a word of trust  $t < r$ , the word increases in trust to
$$t + 0.3(r - t)$$
- The algorithms still need fine-tuning.

## Executive branch

---

Denmark is a constitutional monarchy with an almost unbroken link of monarchs for more than 1,000 years (except for an interregnum of eight years from 1332 to 1340). The current monarch, Queen Margaret II, has largely ceremonial functions; perhaps her most significant formal power lies in her right to appoint the prime minister (Statsminister) and the cabinet of Denmark, who are responsible for administration of the government. However, she must consult with parliamentary leaders to determine the public's will, since the cabinet may be dismissed by a vote of no confidence in the Folketing, the Danish parliament. Cabinet members are occasionally recruited from outside the Folketing. The vote of no confidence is an essential part of danish parliamentary work, since the constitution states "no government can work with a majority against it", opposed to the more common international rule that says government need a majority for it. The rule can seem to have the same meaning, but in politics nothing is absolute, and as proof of this the socialdemocrats has once used this rule to form goverment although the liberal-conservative block was of greater number.

## Executive branch

---

Denmark is a constitutional monarchy with an almost unbroken link of monarchs for more than 1,000 years (except for an interregnum of eight years from 1332 to 1340). The current monarch, Queen Margaret II, has largely ceremonial functions; perhaps her most significant formal power lies in her right to appoint the prime minister (Statsminister) and the cabinet of Denmark, who are responsible for administration of the government. However, she must consult with parliamentary leaders to determine the public's will, since the cabinet may be dismissed by a vote of no confidence in the Folketing, the Danish parliament. Cabinet members are occasionally recruited from outside the Folketing. The vote of no confidence is an essential part of danish parliamentary work, since the constitution states "no government can work with a majority against it", opposed to the more common international rule that says a government needs a majority for it. These rules may seem to have the same meaning, but in politics nothing is absolute, and as proof of this the Social Democrats once used this rule to form government although the liberal-conservative block was of greater number.

## Executive branch

---

Denmark is a constitutional monarchy with an almost unbroken link of monarchs for more than 1,000 years (except for an interregnum of eight years from 1332 to 1340). The current monarch, Queen Margaret II, has largely ceremonial functions; perhaps her most significant formal power lies in her right to appoint the prime minister (Statsminister) and the cabinet of Denmark, who are responsible for administration of the government. However, she must consult with parliamentary leaders to determine the public's will, since the cabinet may be dismissed by a vote of no confidence in the Folketing, the Danish parliament. Cabinet members are occasionally recruited from outside the Folketing. The vote of no confidence is an essential part of Danish parliamentary work, since the constitution states "no government can work with a majority against it", opposed to the more common international rule that says a government needs a majority for it. These rules may seem to have the same meaning, but in politics nothing is absolute, and as proof of this the Social Democrats once used this rule to form government although the liberal-conservative block was of greater number.

## Executive branch

---

Denmark is a constitutional monarchy with an almost unbroken link of monarchs for more than 1,000 years (except for an interregnum of eight years from 1332 to 1340). The current monarch, Queen Margaret II, has largely ceremonial functions; perhaps her most significant formal power lies in her right to appoint the prime minister (Statsminister) and the cabinet of Denmark, who are responsible for administration of the government. However, she must consult with parliamentary leaders to determine the public's will, since the cabinet may be dismissed by a vote of no confidence in the Folketing, the Danish parliament. Cabinet members are occasionally recruited from outside the Folketing. The vote of no confidence is an essential part of danish parliamentary work, since the constitution states "no government can work with a majority against it", opposed to the more common international rule that says a government needs a majority for it. These rules may seem to have the same meaning, but in politics nothing is absolute, and as proof of this the Social Democrats once used this rule to form government although the liberal-conservative block was of greater number.



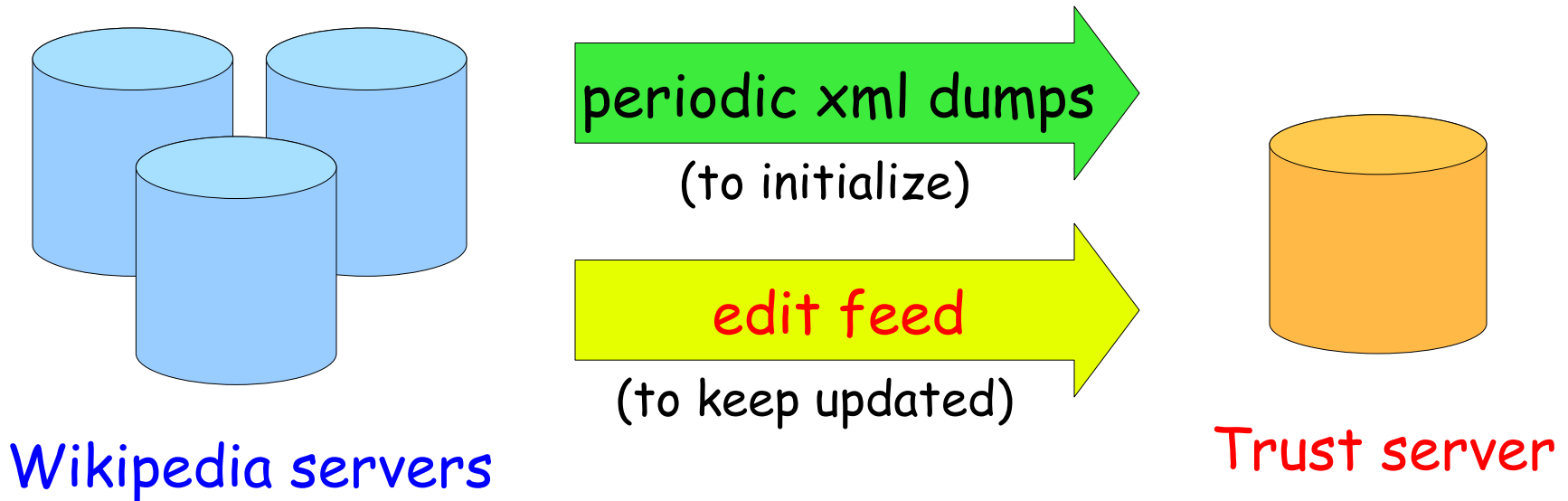
## Executive branch

---

Denmark is a constitutional monarchy with an almost unbroken link of monarchs for more than 1,000 years (except for an interregnum of eight years from 1332 to 1340). The current monarch, Queen Margaret II, has largely ceremonial functions; perhaps her most significant formal power lies in her right to appoint the prime minister (Statsminister) and the cabinet of Denmark, who are responsible for administration of the government. However, she must consult with parliamentary leaders to determine the public's will, since the cabinet may be dismissed by a vote of no confidence in the Folketing, the Danish parliament. Cabinet members are occasionally recruited from outside the Folketing. The vote of no confidence is an essential part of Danish parliamentary work, since the constitution states "no government can work with a majority against it", opposed to the more common international rule that says a government needs a majority for it. These rules may seem to have the same meaning, but in politics nothing is absolute, and as proof of this the Social Democrats once used this rule to form government although the liberal-conservative block was of greater number.

# Batch Implementation

---



- No need to affect the main Wikipedia servers
- People can click "check trust" and visit the trust server.
- Good for experimenting with new ideas
- Necessary to color the past (come up to speed).

# On-Line Implementation

---

## Process edits as they arrive:

- Benefit: real-time colorization of text
- Need to integrate the code in MediaWiki
- Time to process an edit:  $< 1s$  (not much longer than parsing it).
- Storage required: proportional to the size of the last revision (*not to the total history size!*)
- Can be easily used for other Wikis

## My questions:

- Feedback?
- Do you like it?
- Should we try to set up a “trust server” with an edit feed from the Wikipedia?
- Try the demo:

<http://trust.cse.ucsc.edu/>

## Your questions?