Improving Search Through *Efficient* A/B Testing:

A Case Study

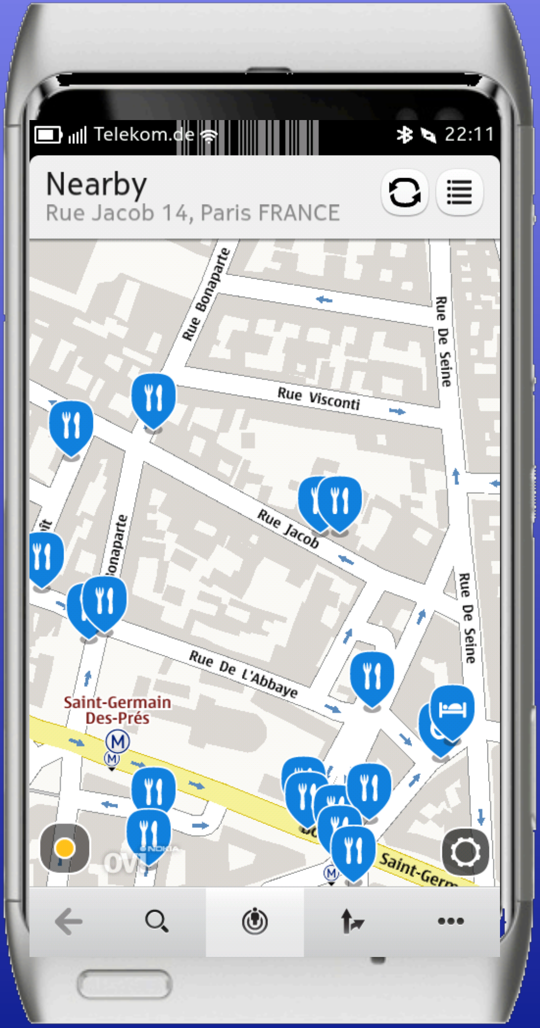
Nokia Maps "Place Discovery" Team, Berlin:

Hannes Kruppa, Steffen Bickel, Mark Waldaukat, Felix Weigel, Ross Turner, Peter Siemen

NOKIA

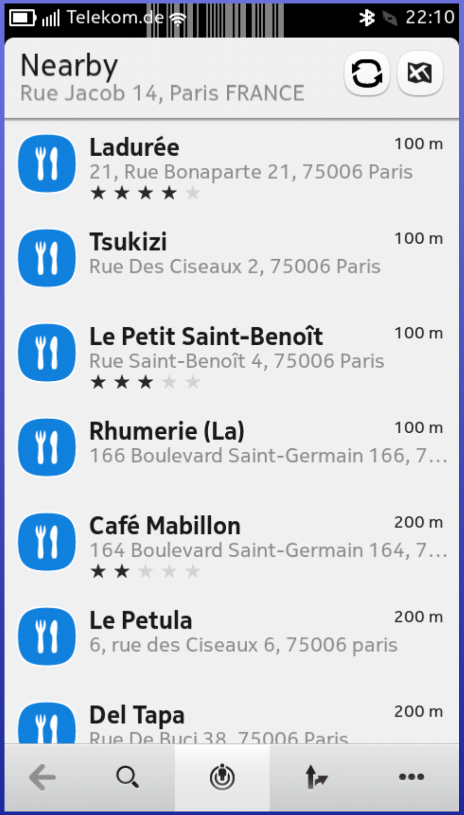
Nokia Maps: *Nearby Places*

"Discover Places You Will Love, Anywhere"



Telekom.de 22:11

Nearby
Rue Jacob 14, Paris FRANCE



Telekom.de 22:10

Nearby
Rue Jacob 14, Paris FRANCE

Ladurée 100 m
21, Rue Bonaparte 21, 75006 Paris
★★★★☆

Tsukizi 100 m
Rue Des Ciseaux 2, 75006 Paris

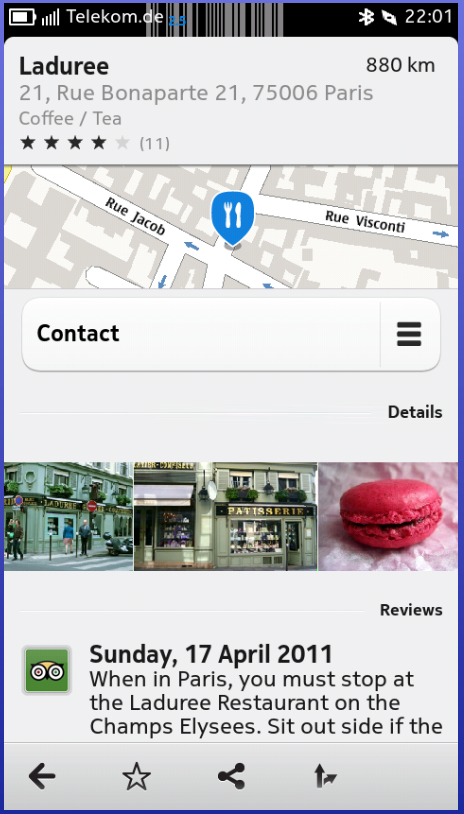
Le Petit Saint-Benoît 100 m
Rue Saint-Benoît 4, 75006 Paris
★★★☆☆

Rhumerie (La) 100 m
166 Boulevard Saint-Germain 166, 7…

Café Mabillon 200 m
164 Boulevard Saint-Germain 164, 7…
★★☆☆☆

Le Petula 200 m
6, rue des Ciseaux 6, 75006 paris

Del Tapa 200 m
Rue De Buci 38, 75006 Paris



Telekom.de 22:01

Laduree 880 km
21, Rue Bonaparte 21, 75006 Paris
Coffee / Tea
★★★★☆ (11)

Contact

Details

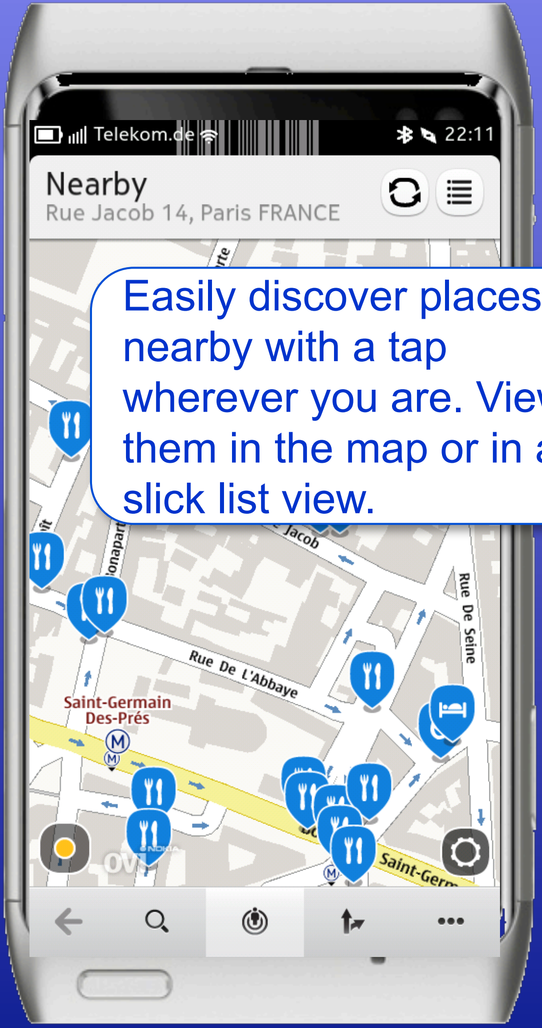
Reviews

Sunday, 17 April 2011
When in Paris, you must stop at the Laduree Restaurant on the Champs Elysees. Sit out side if the

NOKIA

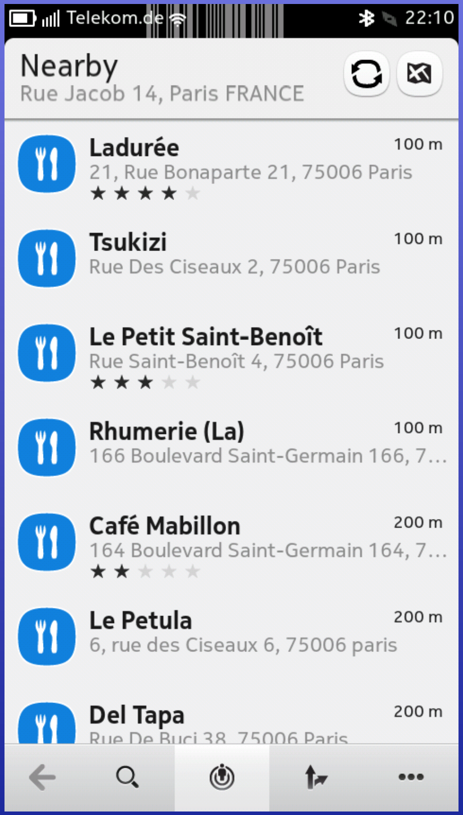
Nokia Maps: *Nearby Places*

"Discover Places You Will Love, Anywhere"



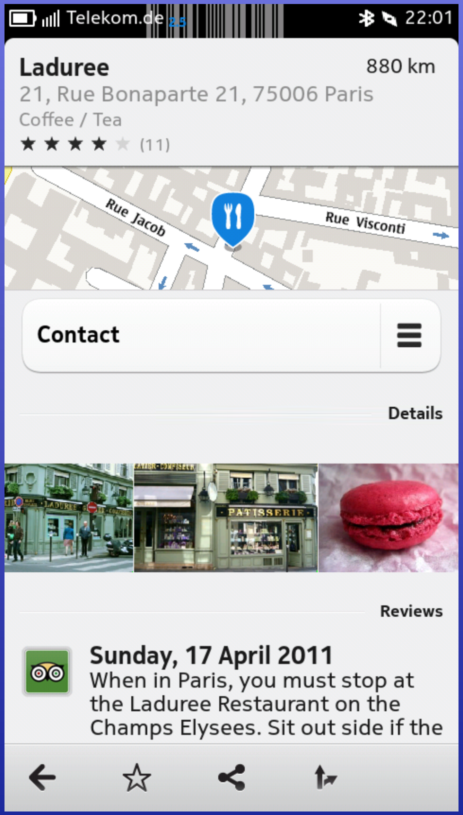
Nearby
Rue Jacob 14, Paris FRANCE

Easily discover places nearby with a tap wherever you are. View them in the map or in a slick list view.



Nearby
Rue Jacob 14, Paris FRANCE

Ladurée 21, Rue Bonaparte 21, 75006 Paris ★★★★☆	100 m
Tsukizi Rue Des Ciseaux 2, 75006 Paris	100 m
Le Petit Saint-Benoît Rue Saint-Benoît 4, 75006 Paris ★★★☆☆	100 m
Rhumerie (La) 166 Boulevard Saint-Germain 166, 7…	100 m
Café Mabillon 164 Boulevard Saint-Germain 164, 7… ★★☆☆☆	200 m
Le Petula 6, rue des Ciseaux 6, 75006 paris	200 m
Del Tapa Rue De Buci 38, 75006 Paris	200 m



Laduree 880 km
21, Rue Bonaparte 21, 75006 Paris
Coffee / Tea
★★★★☆ (11)

Contact

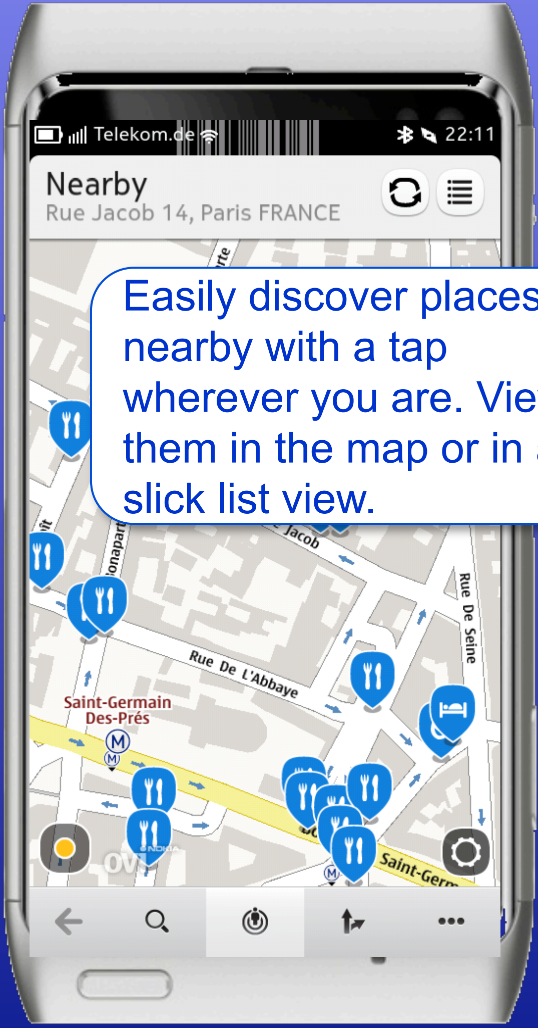
Details

Reviews

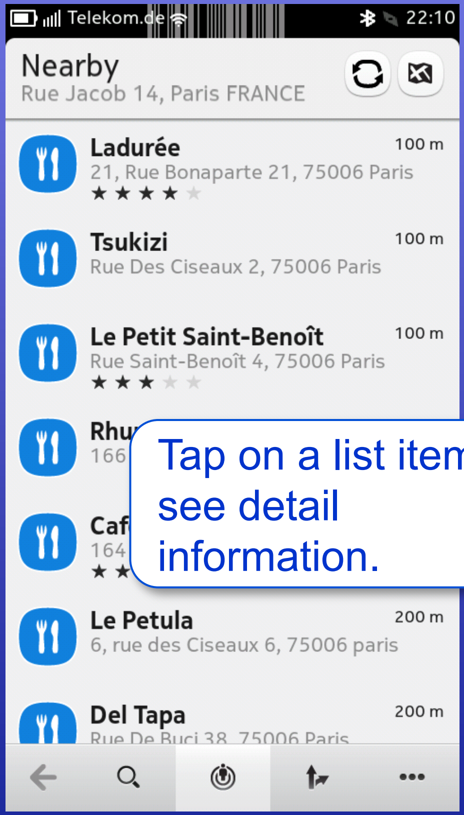
Sunday, 17 April 2011
When in Paris, you must stop at the Laduree Restaurant on the Champs Elysees. Sit out side if the

Nokia Maps: *Nearby Places*

"Discover Places You Will Love, Anywhere"



Easily discover places nearby with a tap wherever you are. View them in the map or in a slick list view.



Nearby
Rue Jacob 14, Paris FRANCE

Ladurée 100 m
21, Rue Bonaparte 21, 75006 Paris
★★★★☆

Tsukizi 100 m
Rue Des Ciseaux 2, 75006 Paris

Le Petit Saint-Benoît 100 m
Rue Saint-Benoît 4, 75006 Paris
★★★☆☆

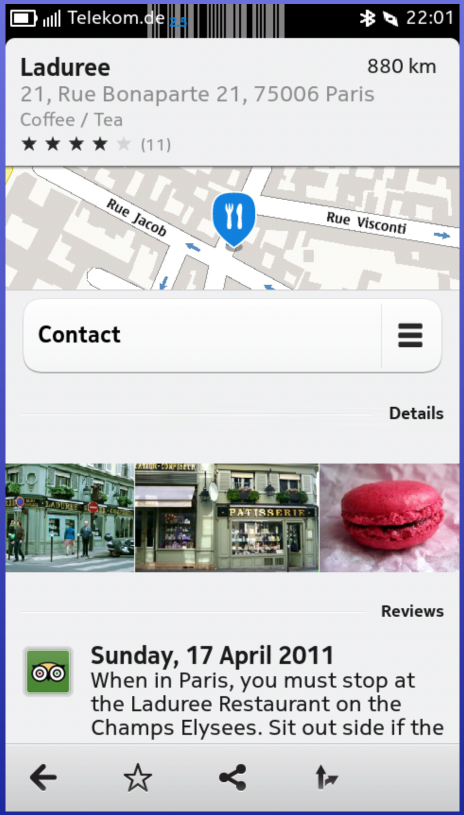
Rhu
166

Caf
164
★★

Le Petula 200 m
6, rue des Ciseaux 6, 75006 paris

Del Tapa 200 m
Rue De Buci 38, 75006 Paris

Tap on a list item to see detail information.



Laduree 880 km
21, Rue Bonaparte 21, 75006 Paris
Coffee / Tea
★★★★☆ (11)

Contact

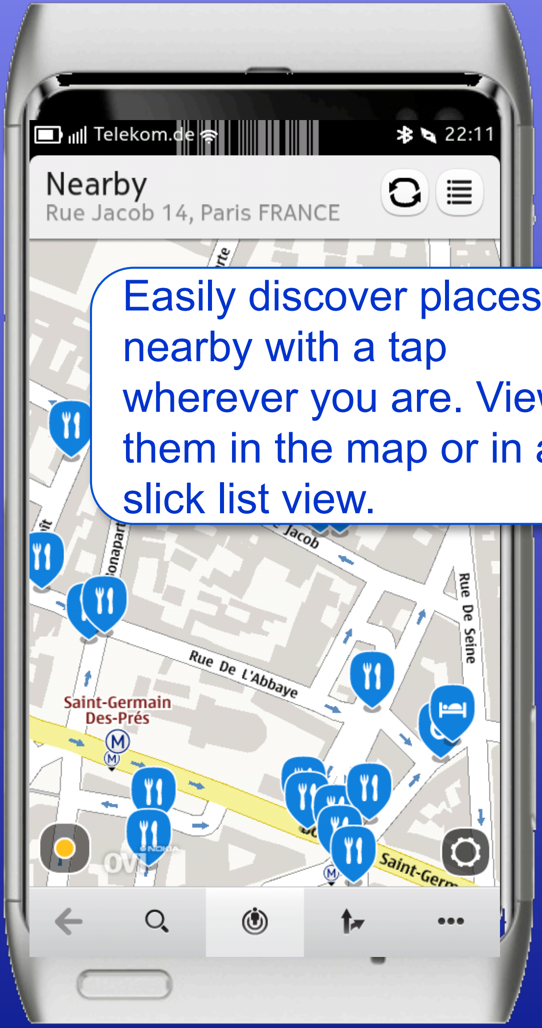
Details

Reviews

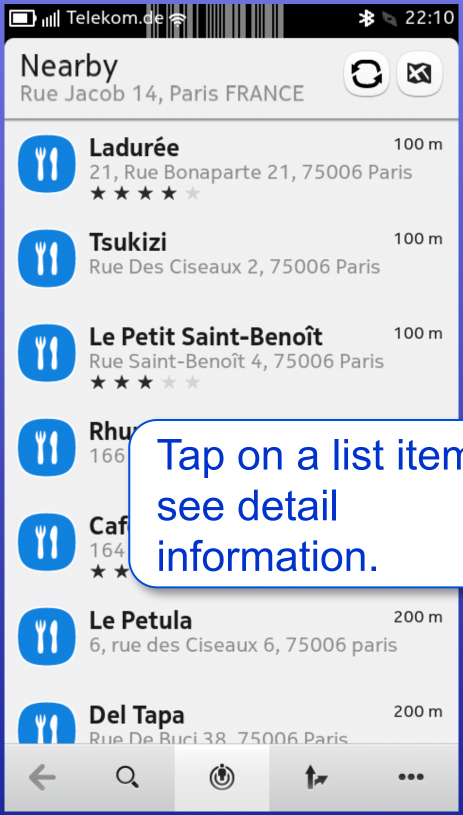
Sunday, 17 April 2011
When in Paris, you must stop at the Laduree Restaurant on the Champs Elysees. Sit out side if the

Nokia Maps: *Nearby Places*

"Discover Places You Will Love, Anywhere"

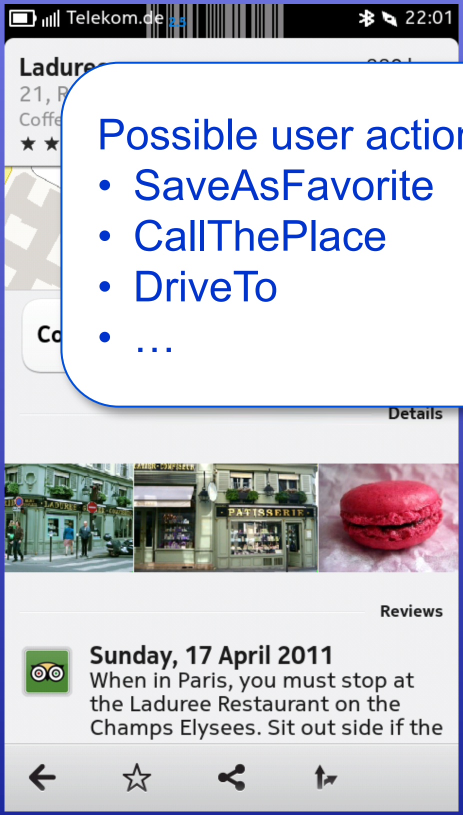


Easily discover places nearby with a tap wherever you are. View them in the map or in a slick list view.



Ladurée 21, Rue Bonaparte 21, 75006 Paris ★★★★☆	100 m
Tsukizi Rue Des Ciseaux 2, 75006 Paris	100 m
Le Petit Saint-Benoît Rue Saint-Benoît 4, 75006 Paris ★★★☆☆	100 m
Rhu 166	
Caf 164 ★★	
Le Petula 6, rue des Ciseaux 6, 75006 paris	200 m
Del Tapa Rue De Buci 38, 75006 Paris	200 m

Tap on a list item to see detail information.



Sunday, 17 April 2011

When in Paris, you must stop at the Laduree Restaurant on the Champs Elysees. Sit out side if the

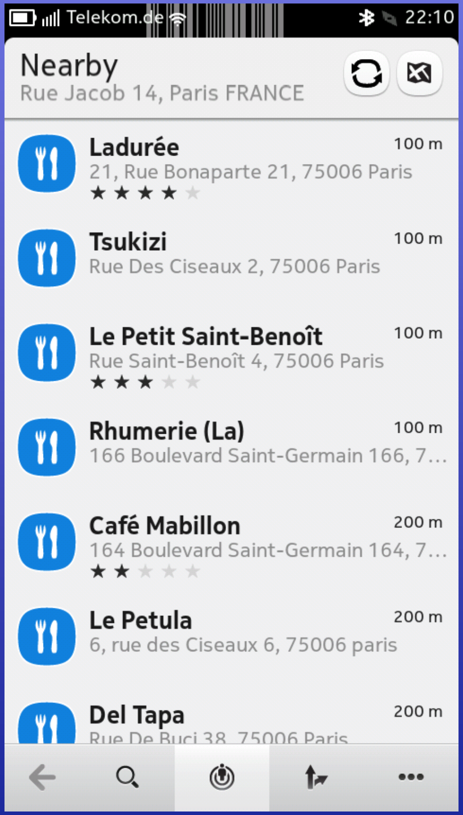
Possible user actions:

- SaveAsFavorite
- CallThePlace
- DriveTo
- …

NOKIA

Problem: Which Places to Show?

- Restaurants? Hotels? Shopping? …
- rank by Ratings?
- Distance?
- Usage?
- Trending?
- ….



Telekom.de 22:10

Nearby
Rue Jacob 14, Paris FRANCE

Ladurée 100 m
21, Rue Bonaparte 21, 75006 Paris
★★★★☆

Tsukizi 100 m
Rue Des Ciseaux 2, 75006 Paris

Le Petit Saint-Benoît 100 m
Rue Saint-Benoît 4, 75006 Paris
★★★☆☆

Rhumerie (La) 100 m
166 Boulevard Saint-Germain 166, 7…

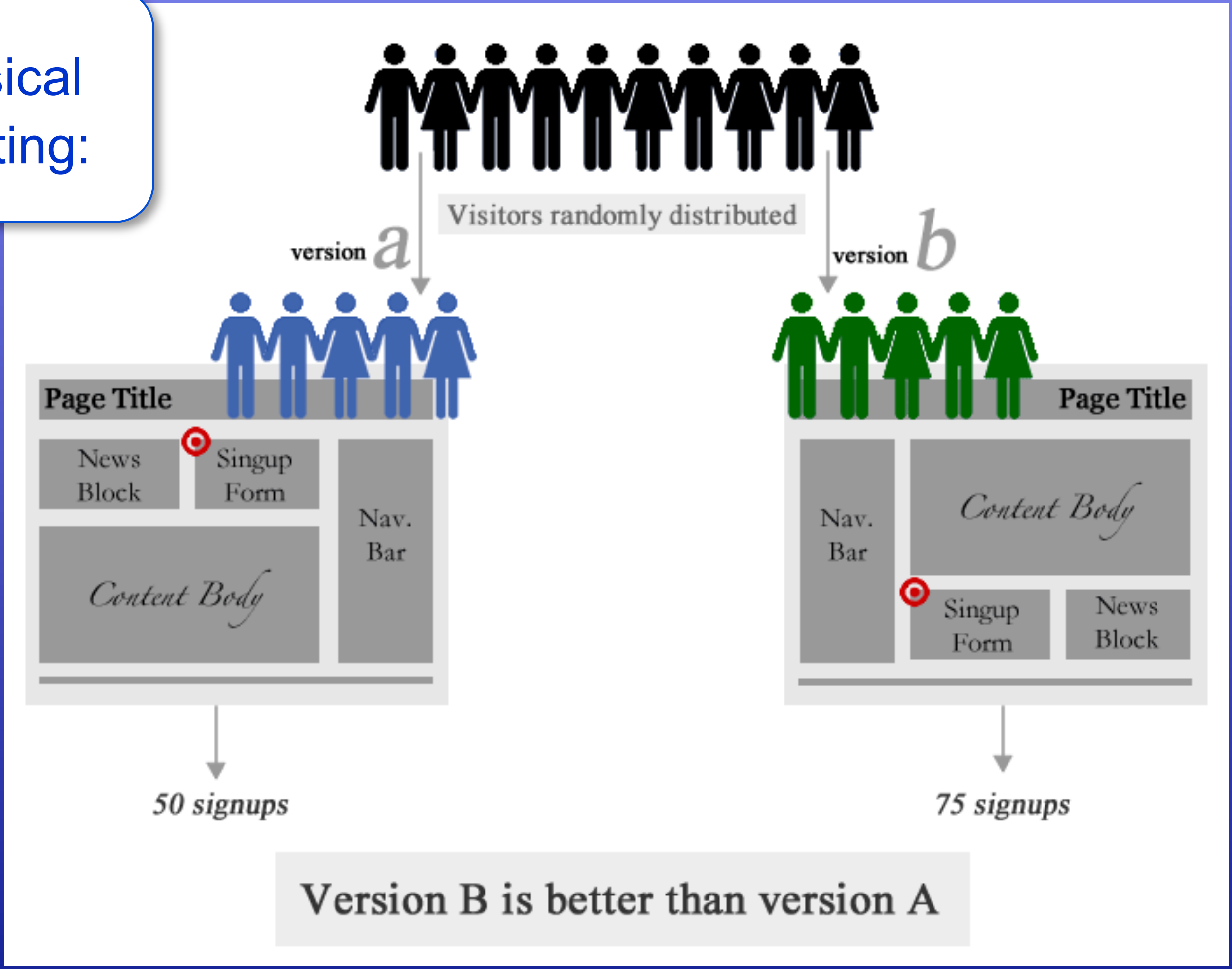
Café Mabillon 200 m
164 Boulevard Saint-Germain 164, 7…
★★☆☆☆

Le Petula 200 m
6, rue des Ciseaux 6, 75006 paris

Del Tapa 200 m
Rue De Buci 38, 75006 Paris

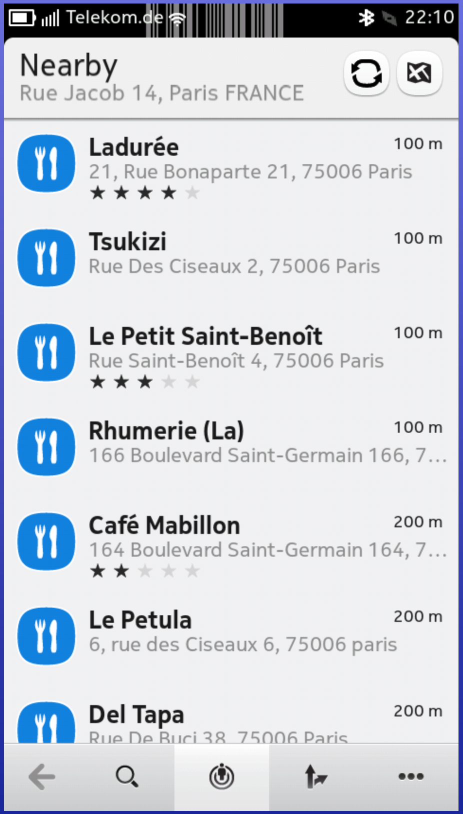
Approach: A/B-Test Different Versions!

Here is classical Web A/B testing:



A/B-Test for *Nearby Places*

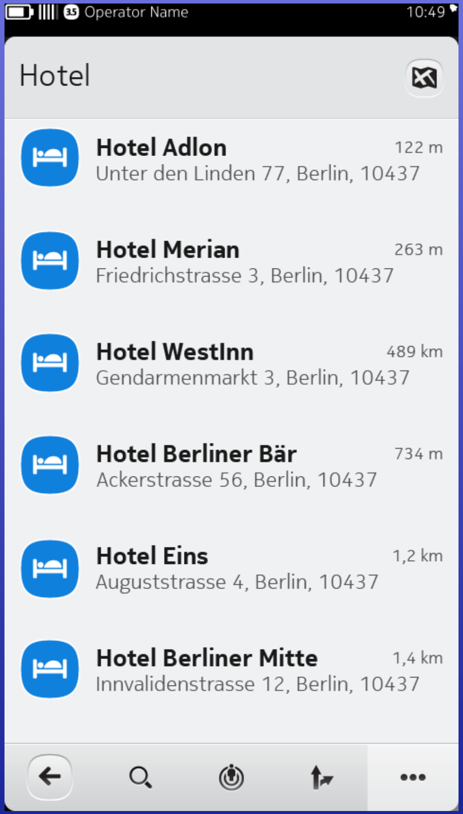
Version A:
Best of Eat'n'Drink



Nearby
Rue Jacob 14, Paris FRANCE

- **Ladurée** 100 m
  21, Rue Bonaparte 21, 75006 Paris
  ★★★★☆
- **Tsukizi** 100 m
  Rue Des Ciseaux 2, 75006 Paris
- **Le Petit Saint-Benoît** 100 m
  Rue Saint-Benoît 4, 75006 Paris
  ★★★☆☆
- **Rhumerie (La)** 100 m
  166 Boulevard Saint-Germain 166, 7…
- **Café Mabillon** 200 m
  164 Boulevard Saint-Germain 164, 7…
  ★★☆☆☆
- **Le Petula** 200 m
  6, rue des Ciseaux 6, 75006 paris
- **Del Tapa** 200 m
  Rue De Buci 38, 75006 Paris

Version B:
Best of Hotels



Hotel

- **Hotel Adlon** 122 m
  Unter den Linden 77, Berlin, 10437
- **Hotel Merian** 263 m
  Friedrichstrasse 3, Berlin, 10437
- **Hotel WestInn** 489 km
  Gendarmenmarkt 3, Berlin, 10437
- **Hotel Berliner Bär** 734 m
  Ackerstrasse 56, Berlin, 10437
- **Hotel Eins** 1,2 km
  Auguststrasse 4, Berlin, 10437
- **Hotel Berliner Mitte** 1,4 km
  Innvalidenstrasse 12, Berlin, 10437

Versions Compete for User engagement:

= Number of Actions performed on places.

NOKIA

There Is A Better Approach For Ranked Lists

[Joachims et al 2008]:

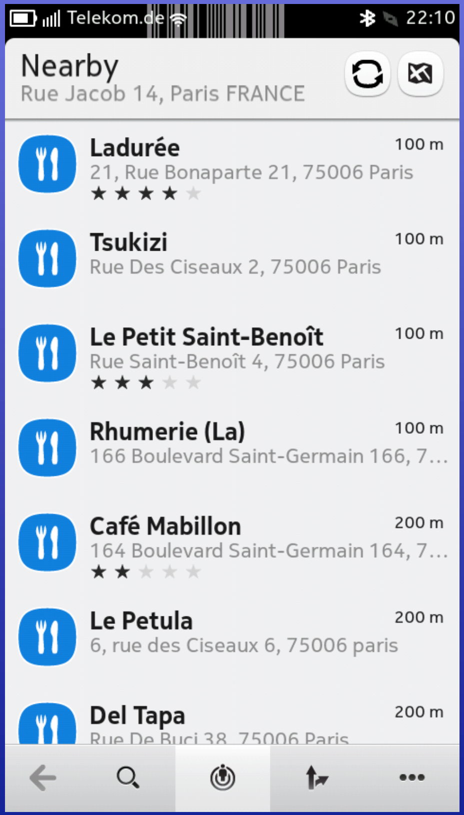
“How Does Clickthrough Data Reflect Retrieval Quality?”

- Classical A/B testing converges slowly for ranked lists
- Classical A/B testing often doesn’t reflect actual relevance

- A/B Tests for Ranked Result Lists: Rank- Interleaving
- Use Rank-Interleaving for faster statistical significance

Efficient A/B Testing: *Rank Interleaving*

Version A:
Best of Eat'n'Drink



Telekom.de 22:10

Nearby
Rue Jacob 14, Paris FRANCE

Ladurée 100 m
21, Rue Bonaparte 21, 75006 Paris
★★★★☆

Tsukizi 100 m
Rue Des Ciseaux 2, 75006 Paris

Le Petit Saint-Benoît 100 m
Rue Saint-Benoît 4, 75006 Paris
★★★☆☆

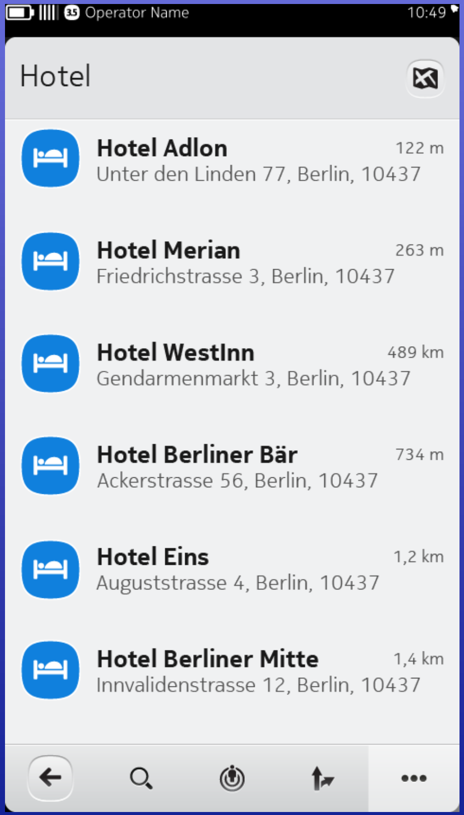
Rhumerie (La) 100 m
166 Boulevard Saint-Germain 166, 7…

Café Mabillon 200 m
164 Boulevard Saint-Germain 164, 7…
★★☆☆☆

Le Petula 200 m
6, rue des Ciseaux 6, 75006 paris

Del Tapa 200 m
Rue De Buci 38, 75006 Paris

Version B:
Best of Hotels



Operator Name 10:49

Hotel

Hotel Adlon 122 m
Unter den Linden 77, Berlin, 10437

Hotel Merian 263 m
Friedrichstrasse 3, Berlin, 10437

Hotel WestInn 489 km
Gendarmenmarkt 3, Berlin, 10437

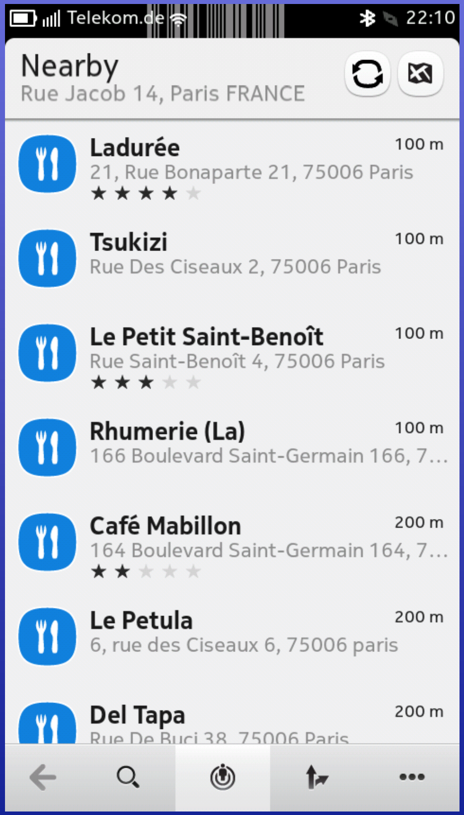
Hotel Berliner Bär 734 m
Ackerstrasse 56, Berlin, 10437

Hotel Eins 1,2 km
Auguststrasse 4, Berlin, 10437

Hotel Berliner Mitte 1,4 km
Invalidenstrasse 12, Berlin, 10437

Efficient A/B Testing: *Rank Interleaving*

Version A:
Best of Eat'n'Drink



Telekom.de 22:10

Nearby
Rue Jacob 14, Paris FRANCE

Ladurée 100 m
21, Rue Bonaparte 21, 75006 Paris
★★★★☆

Tsukizi 100 m
Rue Des Ciseaux 2, 75006 Paris

Le Petit Saint-Benoît 100 m
Rue Saint-Benoît 4, 75006 Paris
★★★☆☆

Rhumerie (La) 100 m
166 Boulevard Saint-Germain 166, 7…

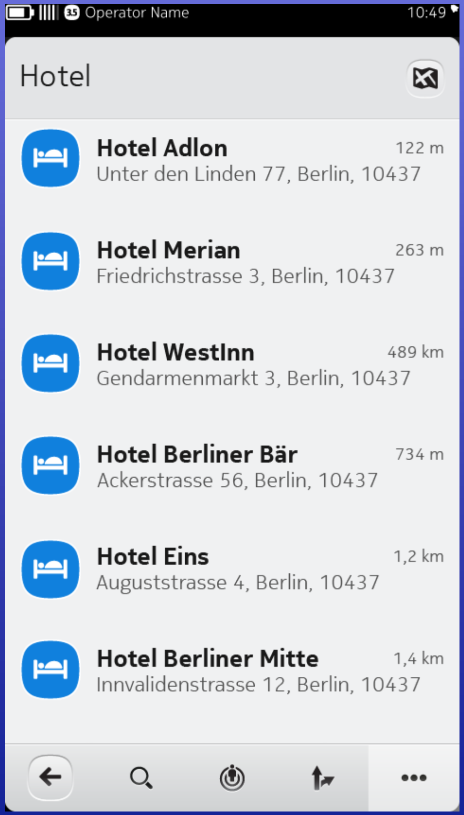
Café Mabillon 200 m
164 Boulevard Saint-Germain 164, 7…
★★☆☆☆

Le Petula 200 m
6, rue des Ciseaux 6, 75006 paris

Del Tapa 200 m
Rue De Buci 38, 75006 Paris

+

Version B:
Best of Hotels



Operator Name 10:49

Hotel

Hotel Adlon 122 m
Unter den Linden 77, Berlin, 10437

Hotel Merian 263 m
Friedrichstrasse 3, Berlin, 10437

Hotel WestInn 489 km
Gendarmenmarkt 3, Berlin, 10437

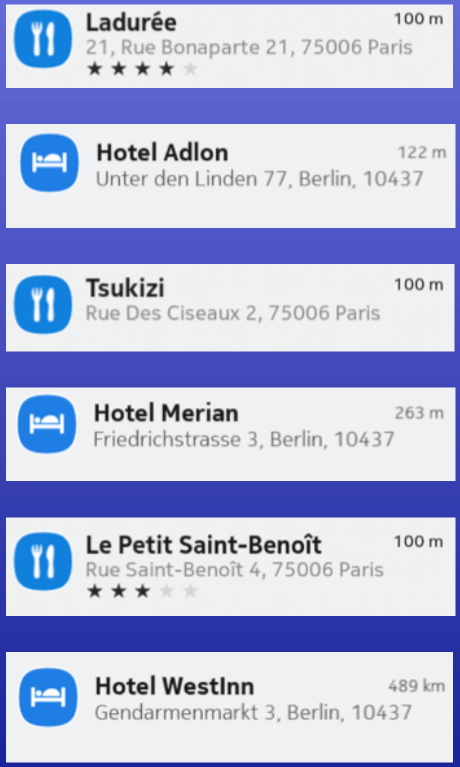
Hotel Berliner Bär 734 m
Ackerstrasse 56, Berlin, 10437

Hotel Eins 1,2 km
Auguststrasse 4, Berlin, 10437

Hotel Berliner Mitte 1,4 km
Invalidenstrasse 12, Berlin, 10437

=

Rank Interleaving:
Version A + B



Ladurée 100 m
21, Rue Bonaparte 21, 75006 Paris
★★★★☆

Hotel Adlon 122 m
Unter den Linden 77, Berlin, 10437

Tsukizi 100 m
Rue Des Ciseaux 2, 75006 Paris

Hotel Merian 263 m
Friedrichstrasse 3, Berlin, 10437

Le Petit Saint-Benoît 100 m
Rue Saint-Benoît 4, 75006 Paris
★★★☆☆

Hotel WestInn 489 km
Gendarmenmarkt 3, Berlin, 10437

NOKIA

Randomized Mixing of Result Lists

- Interleaved list is filled with pairs of results, one item from each version. Coin toss decides who comes first.

Version A

1. *alpha*
2. *beta*
3. *gamma*
4. *delta*
5. *epsilon*

Interleaved Result list

<empty>

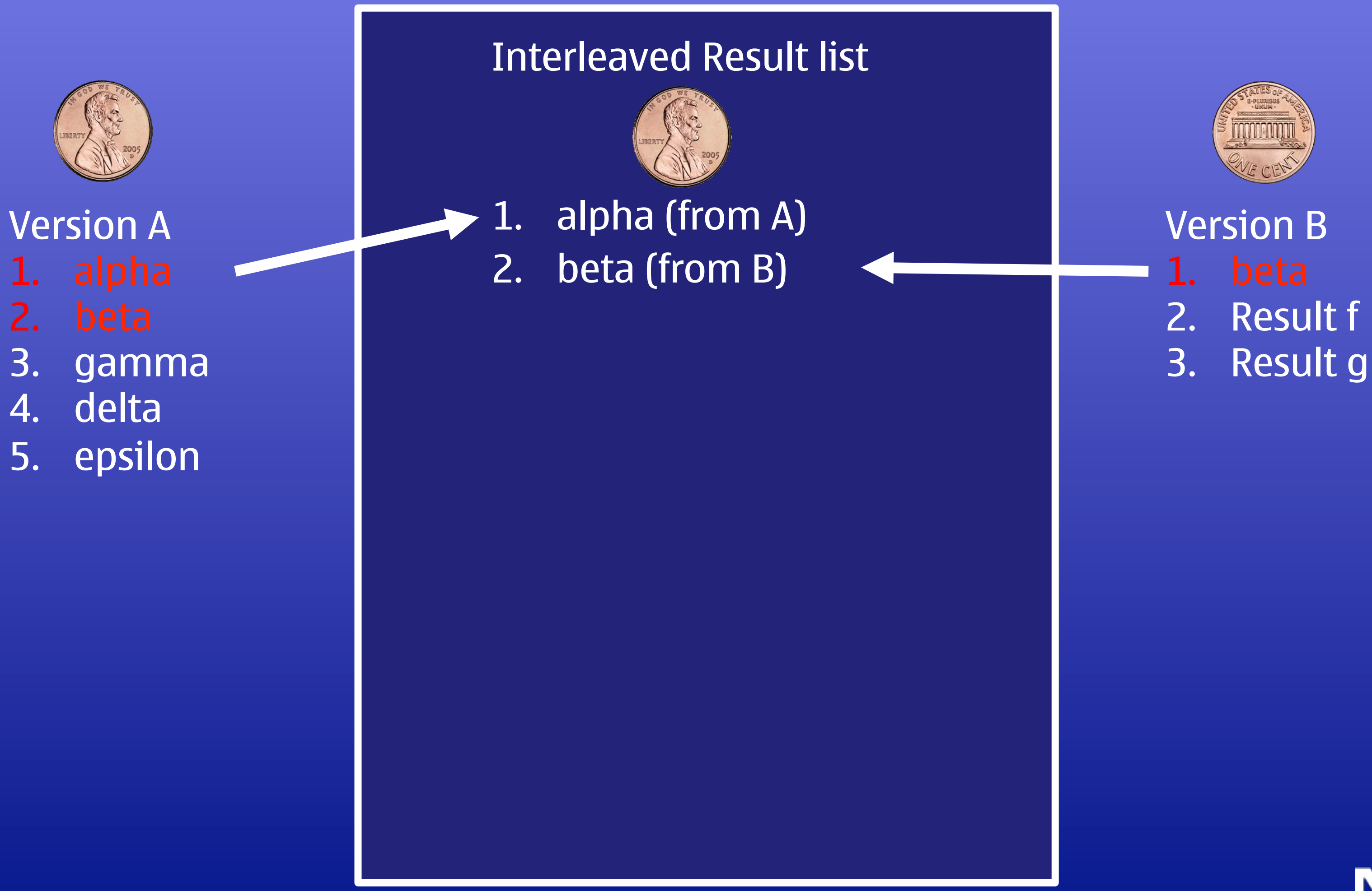
Version B

1. *beta*
2. *kappa*
3. *tau*

NOKIA

A/B Interleaving: Randomized Mixing of Lists

- Interleaved list is filled with pairs of results, one item from each version. Coin toss decides who comes first.



Version A

1. alpha
2. beta
3. gamma
4. delta
5. epsilon

Interleaved Result list

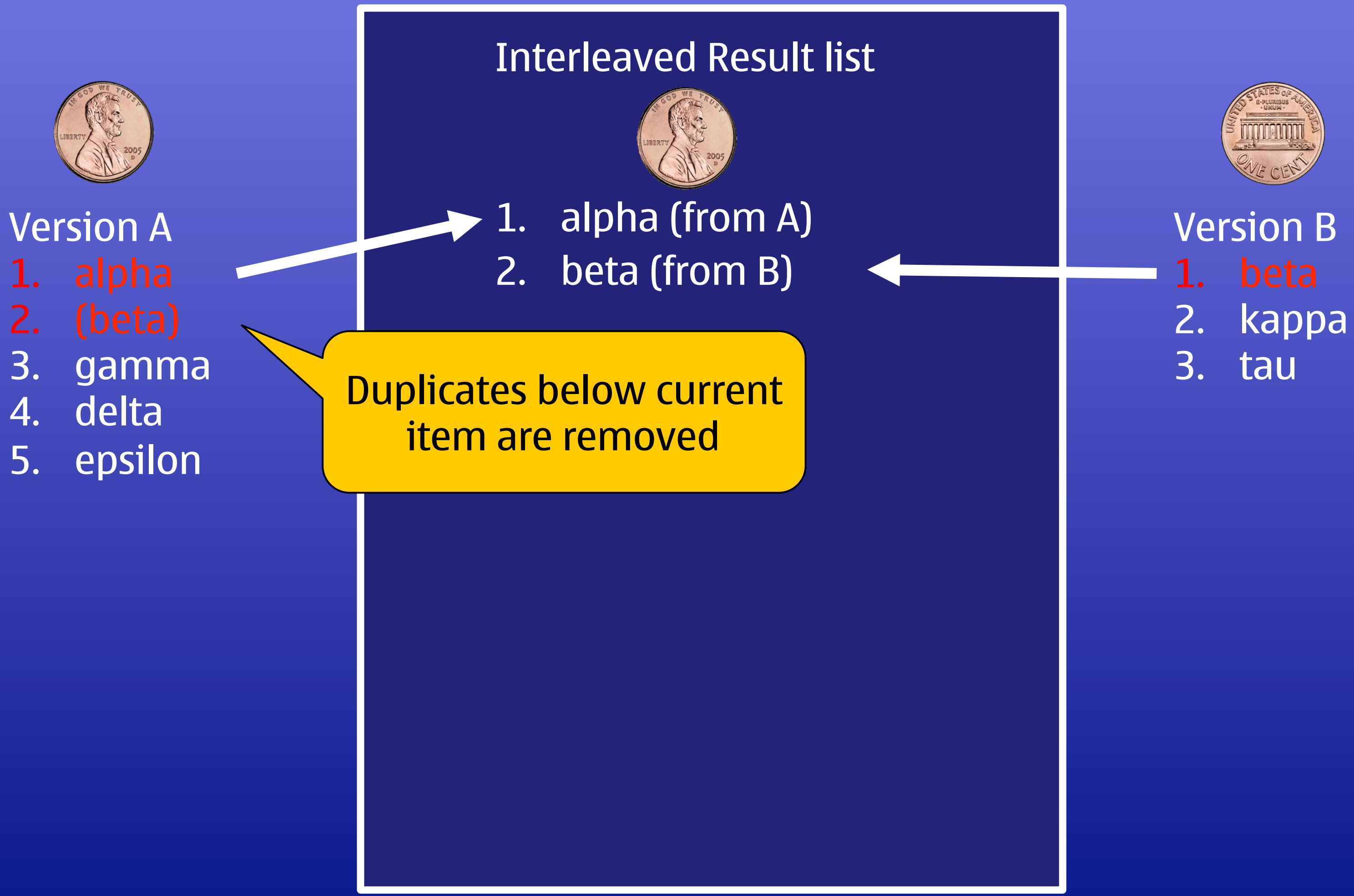
1. alpha (from A)
2. beta (from B)

Version B

1. beta
2. Result f
3. Result g

A/B Interleaving: Randomized Mixing of Lists

- Interleaved list is filled with pairs of results, one item from each version. Coin toss decides who comes first.



Version A

1. alpha
2. (beta)
3. gamma
4. delta
5. epsilon

Interleaved Result list

1. alpha (from A)
2. beta (from B)

Version B

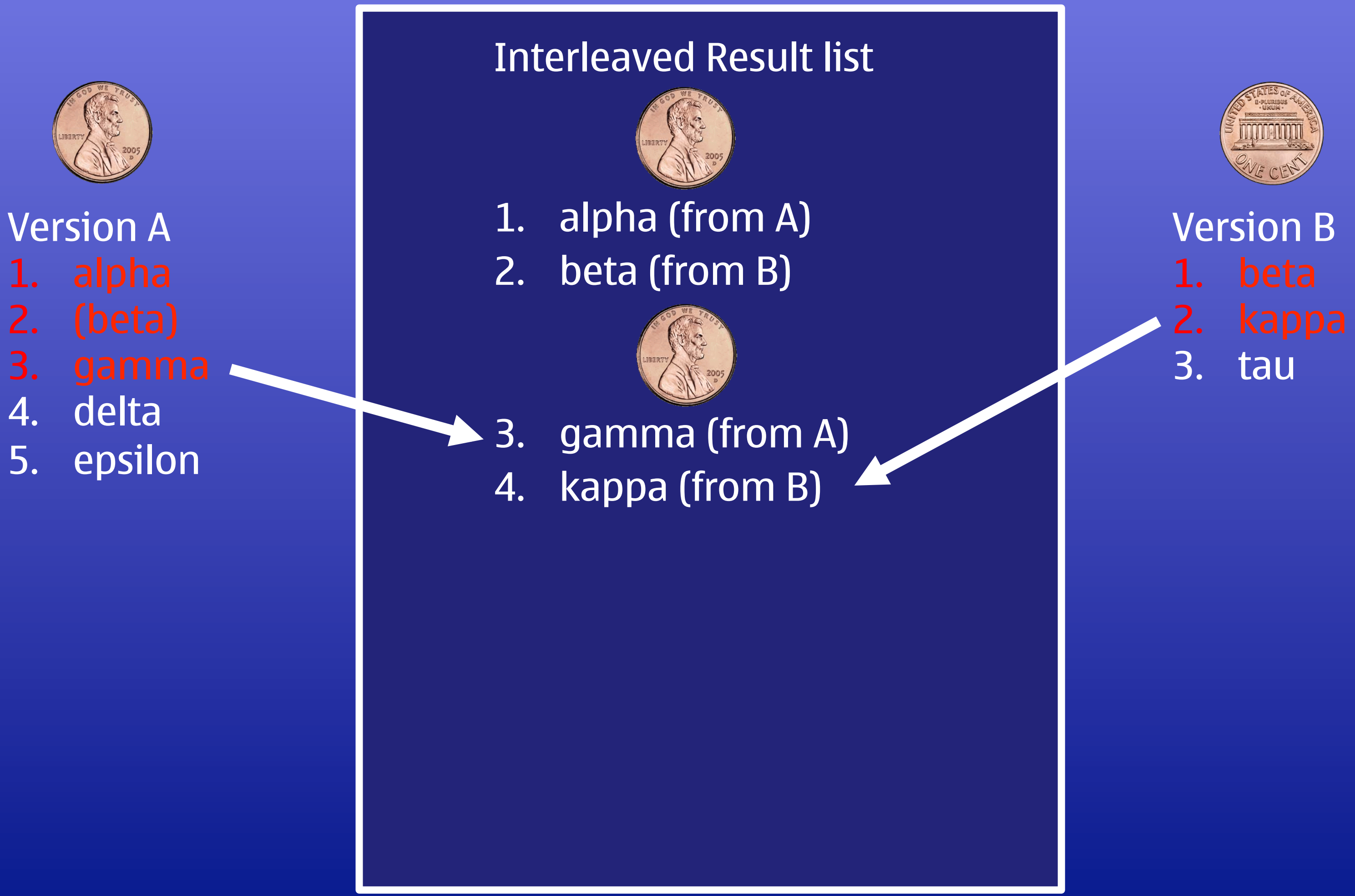
1. beta
2. kappa
3. tau

Duplicates below current item are removed

NOKIA

A/B Interleaving: Randomized Mixing of Lists

- Interleaved list is filled with pairs of results, one item from each version. Coin toss decides who comes first.



Version A

1. alpha
2. (beta)
3. gamma
4. delta
5. epsilon

Interleaved Result list

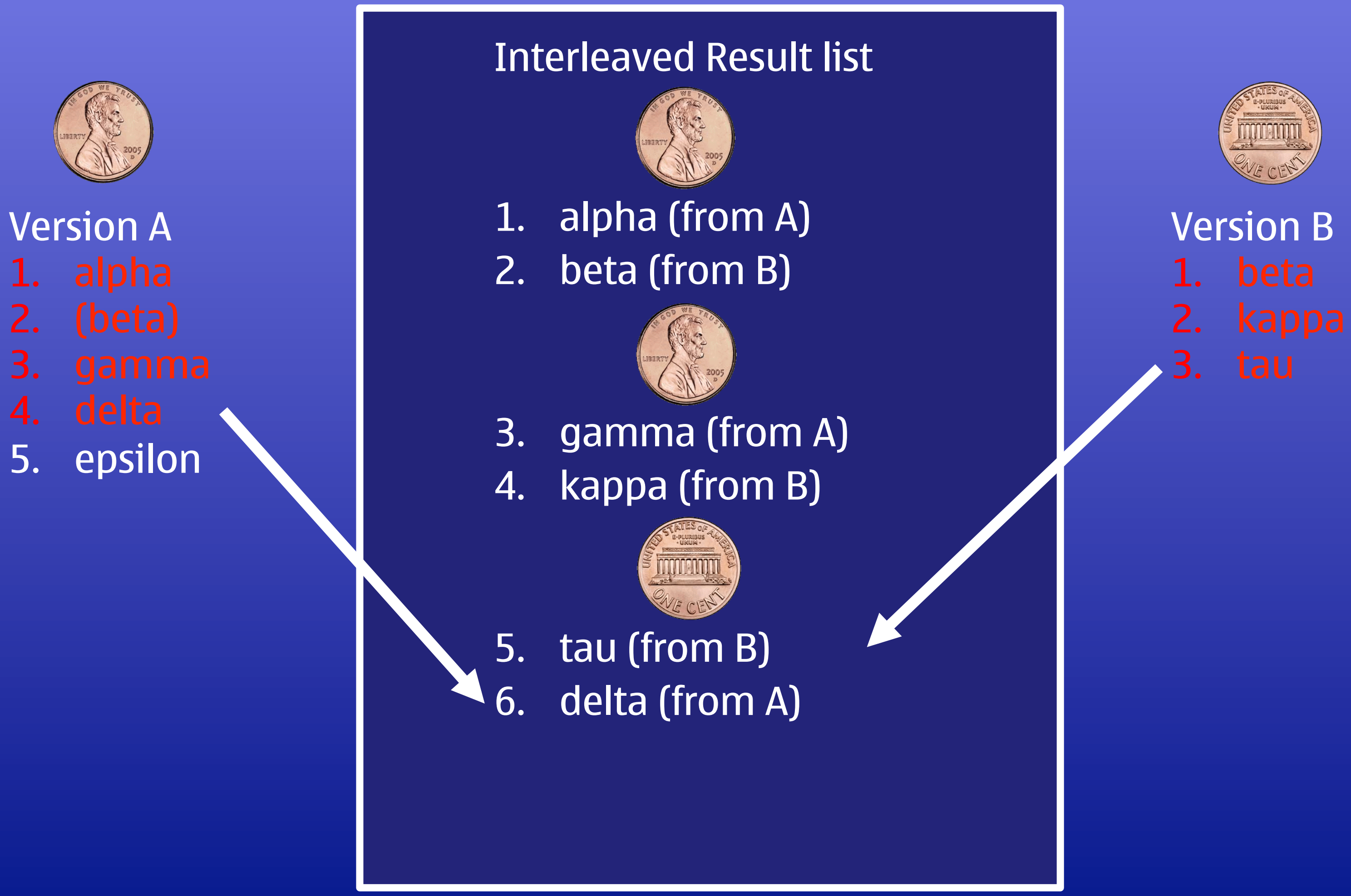
1. alpha (from A)
2. beta (from B)
3. gamma (from A)
4. kappa (from B)

Version B

1. beta
2. kappa
3. tau

A/B Interleaving: Randomized Mixing of Lists

- Interleaved list is filled with pairs of results, one item from each version. Coin toss decides who comes first.



Version A

1. alpha
2. (beta)
3. gamma
4. delta
5. epsilon

Interleaved Result list

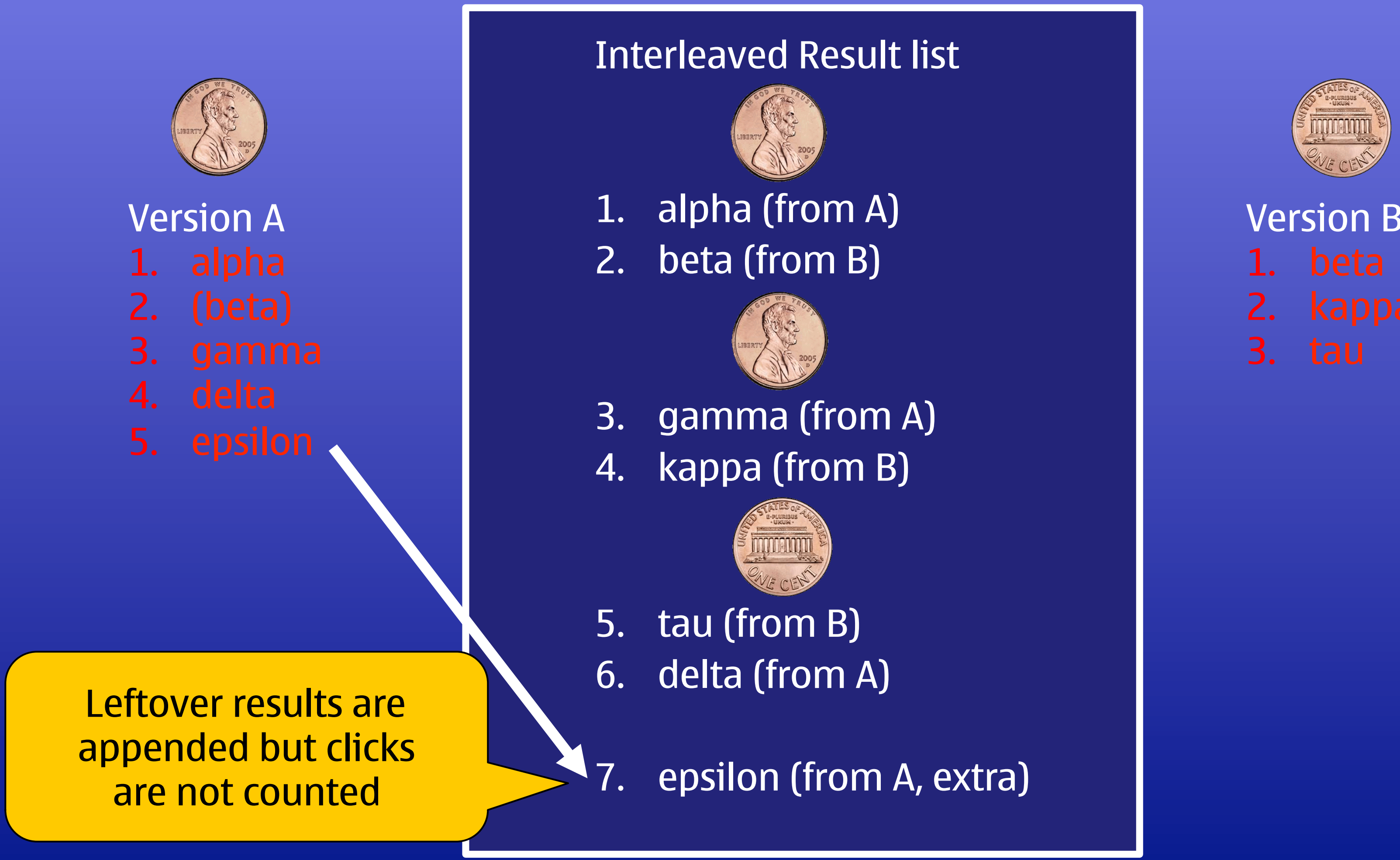
1. alpha (from A)
2. beta (from B)
3. gamma (from A)
4. kappa (from B)
5. tau (from B)
6. delta (from A)

Version B

1. beta
2. kappa
3. tau

A/B Interleaving: Randomized Mixing of Lists

- Interleaved list is filled with pairs of results, one item from each version. Coin toss decides who comes first.



Version A

1. alpha
2. (beta)
3. gamma
4. delta
5. epsilon

Interleaved Result list

1. alpha (from A)
2. beta (from B)
3. gamma (from A)
4. kappa (from B)
5. tau (from B)
6. delta (from A)
7. epsilon (from A, extra)

Version B

1. beta
2. kappa
3. tau

Leftover results are appended but clicks are not counted

A/B Interleaving: Randomized Mixing of Lists

- Interleaved list is filled with pairs of results, one item from each version. Coin toss decides who comes first.

Version A

1. alpha
2. (beta)
3. gamma
4. delta
5. epsilon

Final list shown to user

1. alpha (from A)
2. beta (from B)
3. gamma (from A)
4. kappa (from B)
5. tau (from B)
6. delta (from A)
7. epsilon (from A, extra)

Version B

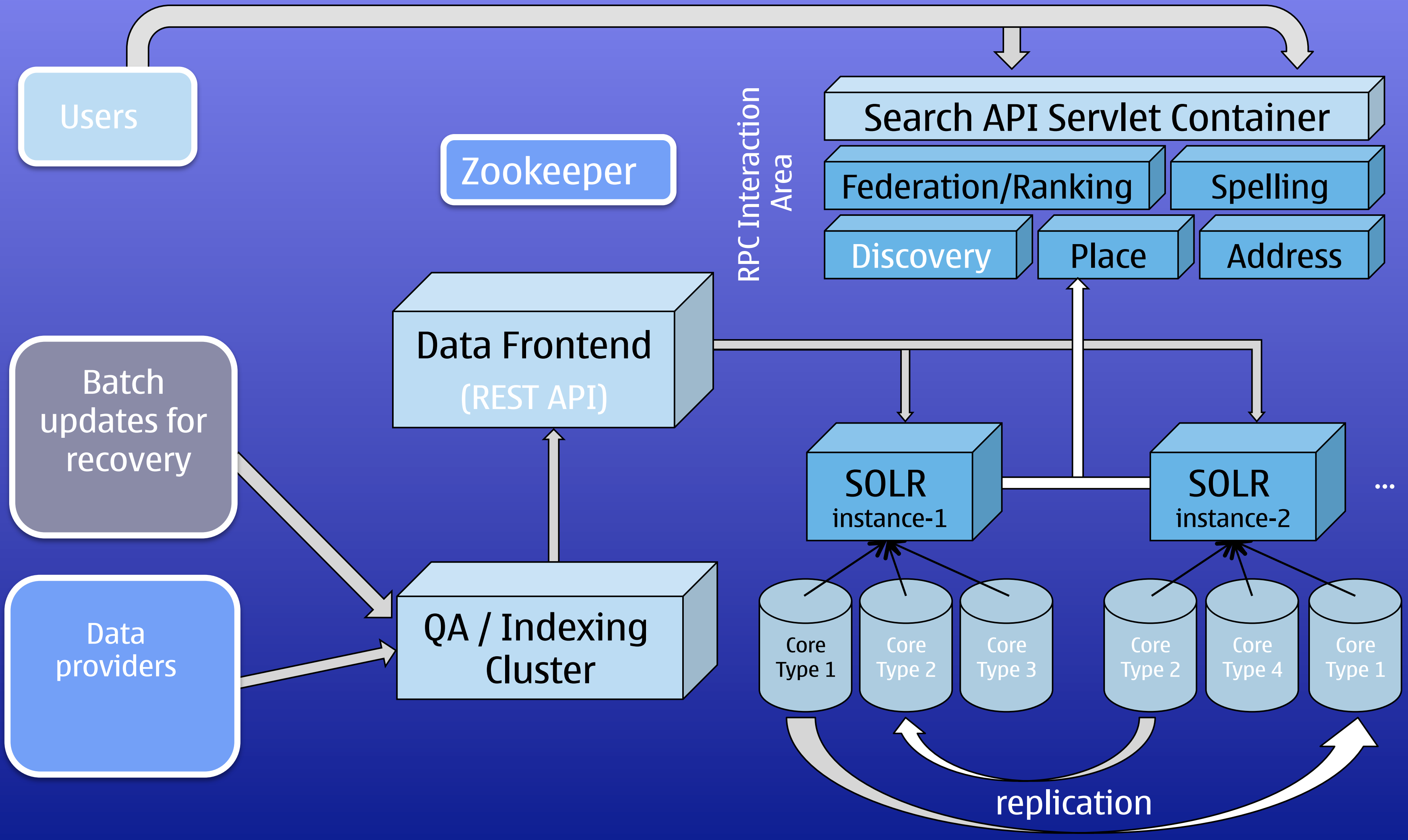
1. beta
2. kappa
3. tau

Declaring A Winner

- Statistical Significance Test
- Input (after hadoop-based log-processing...)
  - Number of clicks on version A
  - Number of clicks on version B
- G-Test:
  - improved version of Pearson's Chi-squared test.
  - G > 6.635 corresponds to 99% confidence level
- Null hypothesis:
  - Frequency of counts is equally distributed over both versions.
- Test statistic:

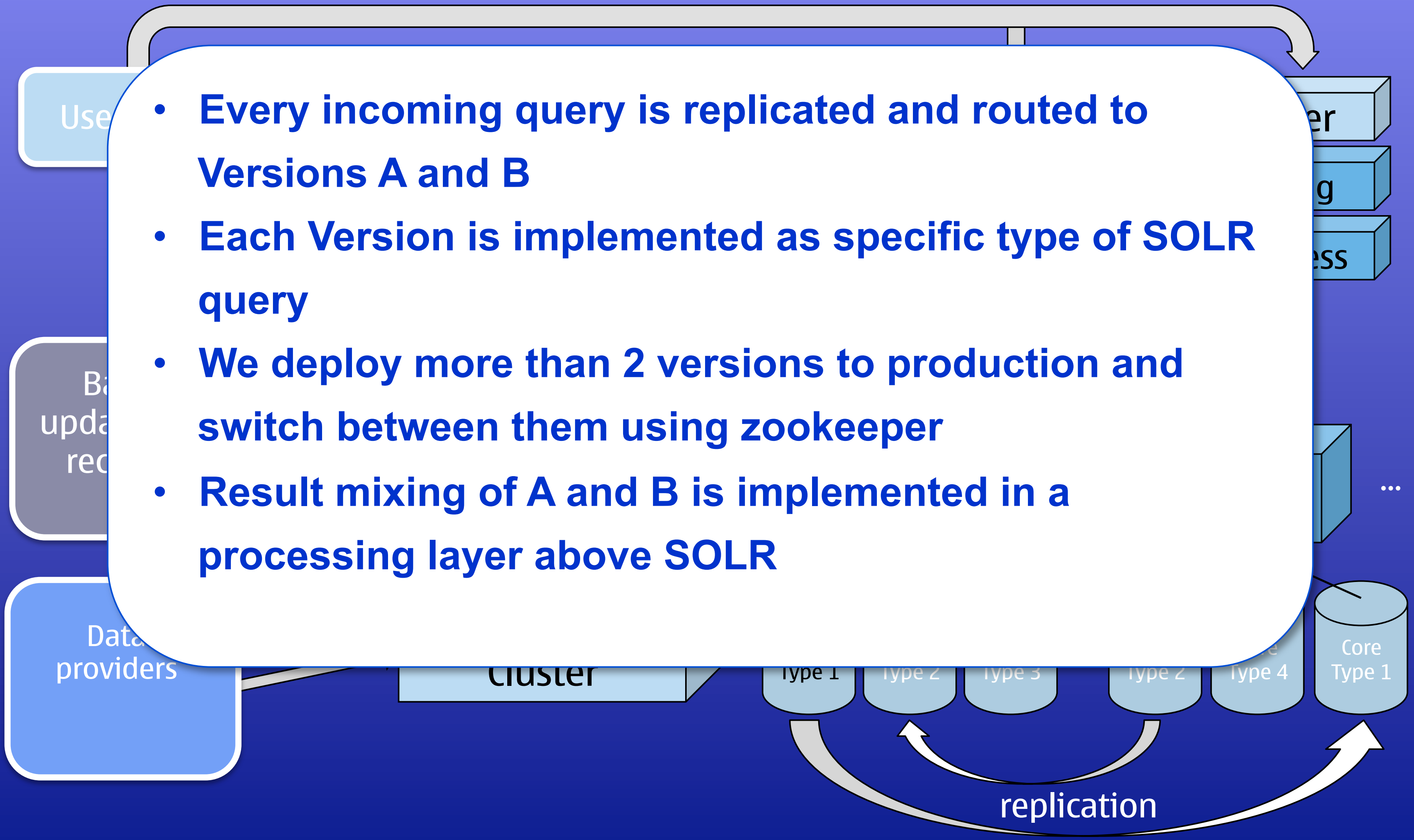
$$G = 2 \sum_{i \in \{A,B\}} [\text{counts i}] \ \ln\left(\frac{[\text{counts i}]}{[\text{total counts/2}]}\right)$$

Managing Multiple Versions



NOKIA

Managing Multiple Versions



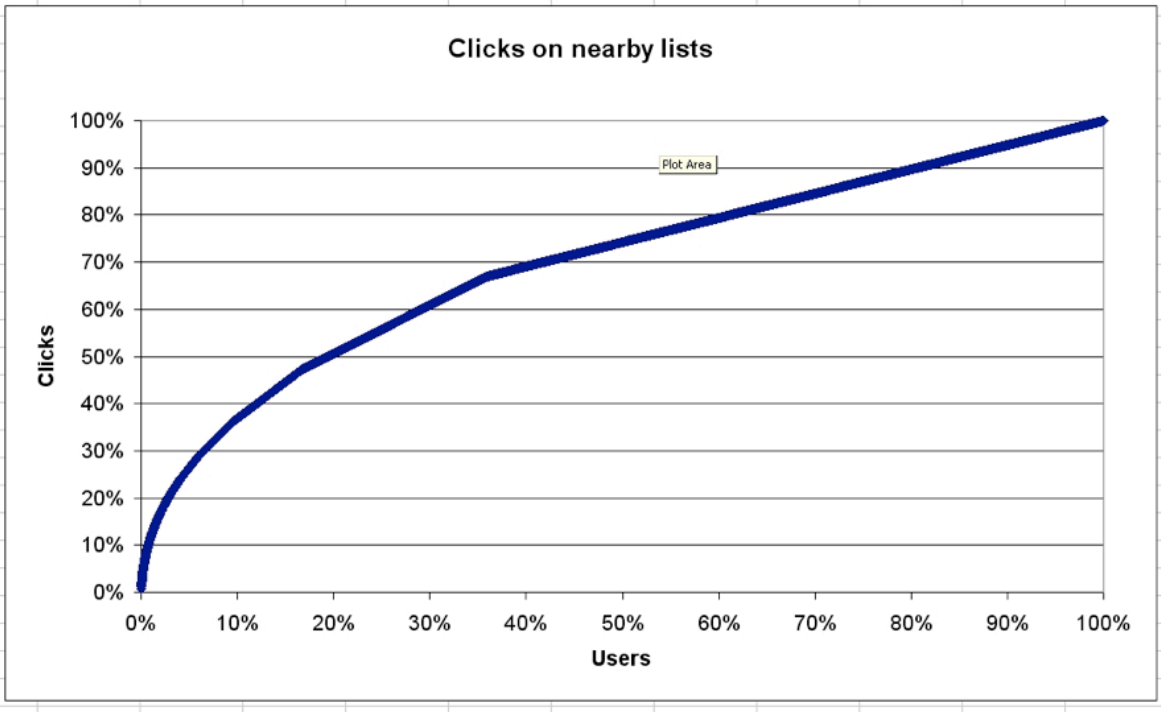
- Every incoming query is replicated and routed to Versions A and B
- Each Version is implemented as specific type of SOLR query
- We deploy more than 2 versions to production and switch between them using zookeeper
- Result mixing of A and B is implemented in a processing layer above SOLR

Caveat 1: Randomization

- don’t confuse users with changing results, i.e.: provide a consistent user experience
- Solution:
  - Random generator is seeded with USER-ID for each query.
  - Each user gets his personal random generator.

Caveat 2: Healthy Click Data

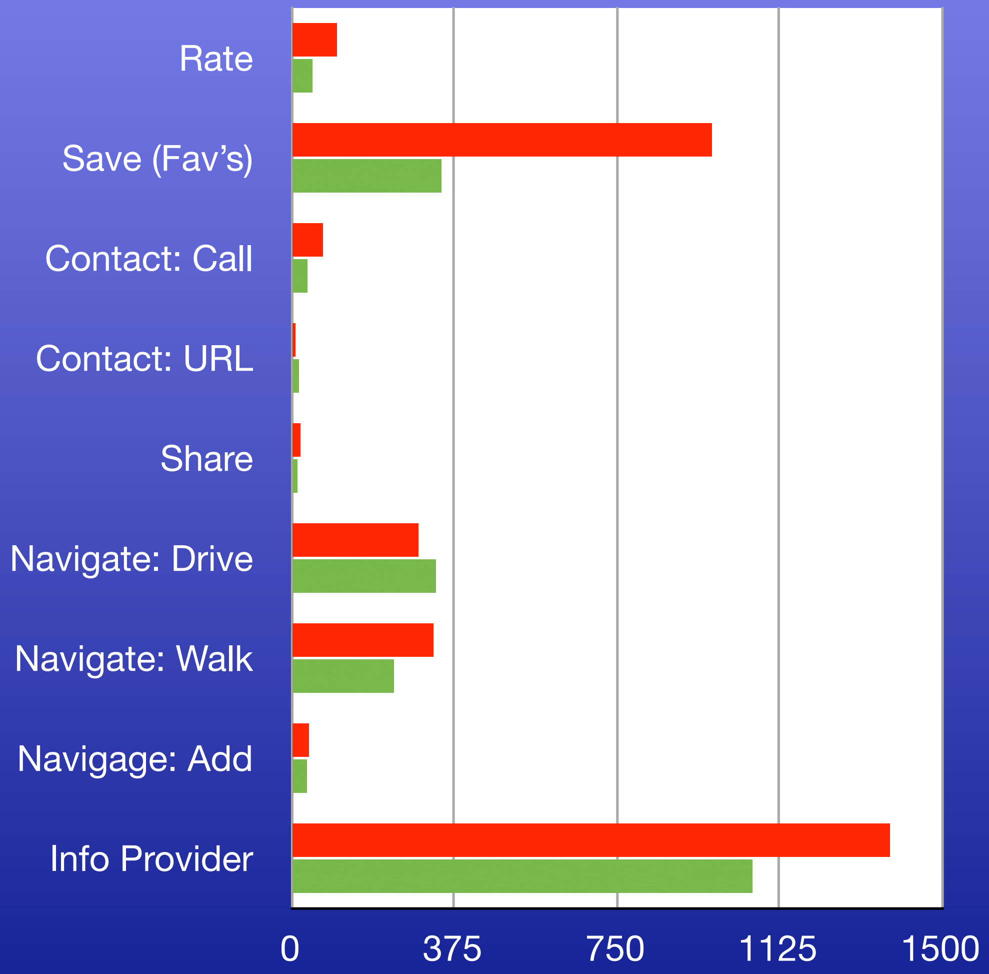
- we are relying on the integrity of transmitted user actions
- sensitive to log contamination (unidentified QA, spam)
- user-clicks plot:



Caveat 3: A/B Clicks vs. Coverage

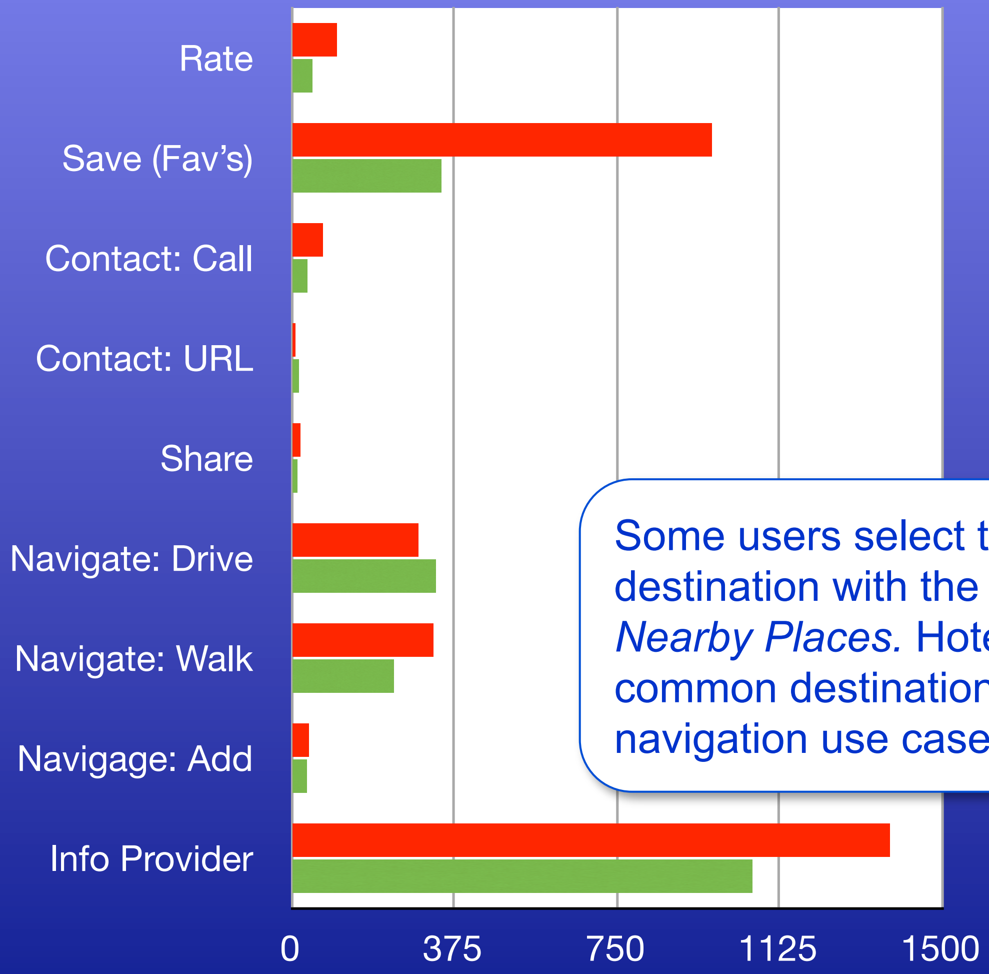
- Coverage = non-empty responses (in percent)
- For example
    - A/B interleaving of eat&drink vs. eat&drink + going out
    - difference is not significant
    - But coverage different, percentage of responses with POIs nearby:
        - 60% eat&drink
        - 62% eat&drink + going out
- Higher coverage decides in case there is no statistical difference

Case Study: Eat'n'Drink versus Hotels: Not the User Behaviour we had expected!



NOKIA

Case Study: Eat’n’Drink versus Hotels: Not the User Behaviour we had expected!



Some users select their driving destination with the help of *Nearby Places.* Hotels are a common destination in the car navigation use case.

Summary

- use A/B Rank Interleaving to optimize result relevance
- Rank Interleaving is easy to implement. Works.
- in a distributed search architecture manage your A/B test configurations conveniently using Zookeeper
- harness your hadoop/search analytics stack for A/B test evaluations
- don’t make assumptions about your users!

Thanks!

Get in touch: *hannes.kruppa@nokia.com*

Nokia Maps "Place Discovery" Team, Berlin:

Hannes Kruppa, Steffen Bickel, Mark Waldaukat,
Felix Weigel, Ross Turner, Peter Siemen