Feeling Lucky? Multi-armed bandits for Ordering Judgements in Pooling-based Evaluation

David E. Losada

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CiTIUS Centro Singular de Investigación en Tecnoloxías da Información

Javier Parapar, Álvaro Barreiro

UNIVERSIDADE DA CORUÑA

IR Lab

ACM SAC, 2016

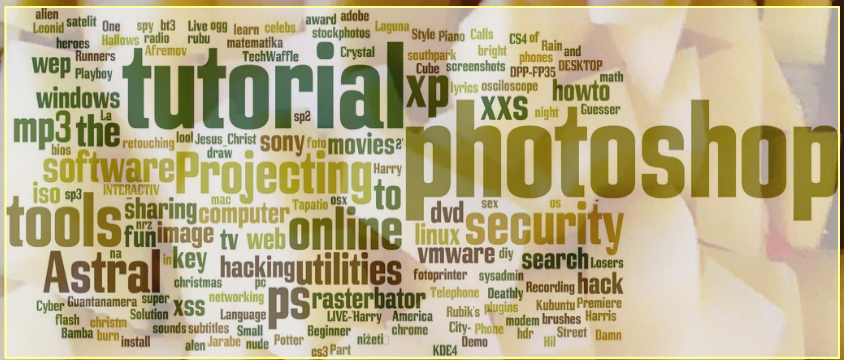
Evaluation is crucial

compare **retrieval** algorithms, design new **search** solutions, ...


information retrieval evaluation:
3 main ingredients
docs

information retrieval evaluation:
3 main ingredients

queries



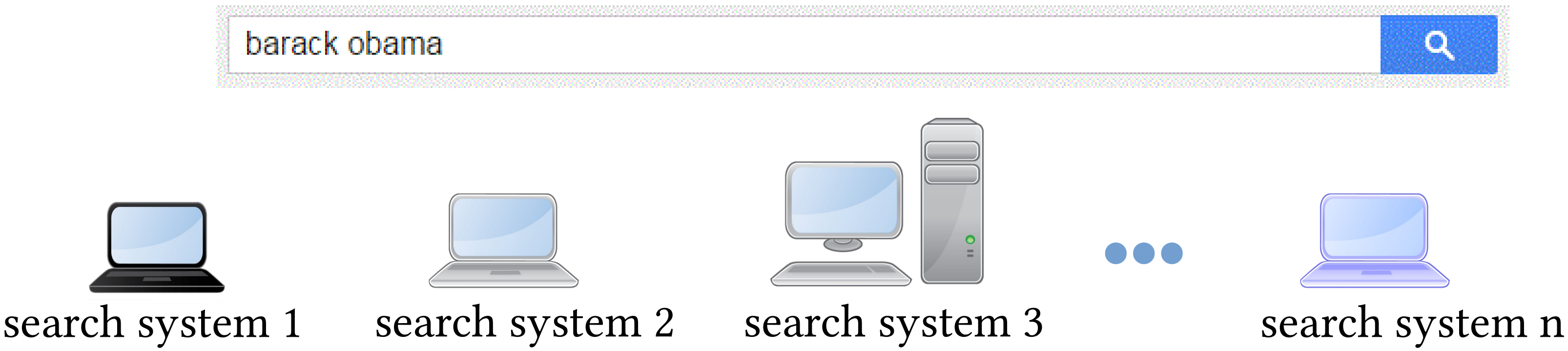
information retrieval evaluation:
3 main ingredients

relevance judgements

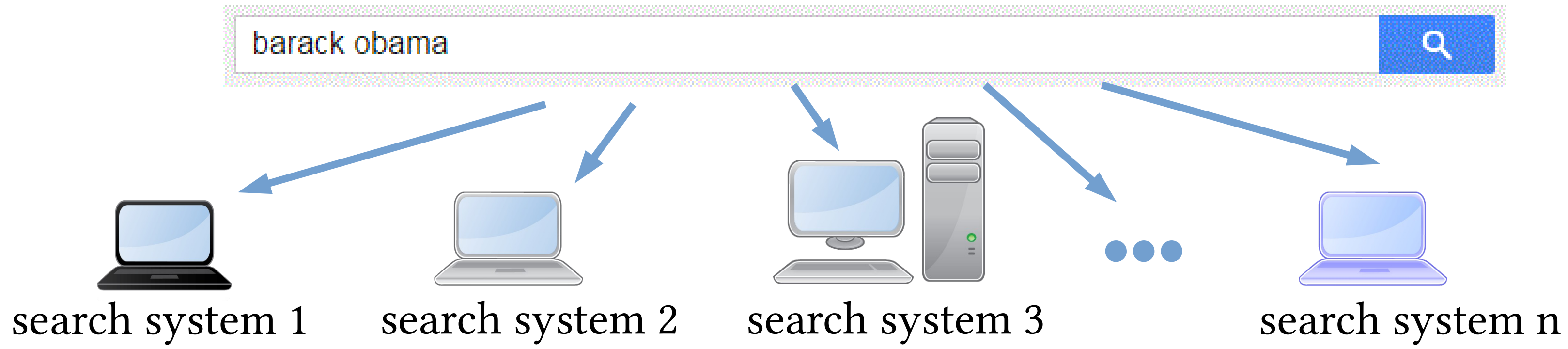
!

relevance assessments are incomplete

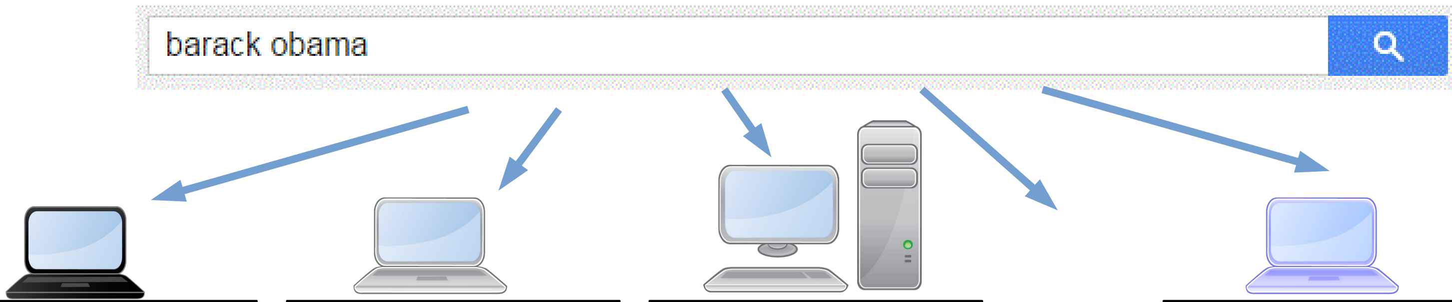
relevance assessments are incomplete



relevance assessments are incomplete



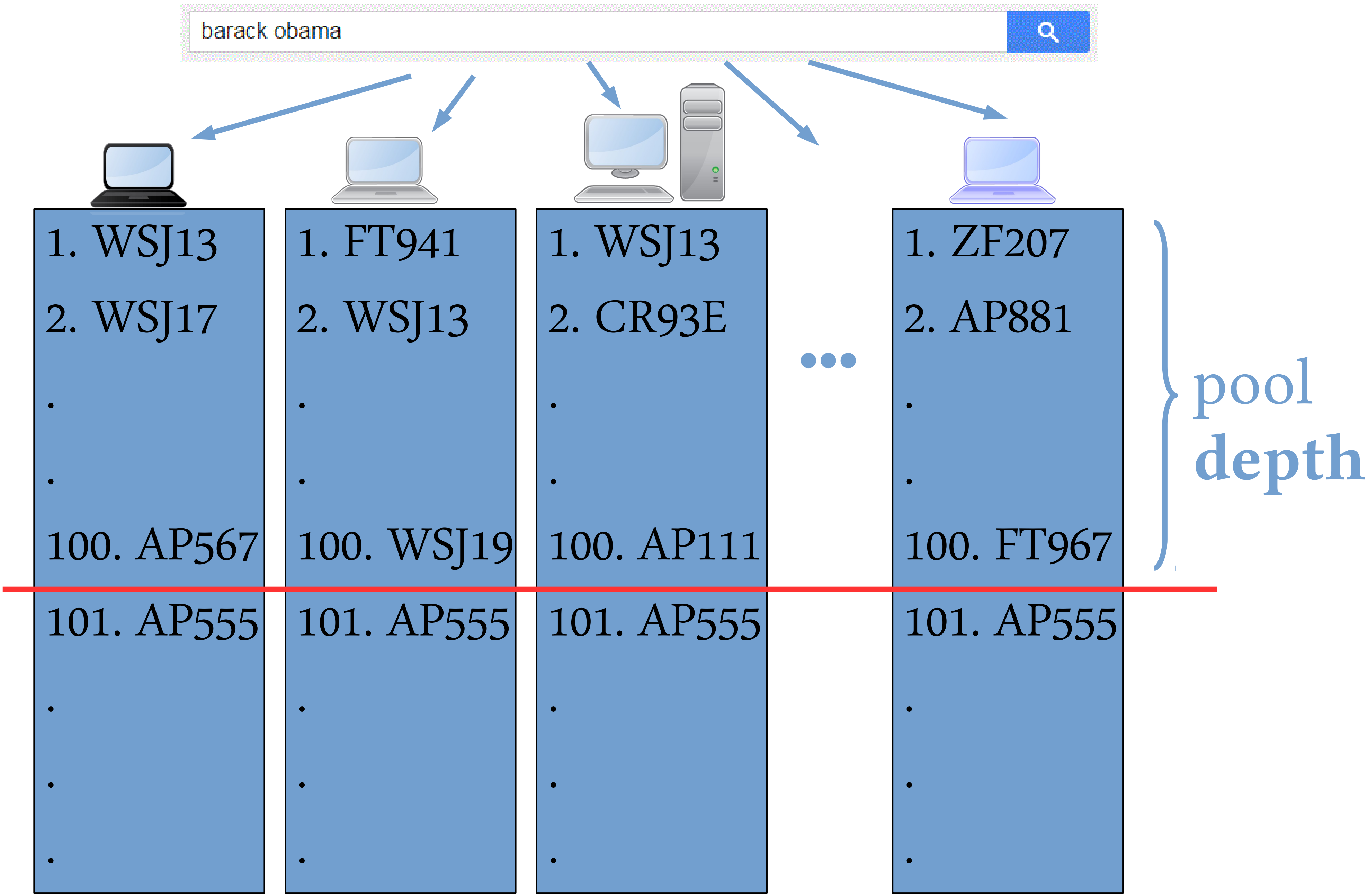
relevance assessments are incomplete



1. WSJ13	1. FT941	1. WSJ13		1. ZF207
2. WSJ17	2. WSJ13	2. CR93E	•••	2. AP881
.	.	.		.
.	.	.		.
100. AP567	100. WSJ19	100. AP111		100. FT967
101. AP555	101. AP555	101. AP555		101. AP555
.	.	.		.
.	.	.		.
.	.	.		.

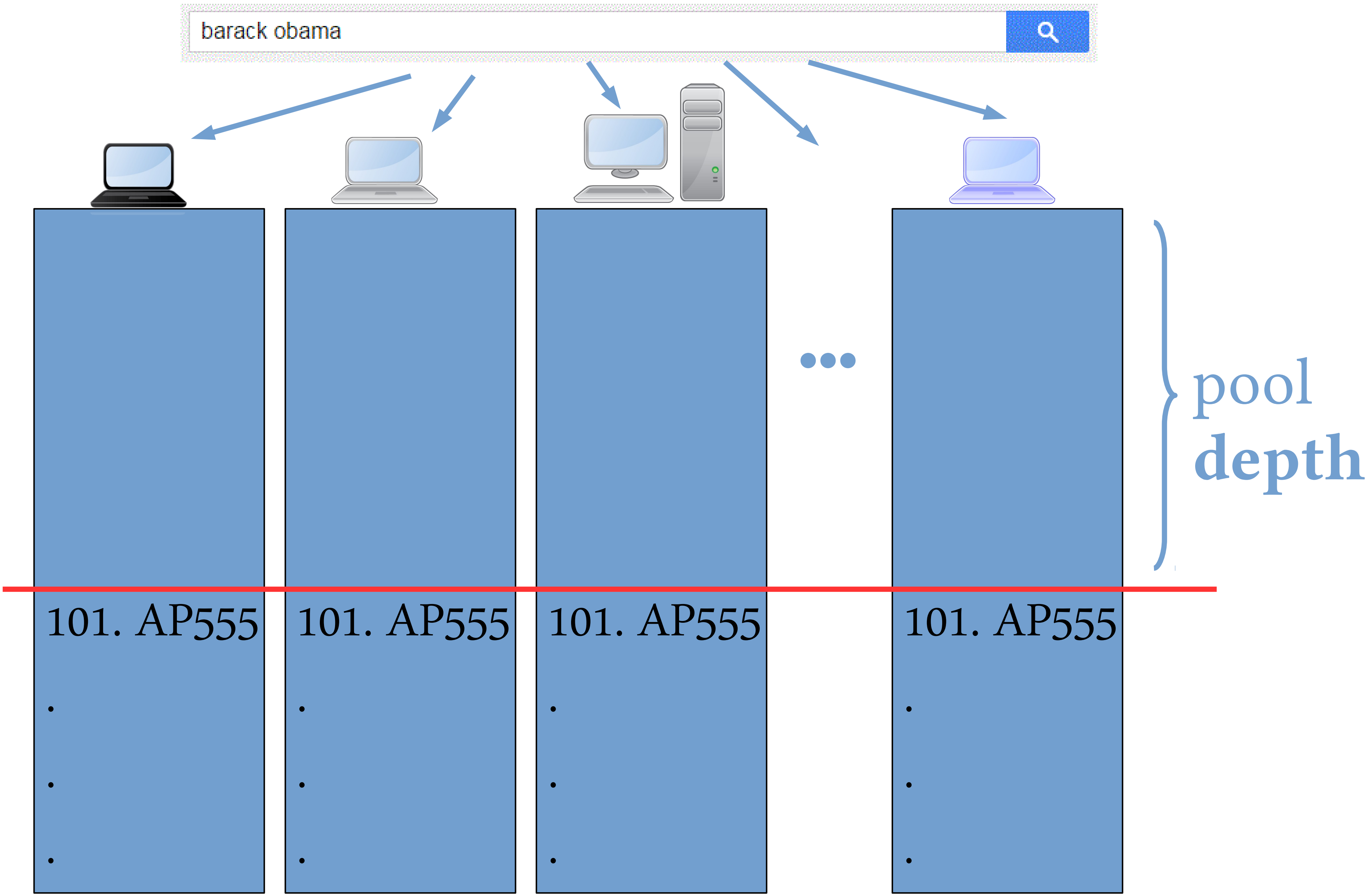
rankings of docs by **estimated relevance** (runs)

relevance assessments are **incomplete**



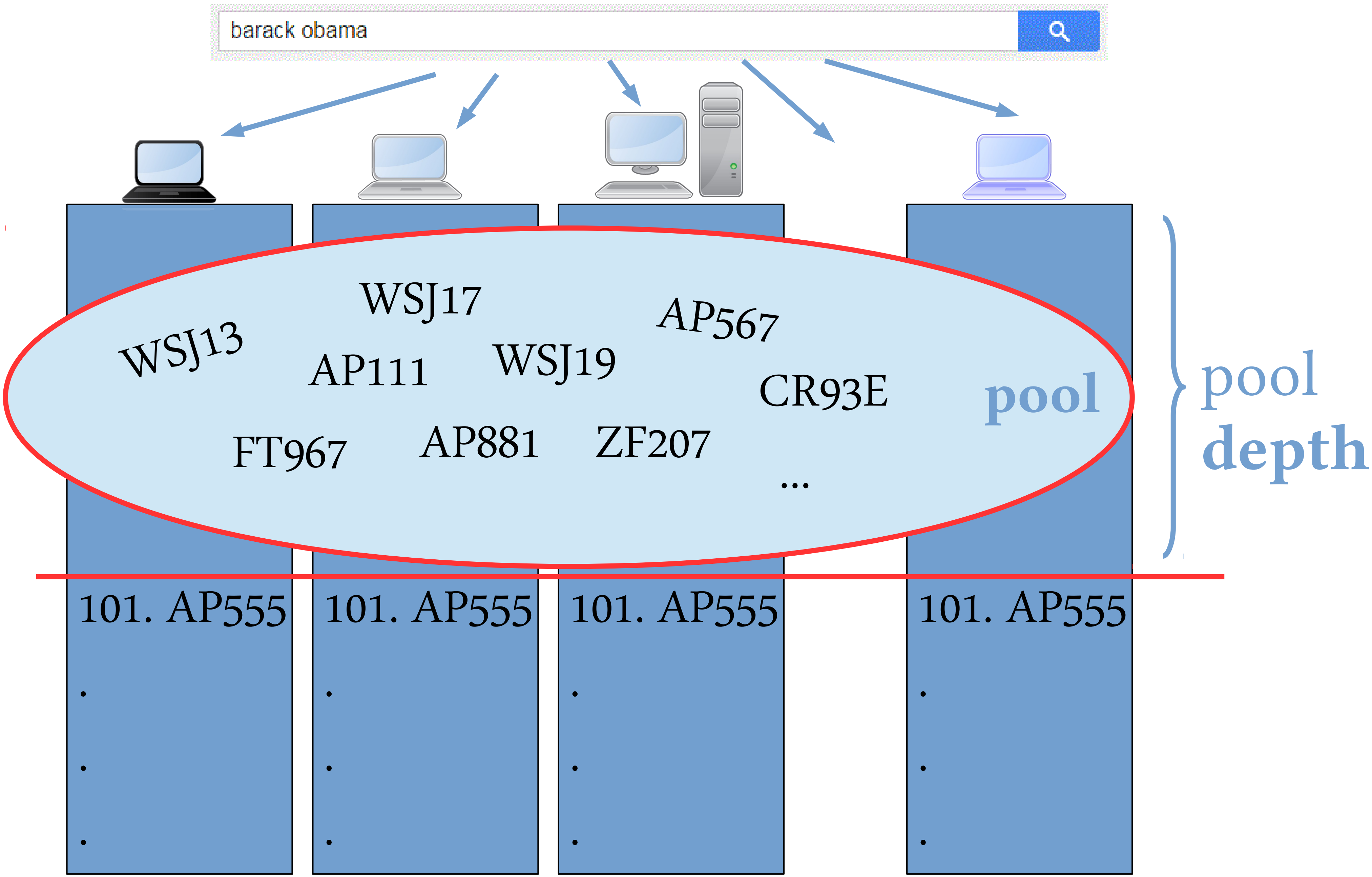
rankings of docs by **estimated relevance** (runs)

relevance assessments are incomplete



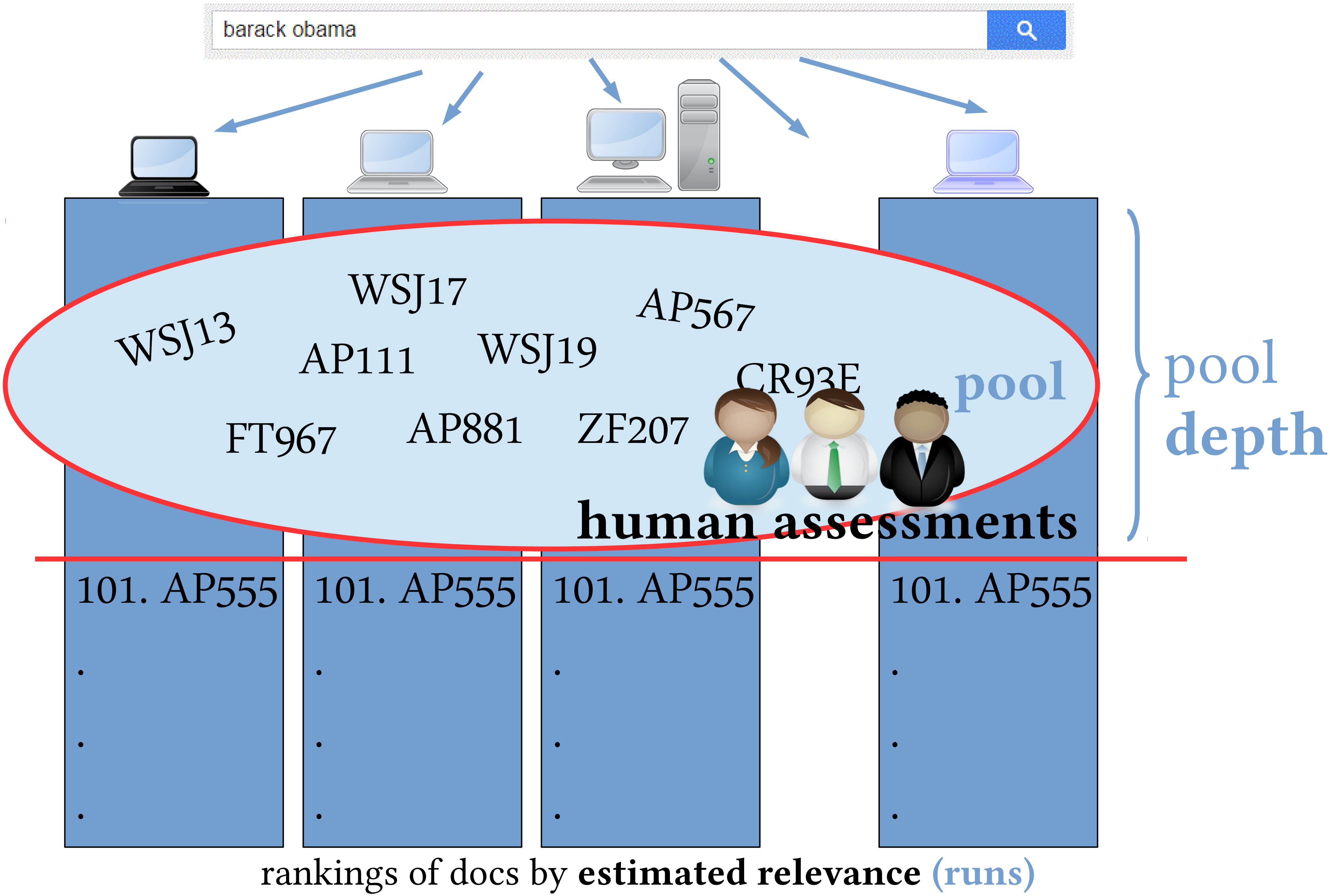
rankings of docs by **estimated relevance (runs)**

relevance assessments are incomplete



rankings of docs by **estimated relevance** (runs)

relevance assessments are incomplete



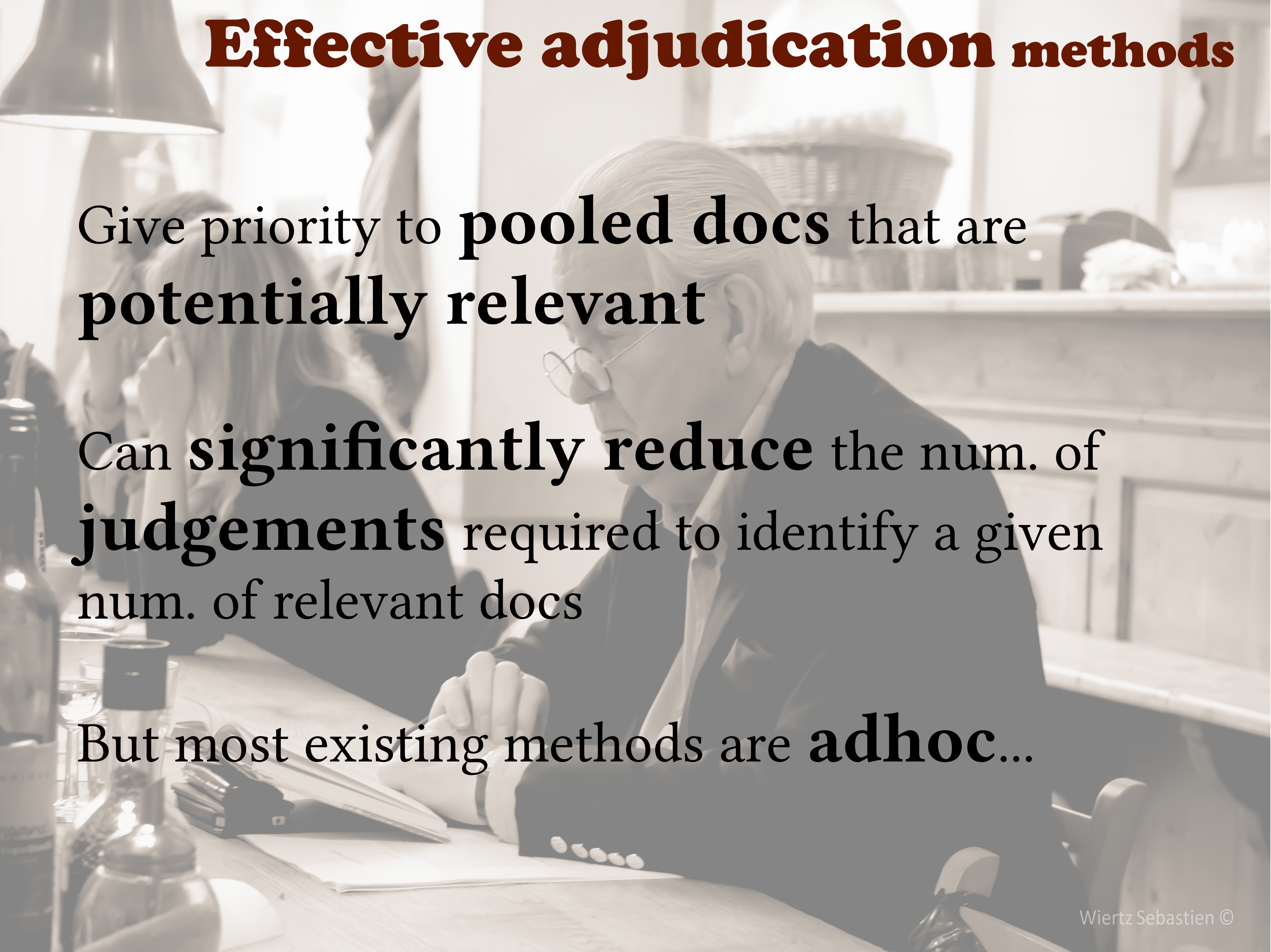

finding relevant docs is the key
Most productive use of assessors' time
is spent on judging relevant docs
(Sanderson & Zobel, 2005)

Effective adjudication methods

Give priority to **pooled docs** that are **potentially relevant**

Can **significantly reduce** the num. of **judgements** required to identify a given num. of relevant docs

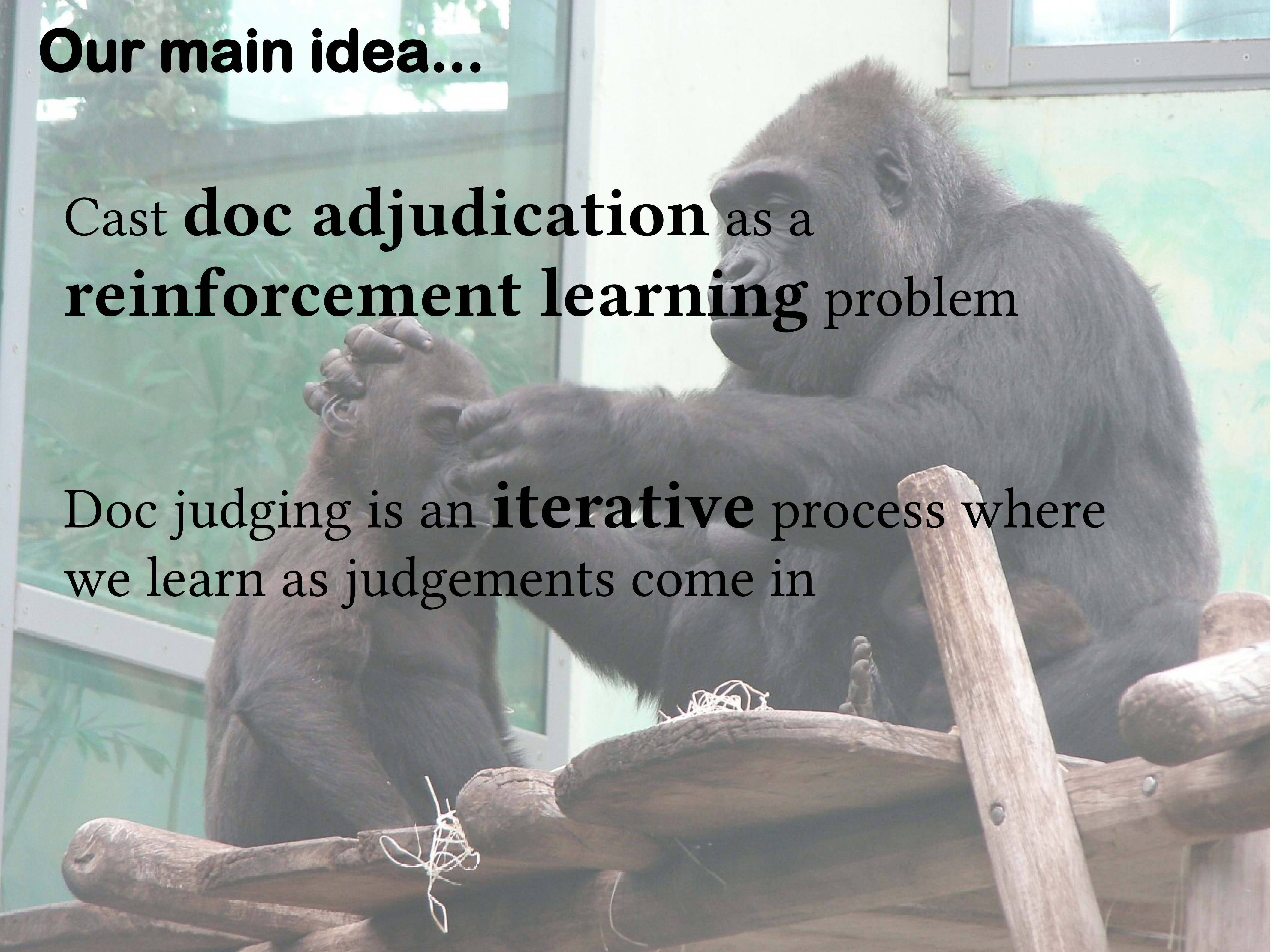
But most existing methods are **adhoc**...



Our main idea…

Cast **doc adjudication** as a **reinforcement learning** problem

Doc judging is an **iterative** process where we learn as judgements come in



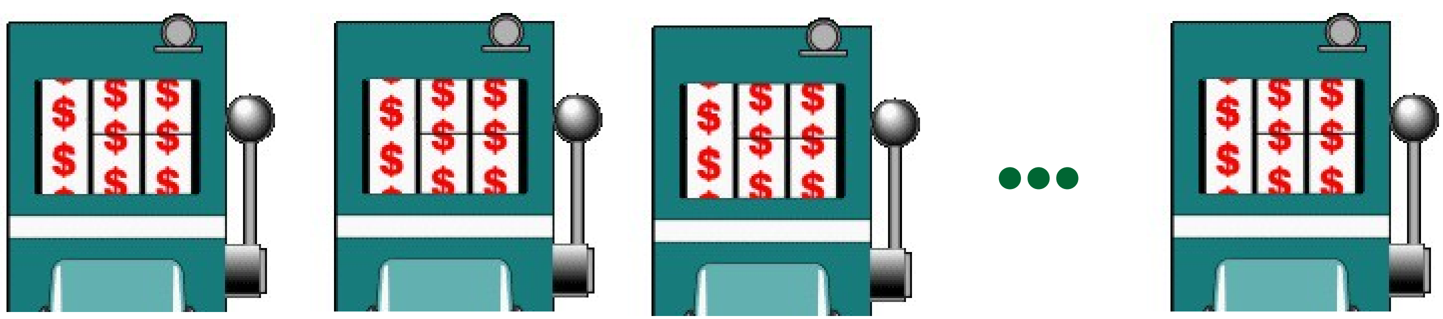
Doc adjudication as a reinforcement learning problem

Initially we **know nothing** about the quality of the runs

As **judgements come in**...

And we can **adapt** and allocate more docs for judgement from the most promising runs

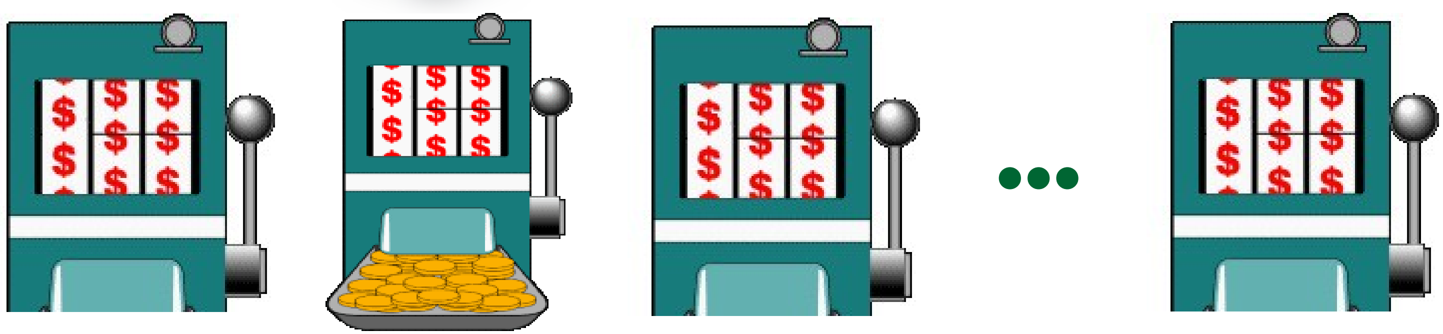
Multi-armed bandits



unknown probabilities of giving a prize

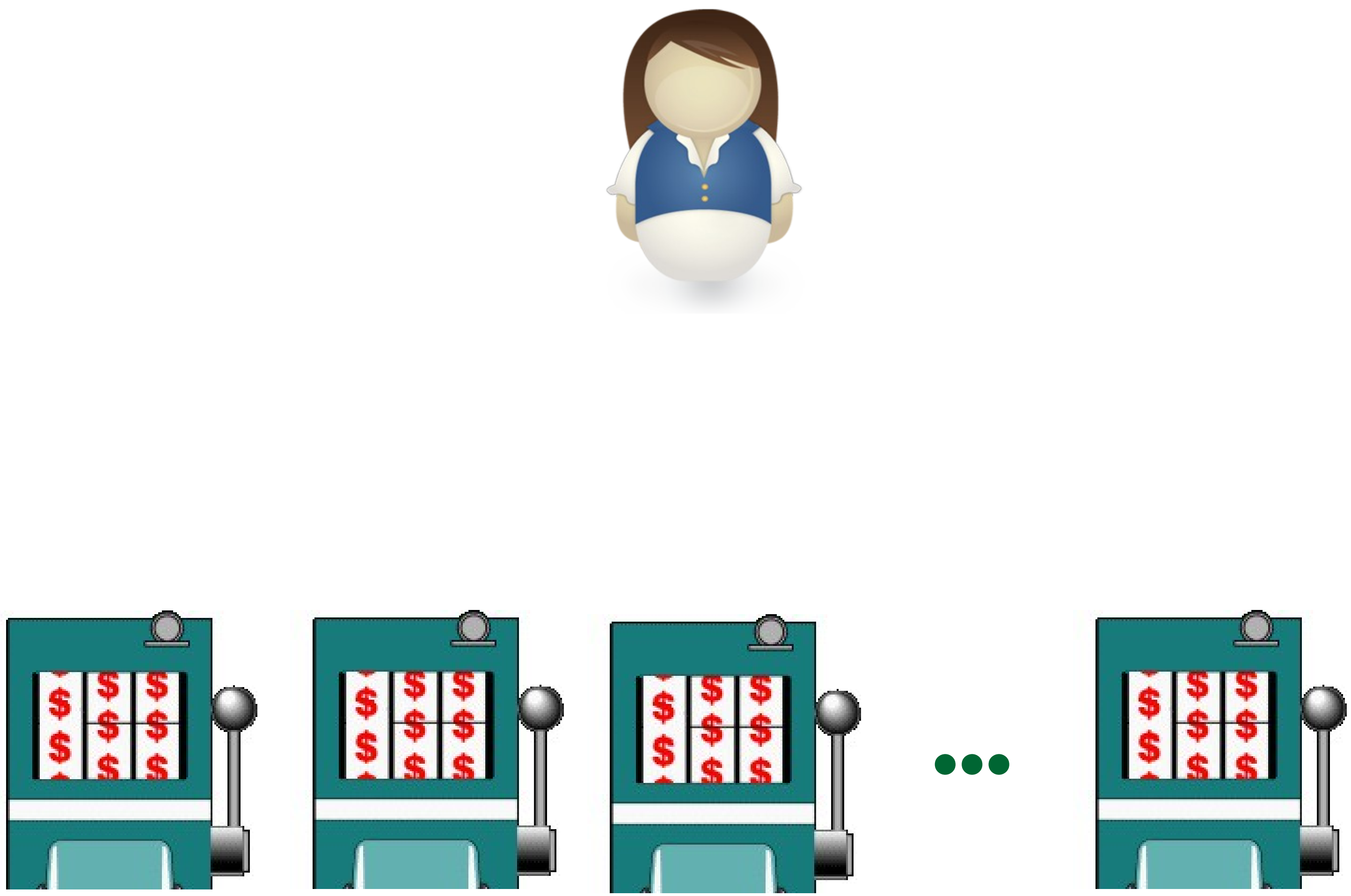
Multi-armed bandits

play and observe the **reward**



unknown probabilities of giving a prize

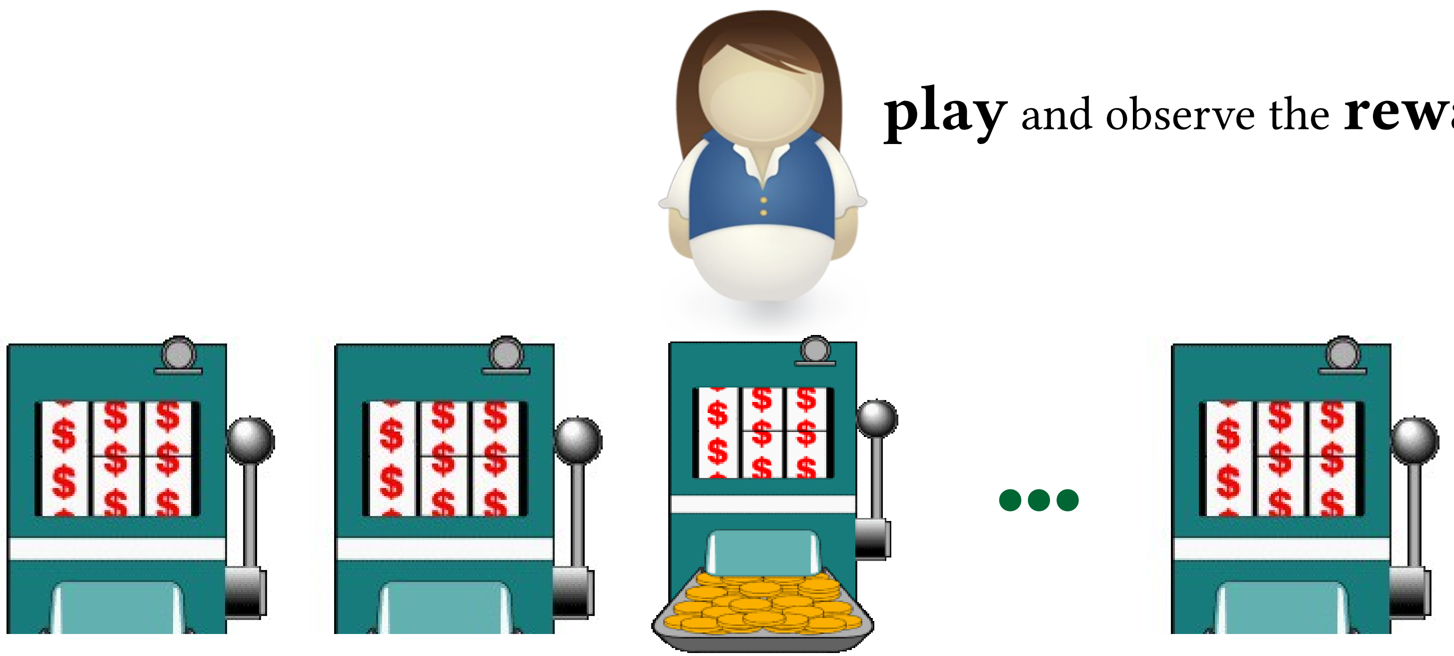
Multi-armed bandits



unknown probabilities of giving a prize

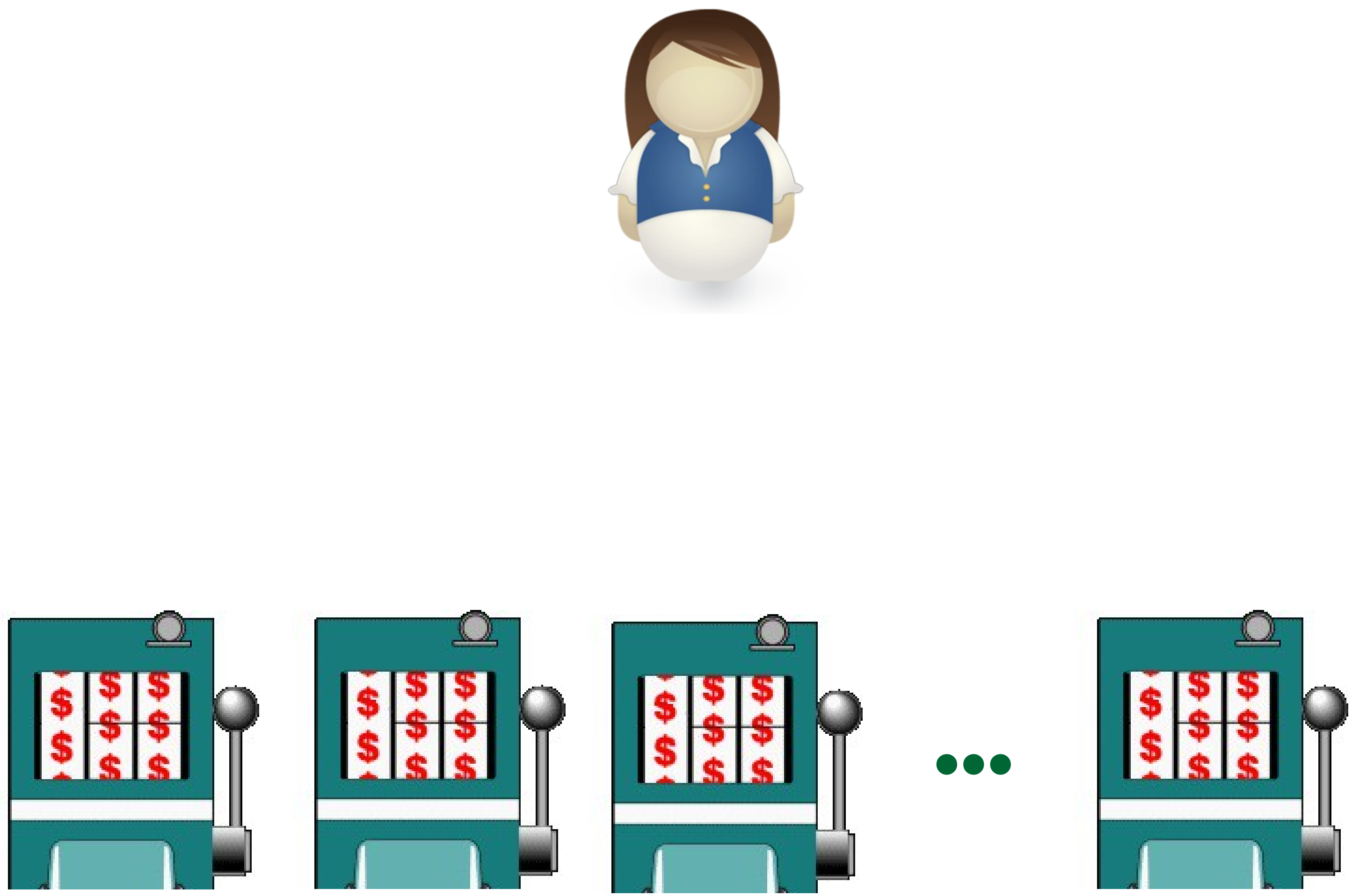
Multi-armed bandits

play and observe the **reward**



unknown probabilities of giving a prize

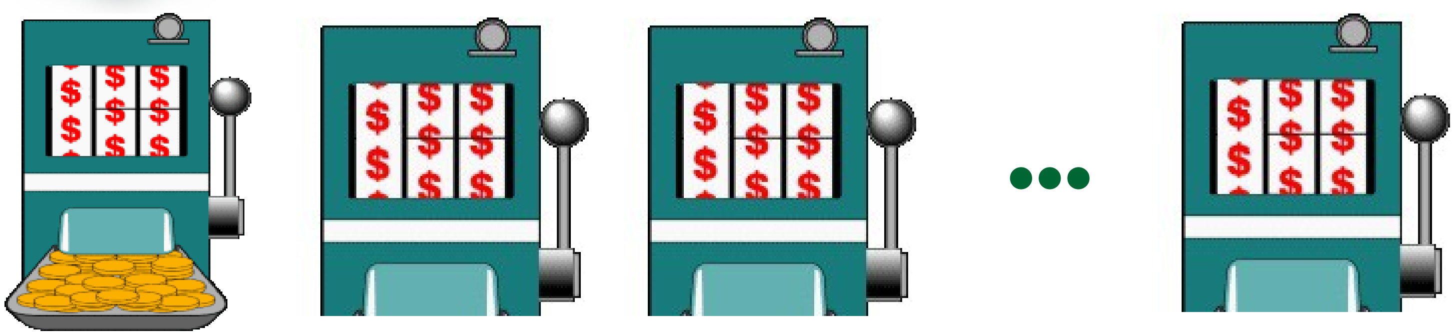
Multi-armed bandits



unknown probabilities of giving a prize

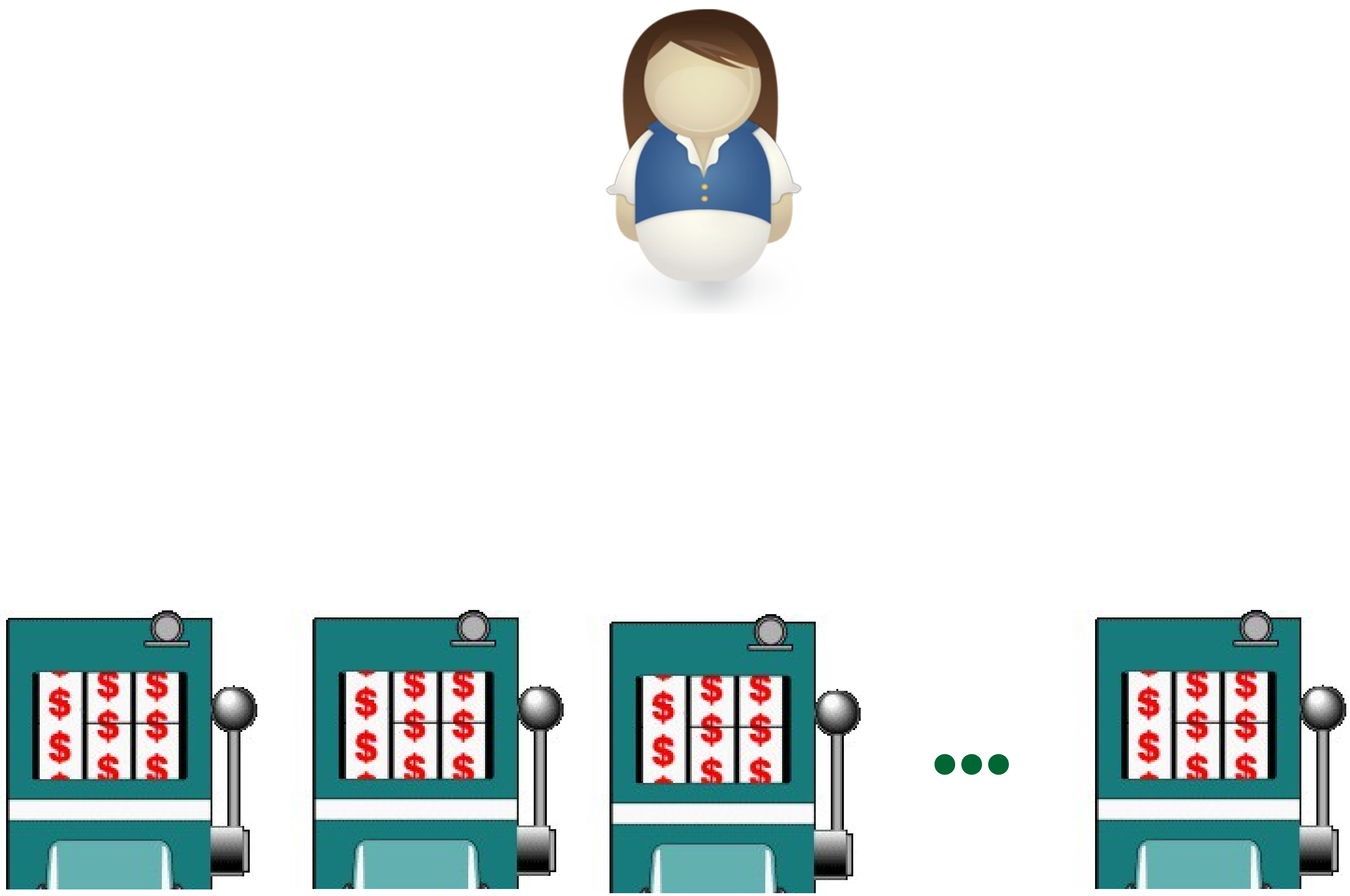
Multi-armed bandits

play and observe the **reward**



unknown probabilities of giving a prize

Multi-armed bandits



unknown probabilities of giving a prize

exploration vs exploitation

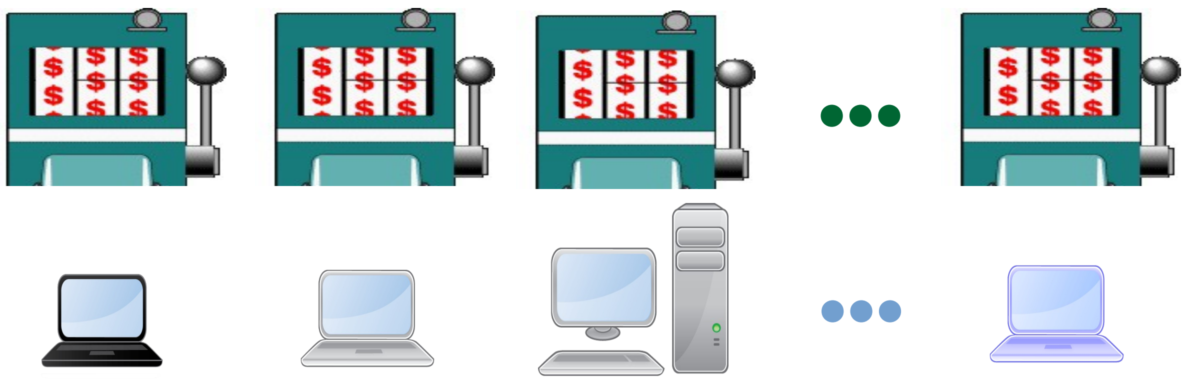
explores **uncertain** actions
gets **more info** about expected payoffs
may produce greater total reward
in the long run

exploits **current knowledge**
spends no time sampling inferior actions
maximizes expected reward
on the **next** action

allocation methods: choose next action (play) based on past plays and obtained rewards
implement different ways to **trade** between exploration and exploitation

Multi-armed bandits for ordering judgements

machines = runs



play a machine = select a run and get the next (unjudged) doc

1. WSJ13
2. CR93E

.

.

(binary) reward = relevance/non-relevance of the selected doc

Allocation methods tested

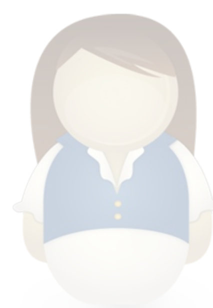
random

Upper Confidence Bound (UCB)

computes upper confidence bounds for avg rewards

conf. intervals get **narrower with the number of plays**

selects the machine with the **highest optimistic estimate**

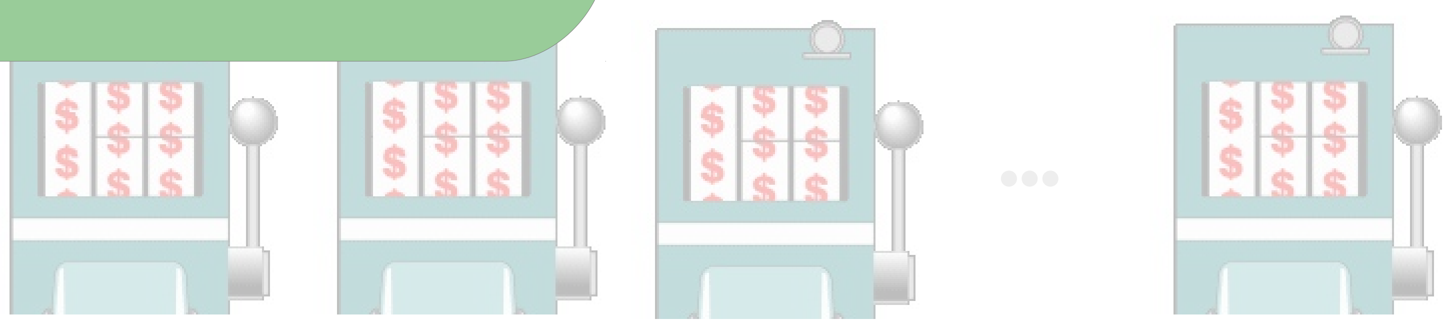


ϵ_n-greedy

with **prob 1-ϵ** plays the machine with the highest avg reward

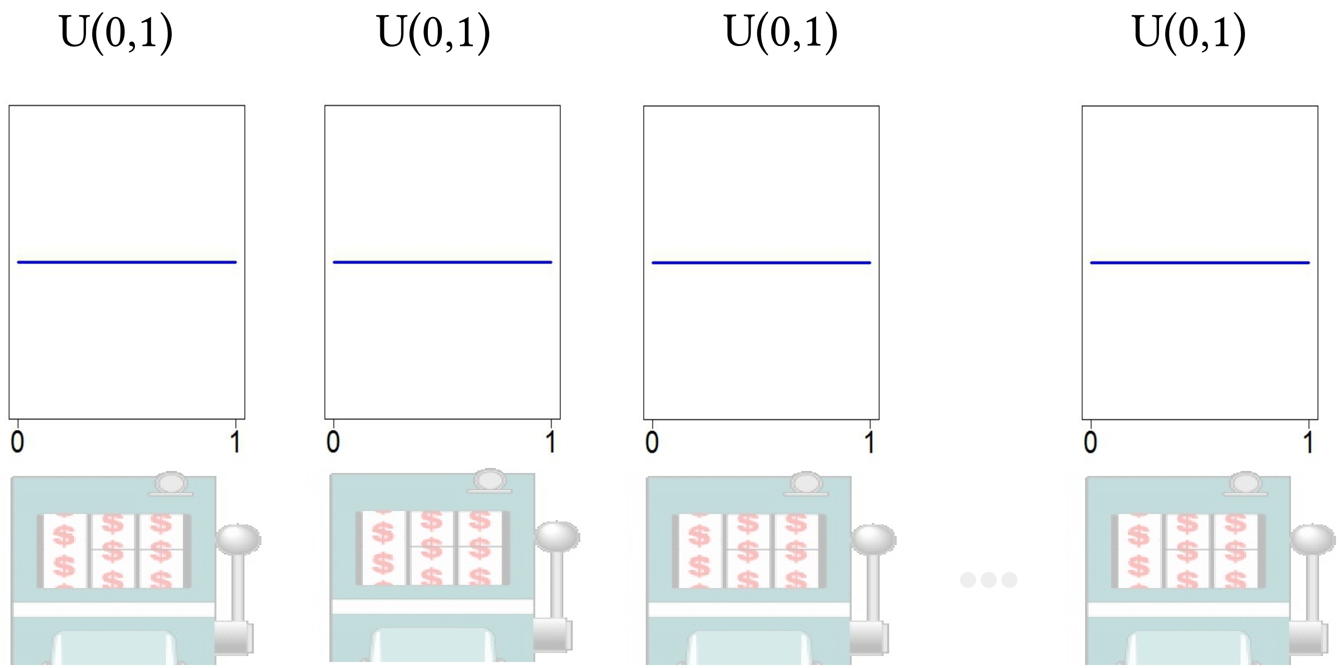
with **prob ϵ** plays a random machine

prob of exploration (ϵ) decreases with the num. of plays



Allocation methods tested: Bayesian bandits

prior probabilities of giving a relevant doc: **Uniform(0,1)** (or, equivalently, Beta(α,β), α,β=1)

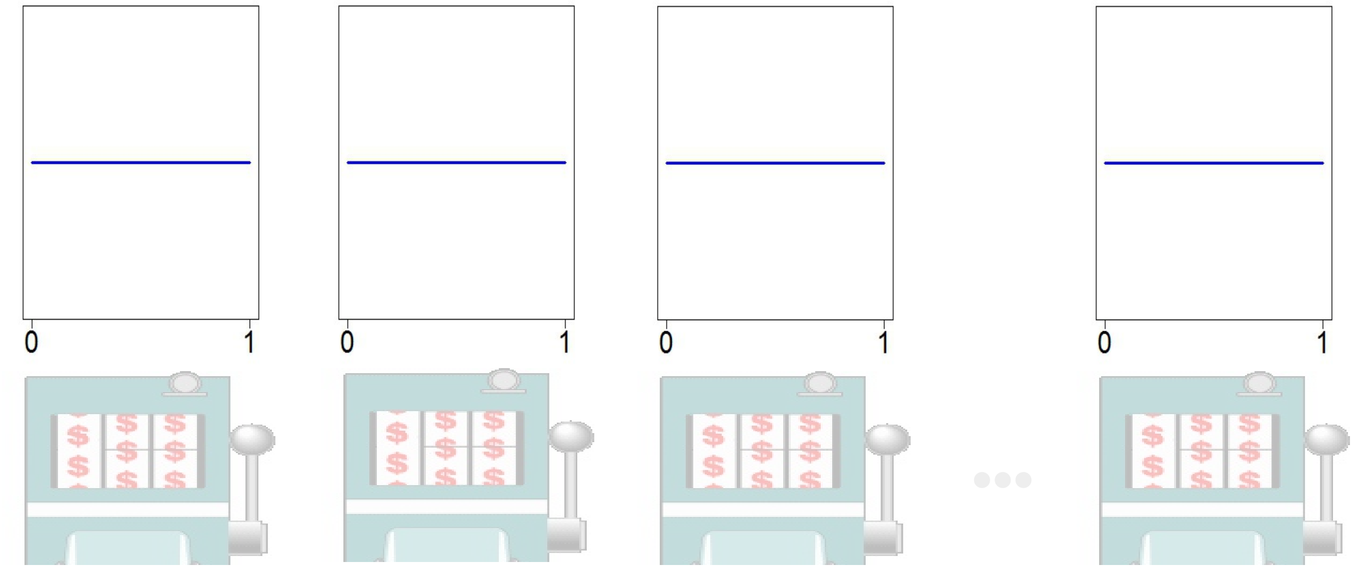


evidence (O ∈ {0,1}) is Bernoulli (or, equivalently, Binomial(1,p))

posterior probabilities of giving a relevant doc: Beta(α+O, β+1-O) (Beta: conjugate prior for Binomial)

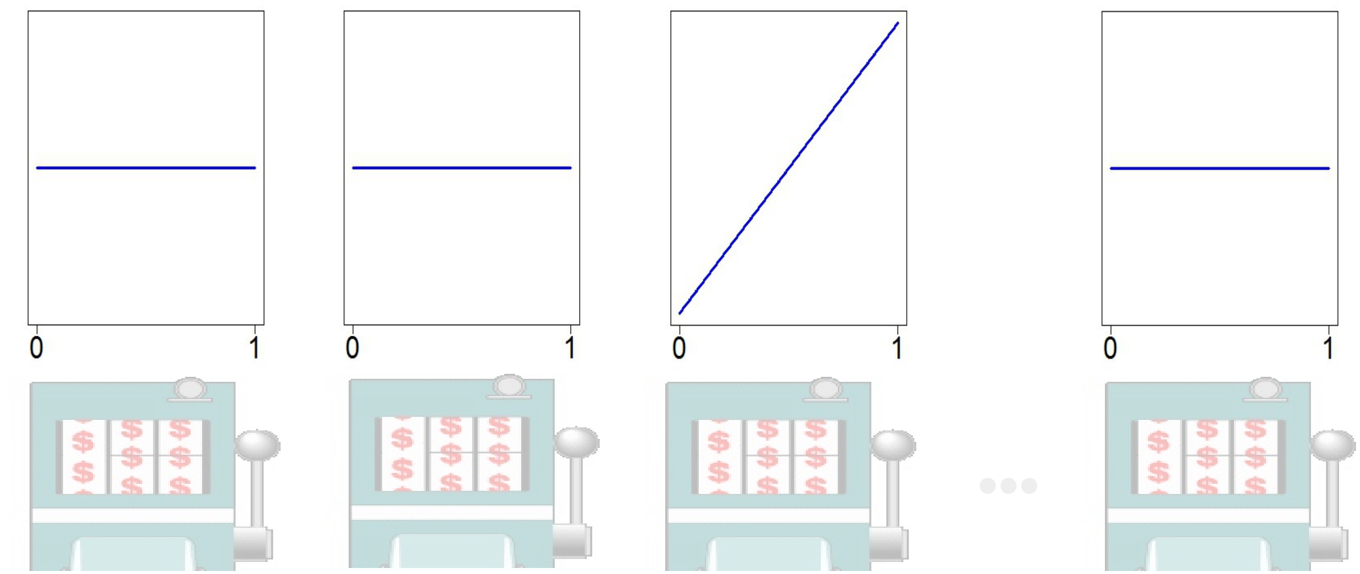
Allocation methods tested: Bayesian bandits

we **iteratively update** our estimations using Bayes:



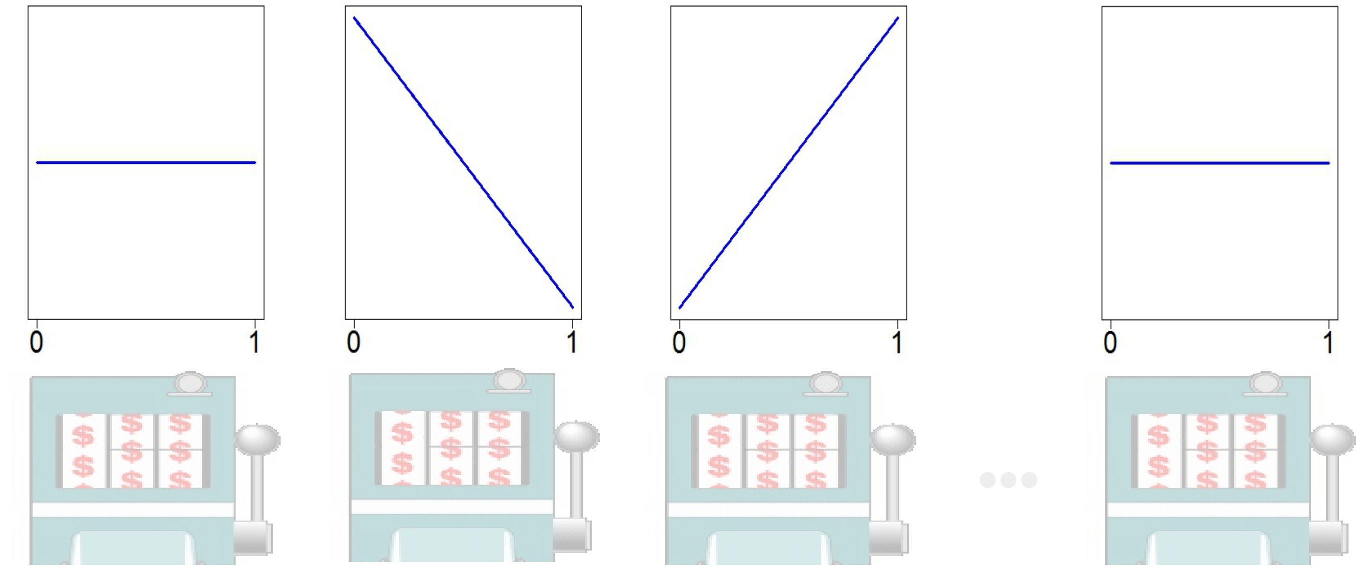
Allocation methods tested: Bayesian bandits

we **iteratively update** our estimations using Bayes:



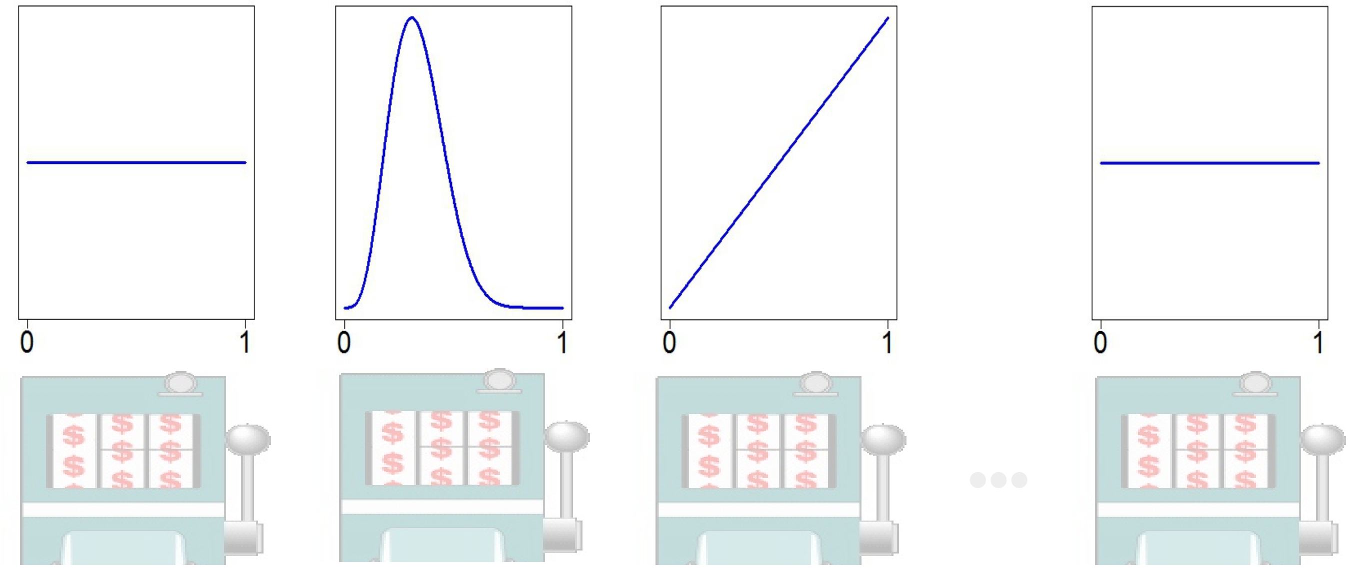
Allocation methods tested: Bayesian bandits

we **iteratively update** our estimations using Bayes:



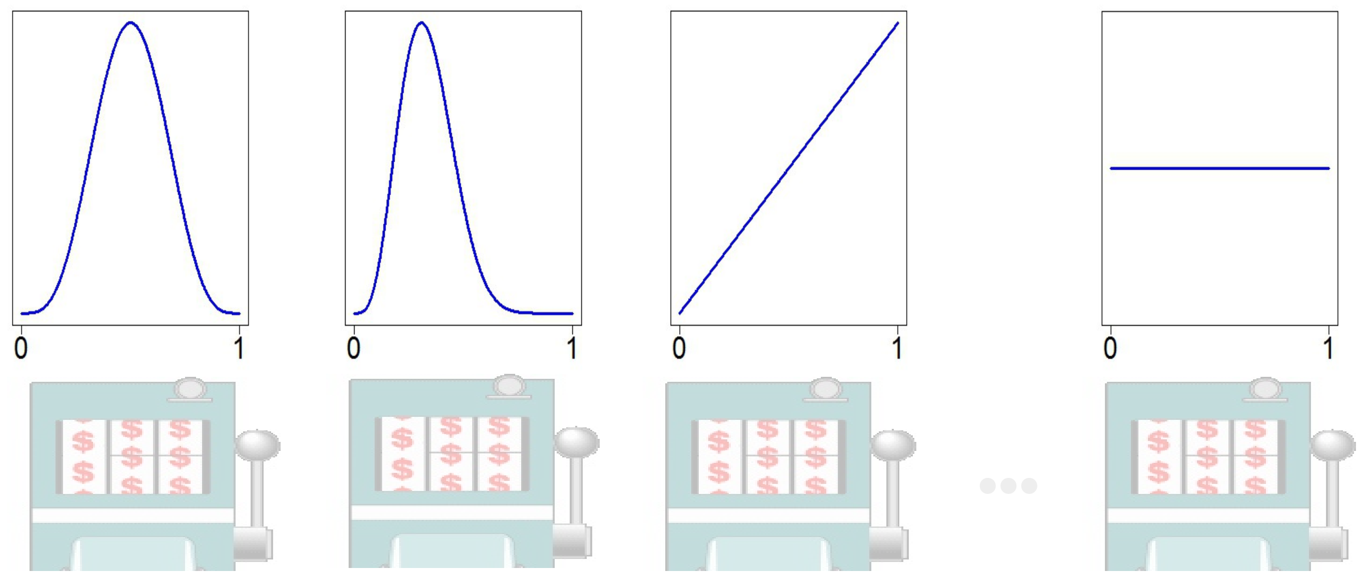
Allocation methods tested: Bayesian bandits

we **iterately update** our estimations using Bayes:



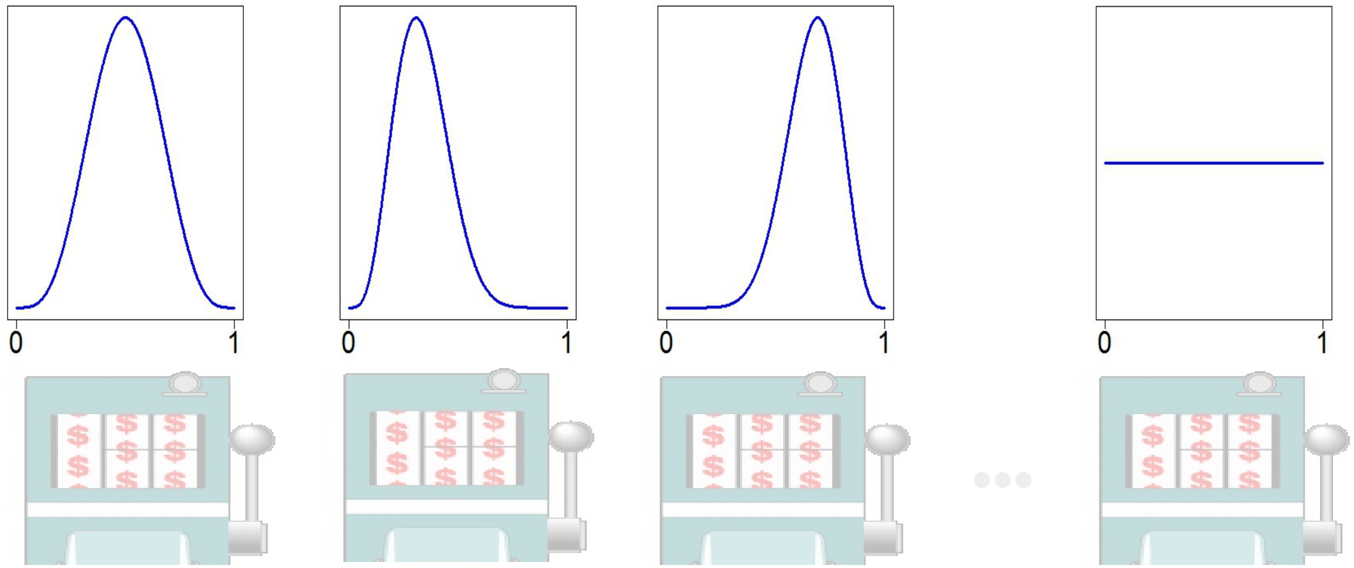
Allocation methods tested: Bayesian bandits

we **iteratively update** our estimations using Bayes:



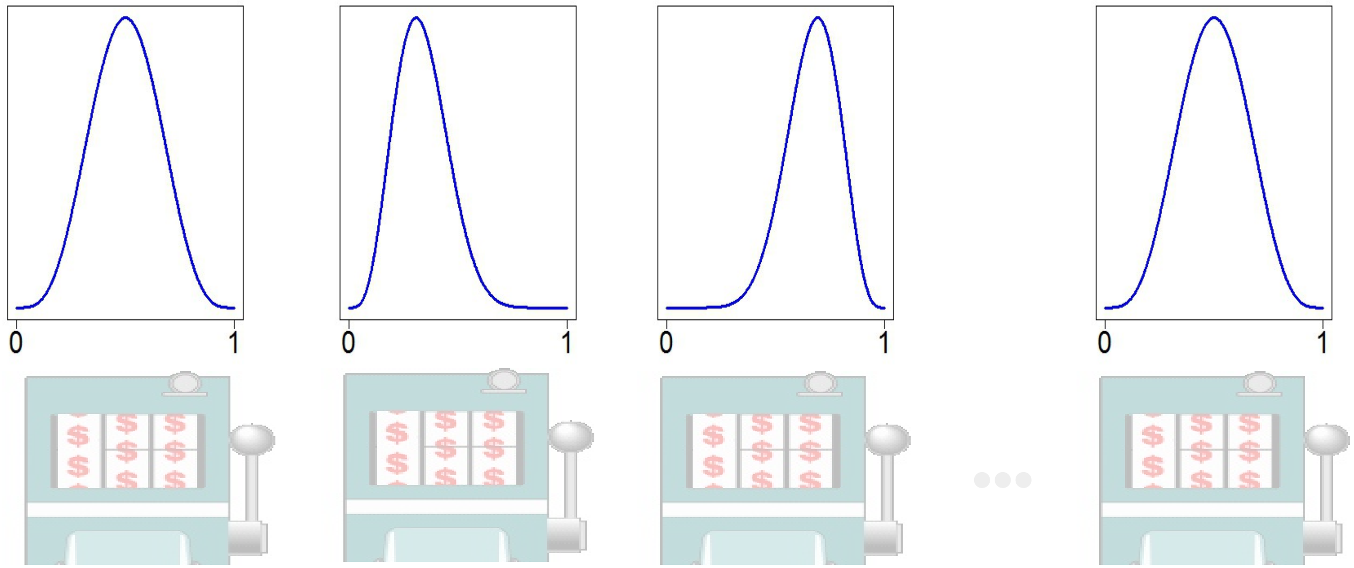
Allocation methods tested: Bayesian bandits

we **iteratively update** our estimations using Bayes:



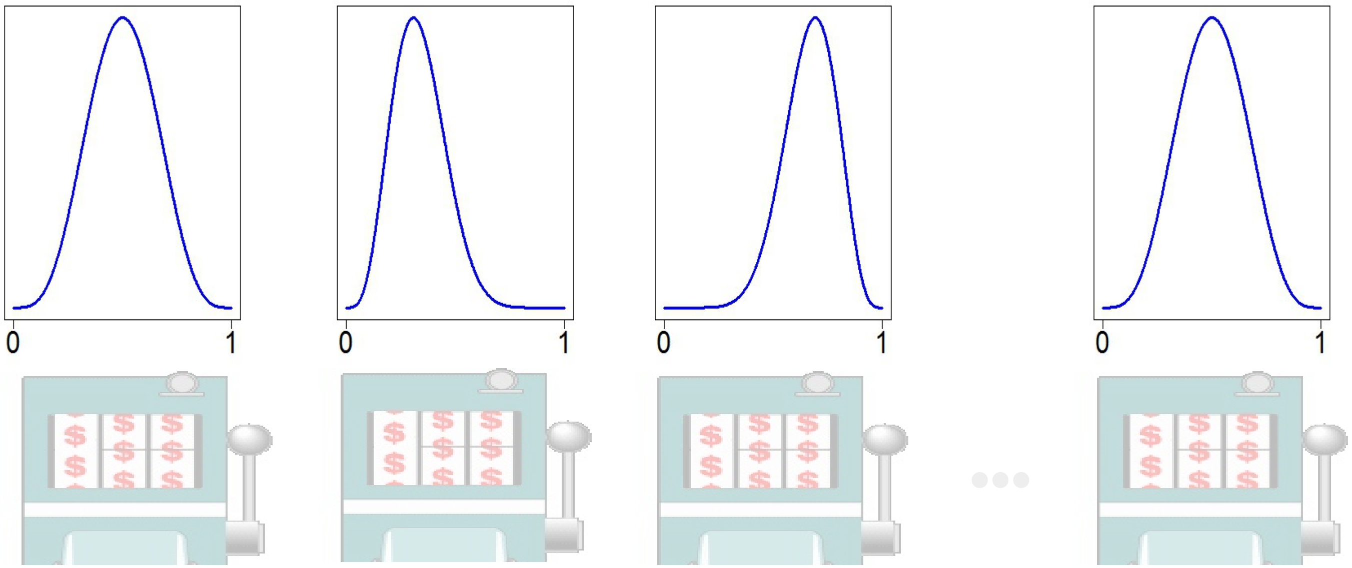
Allocation methods tested: Bayesian bandits

we **iteratively update** our estimations using Bayes:



Allocation methods tested: Bayesian bandits

we **iteratively update** our estimations using Bayes:

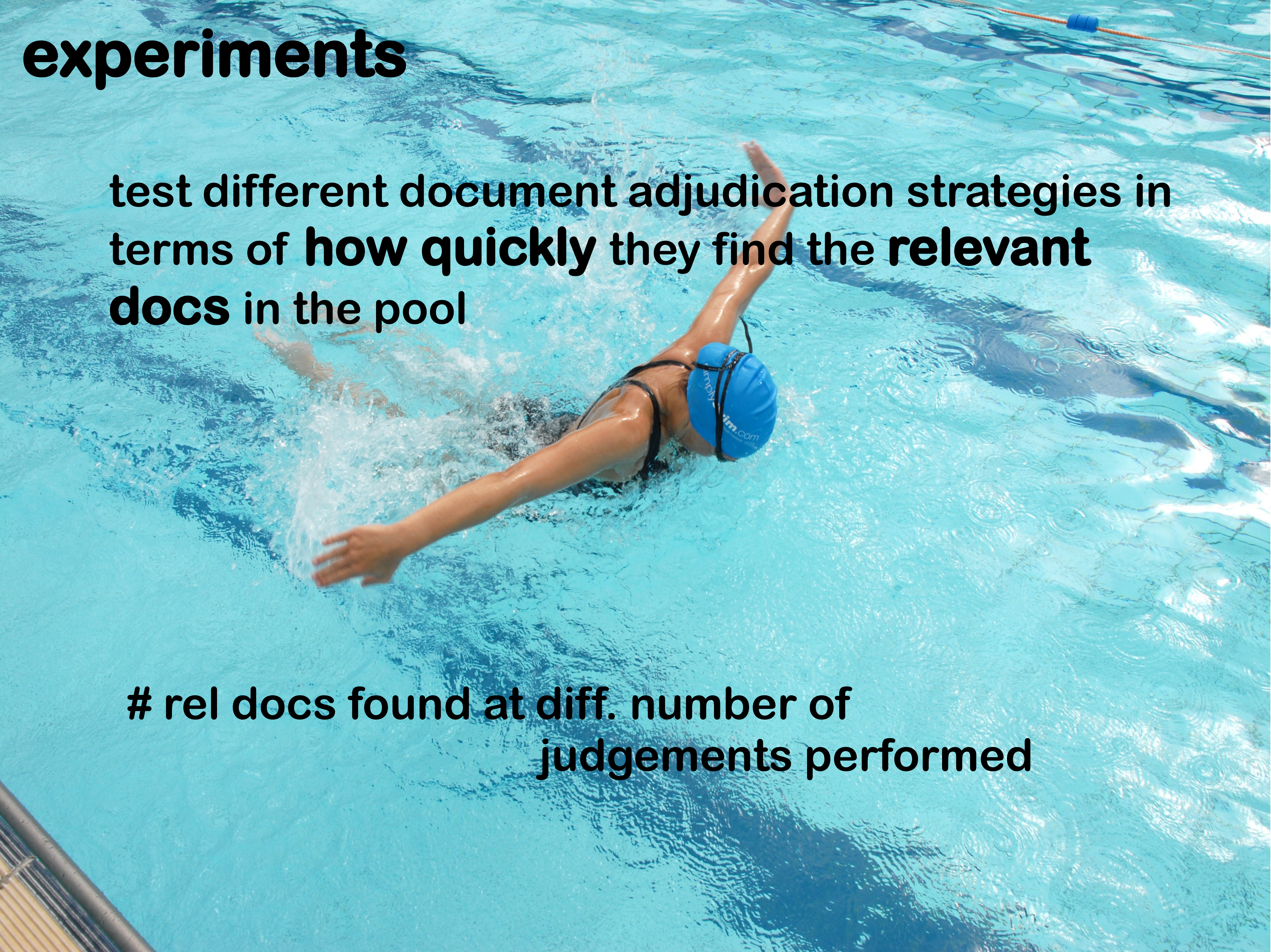


two strategies to select the **next machine**:

Bayesian Learning Automaton (BLA): draws a **sample from each the posterior** distribution and selects the machine yieding the **highest** sample

MaxMean (MM): selects the machine with the **highest** expectation of the posterior distribution

experiments



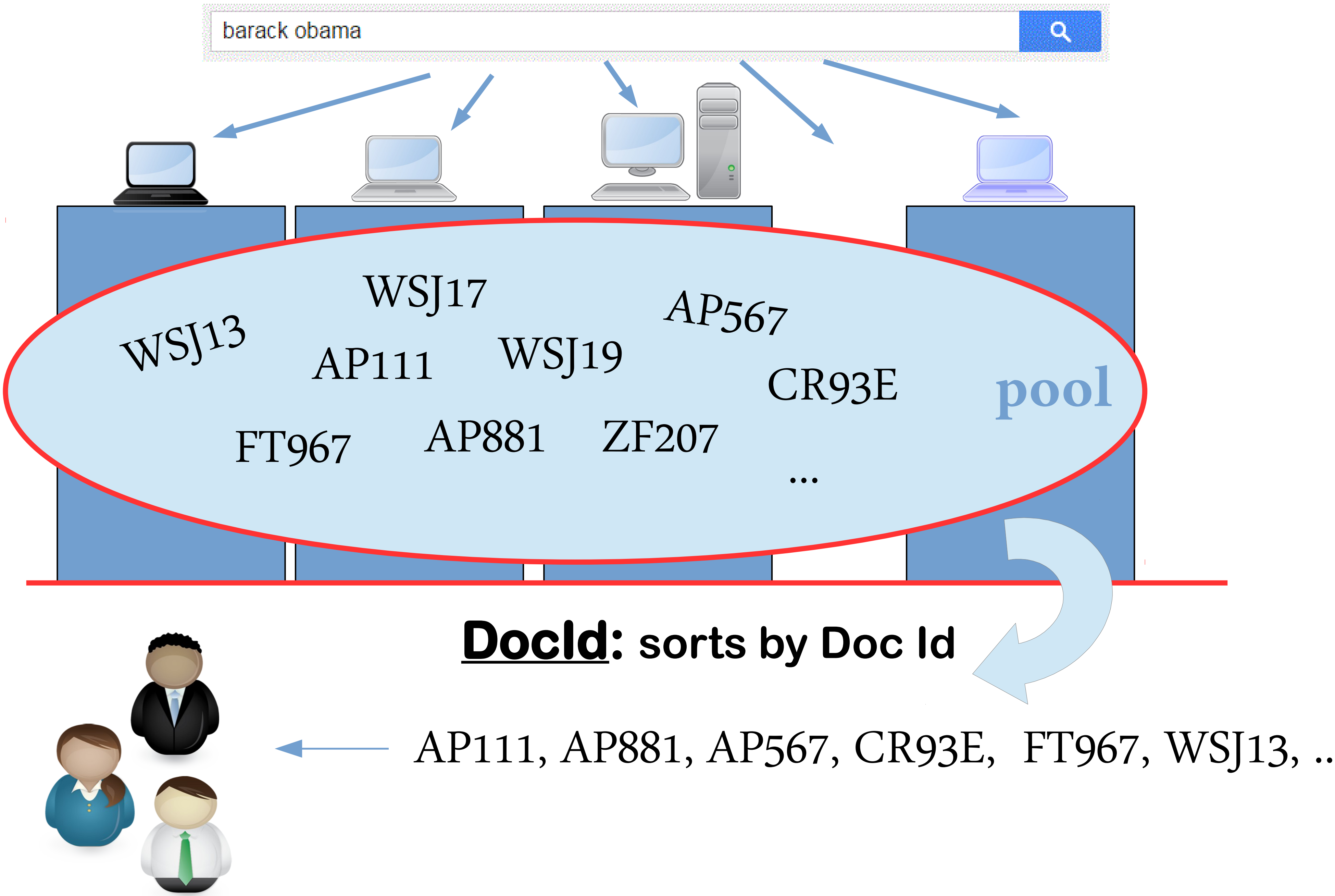
test different document adjudication strategies in terms of **how quickly** they find the **relevant docs** in the pool

\# rel docs found at diff. number of judgements performed

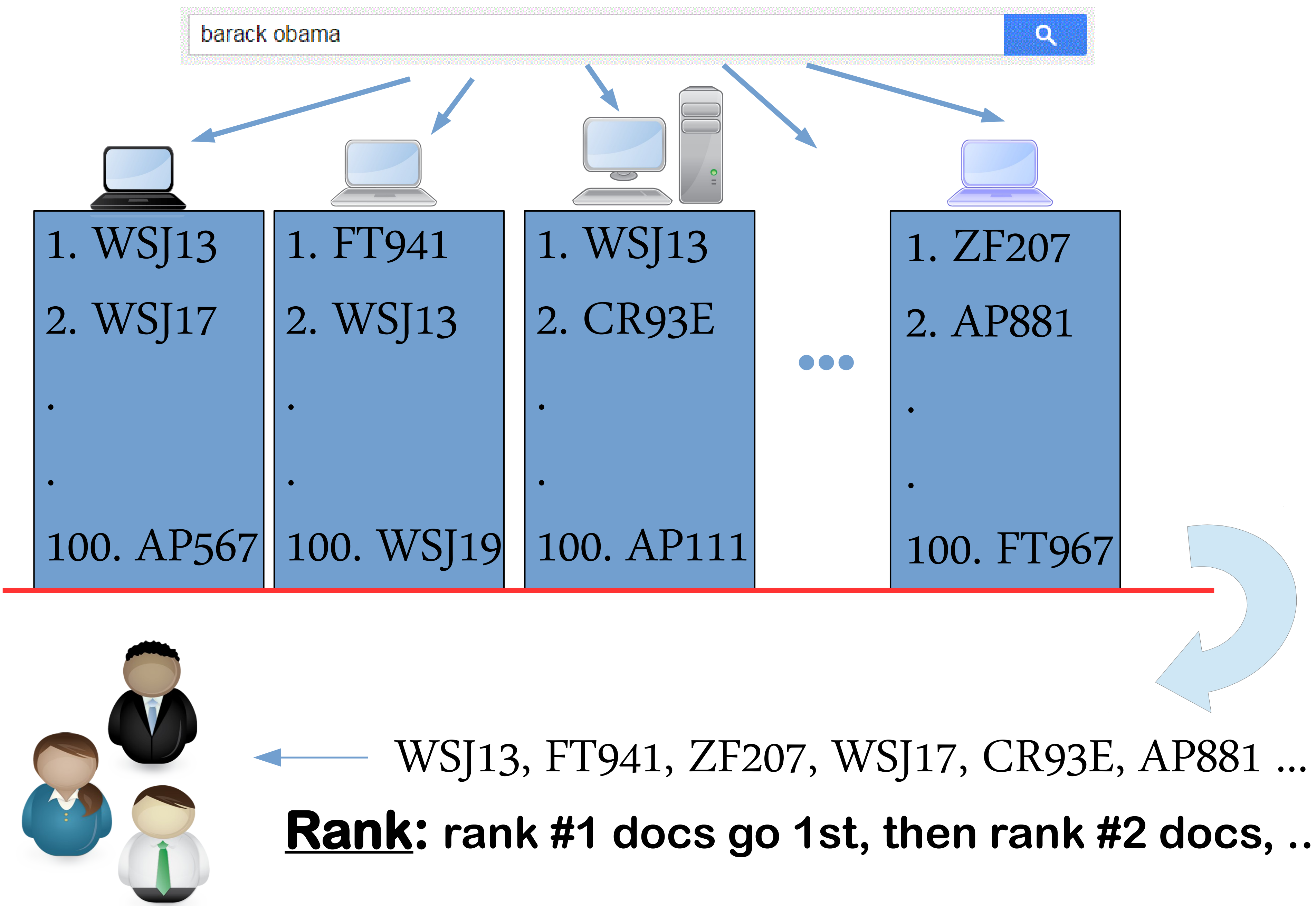
experiments: data

	TREC5	*TREC6*	*TREC7*	*TREC8*
# queries	50	50	50	50
# automatic runs	77	31	77	71
# manual runs	24	15	7	0
# assessed docs	133681	72270	80345	86830
avg. # docs judged per query	2673.6	1445.4	1606.9	1736.6
% of rel docs in the pool	4.1%	6.4%	5.8%	5.4%
avg. # rels per query	110.48	92.22	93.48	94.56

experiments: baselines



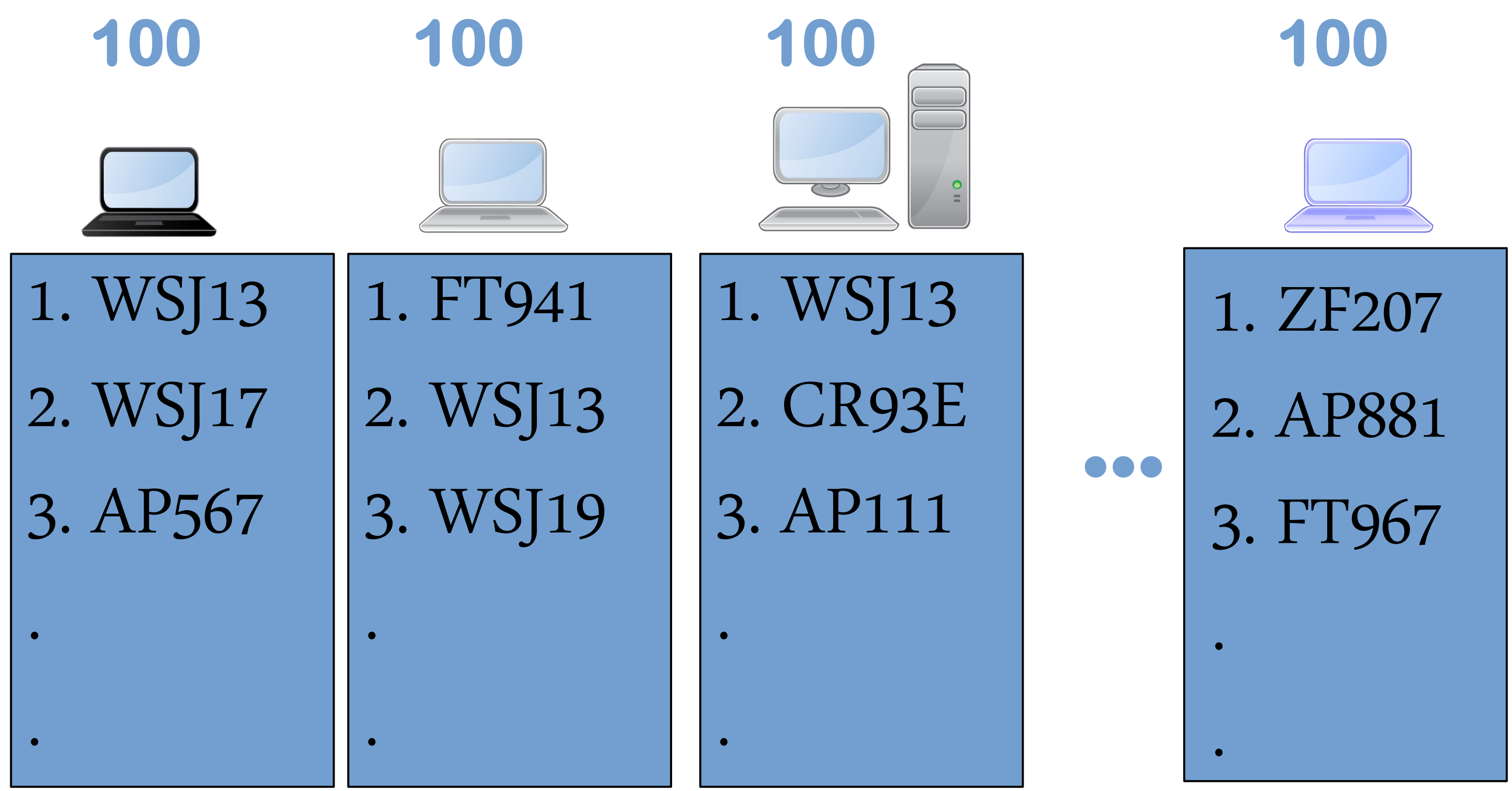
experiments: baselines



experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

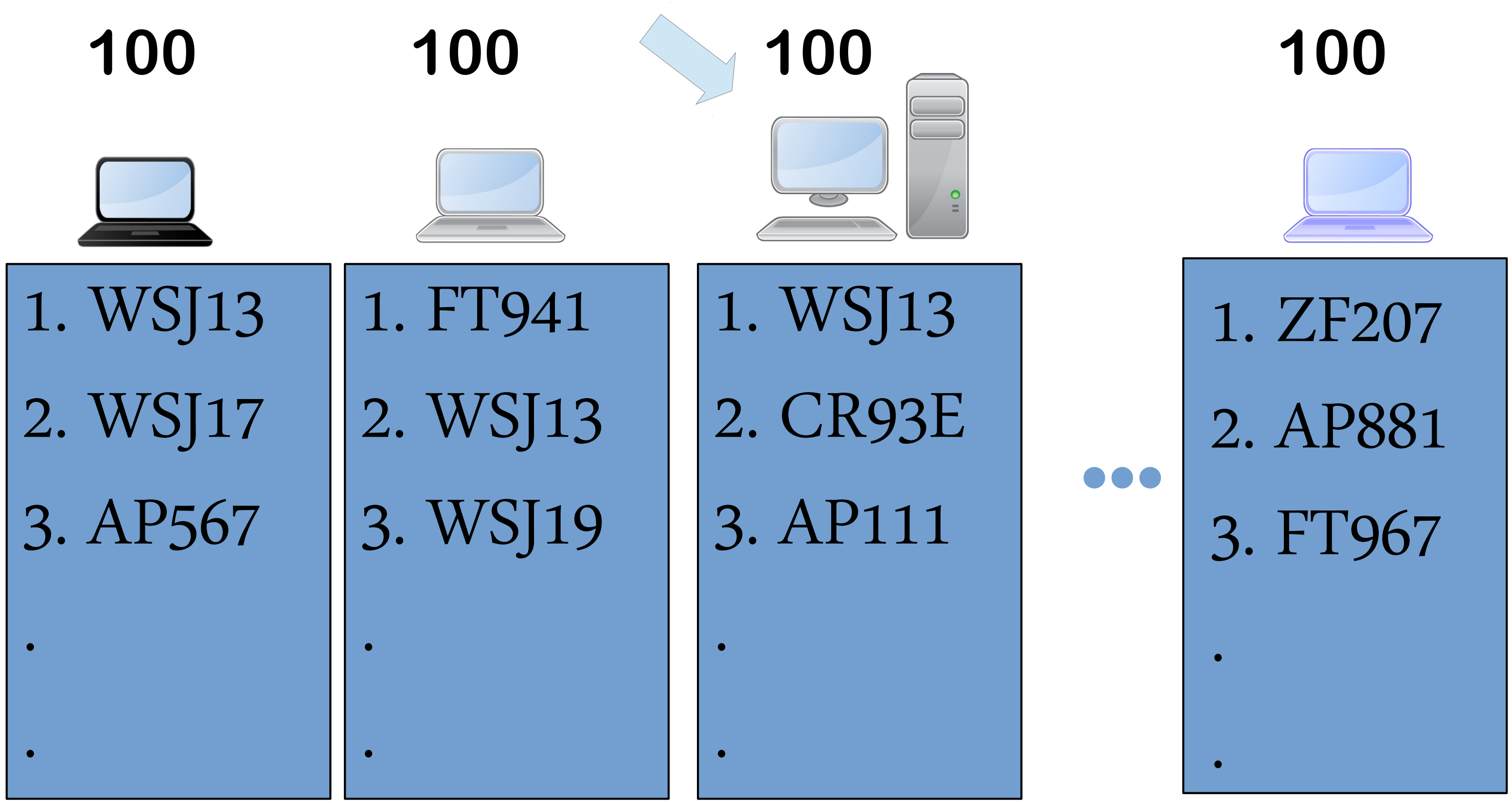
starts with uniform priorities for all runs (e.g. max priority=100)
selects a random run (from those with max priority)



experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

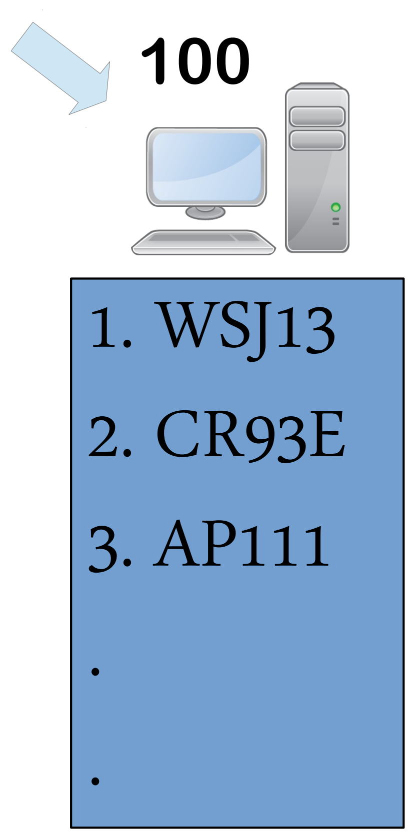
starts with uniform priorities for all runs (e.g. max priority=100)
selects a random run (from those with max priority)



experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

extracts & judges docs from the selected run
stays in the run until a non-rel doc is found

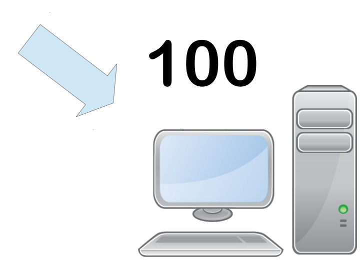


experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

extracts & judges docs from the selected run
stays in the run until a non-rel doc is found

100



WSJ13

1. ~~WSJ13~~
2. CR93E
3. AP111

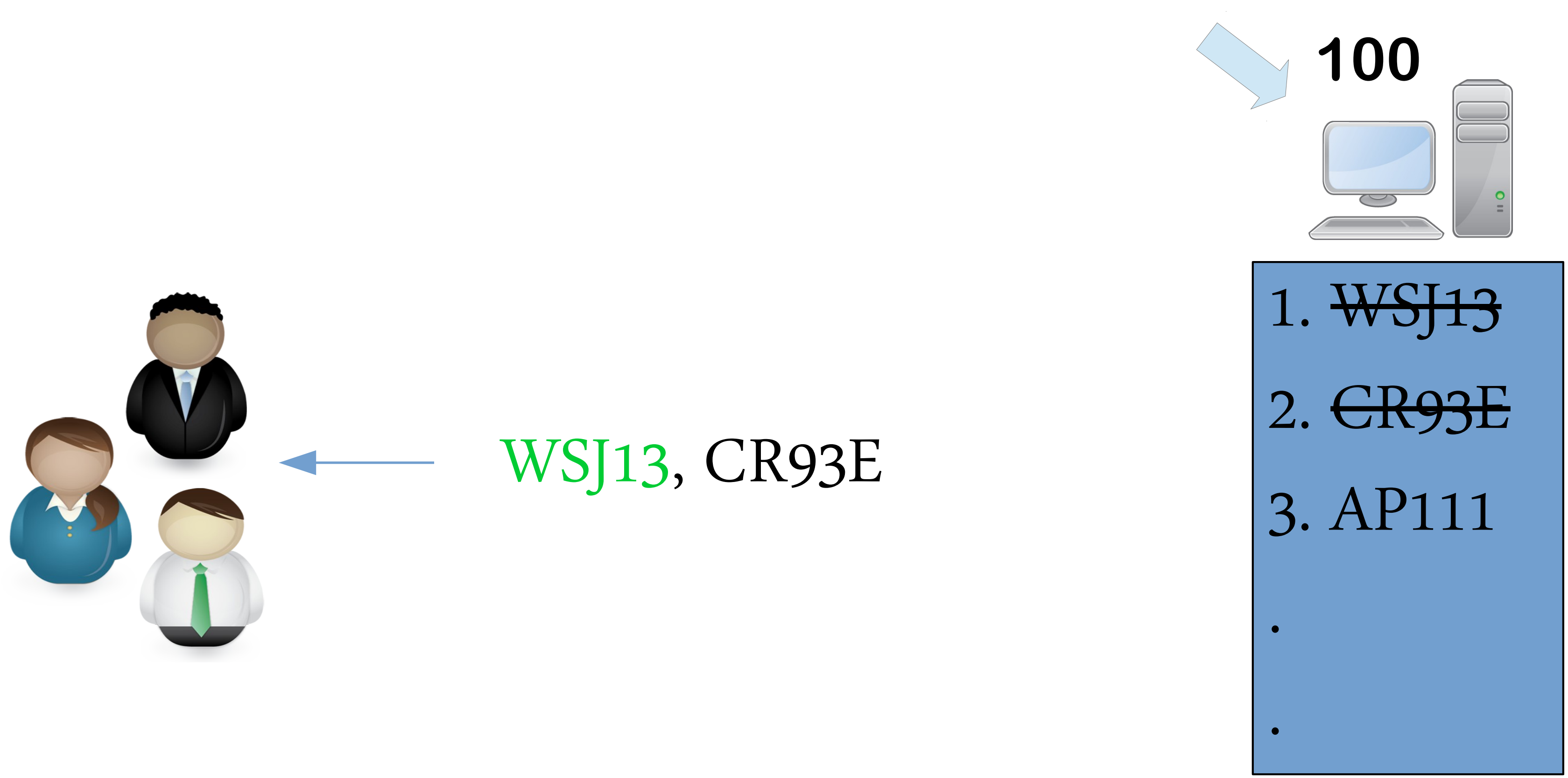
.

.

experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

extracts & judges docs from the selected run
stays in the run until a non-rel doc is found

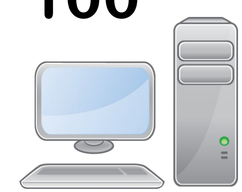


experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

extracts & judges docs from the selected run
stays in the run until a non-rel doc is found

100



1. ~~WSJ13~~
2. ~~CR93E~~
3. ~~AP111~~

.

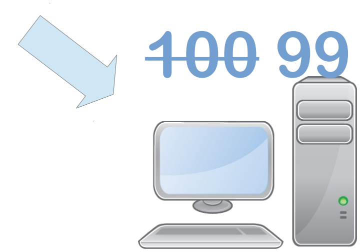
.

WSJ13, CR93E, AP111

experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

extracts & judges docs from the selected run
stays in the run until a non-rel doc is found
when a non-rel doc is found, priority is decreased



WSJ13, CR93E, AP111

1. ~~WSJ13~~
2. ~~CR93E~~
3. ~~AP111~~

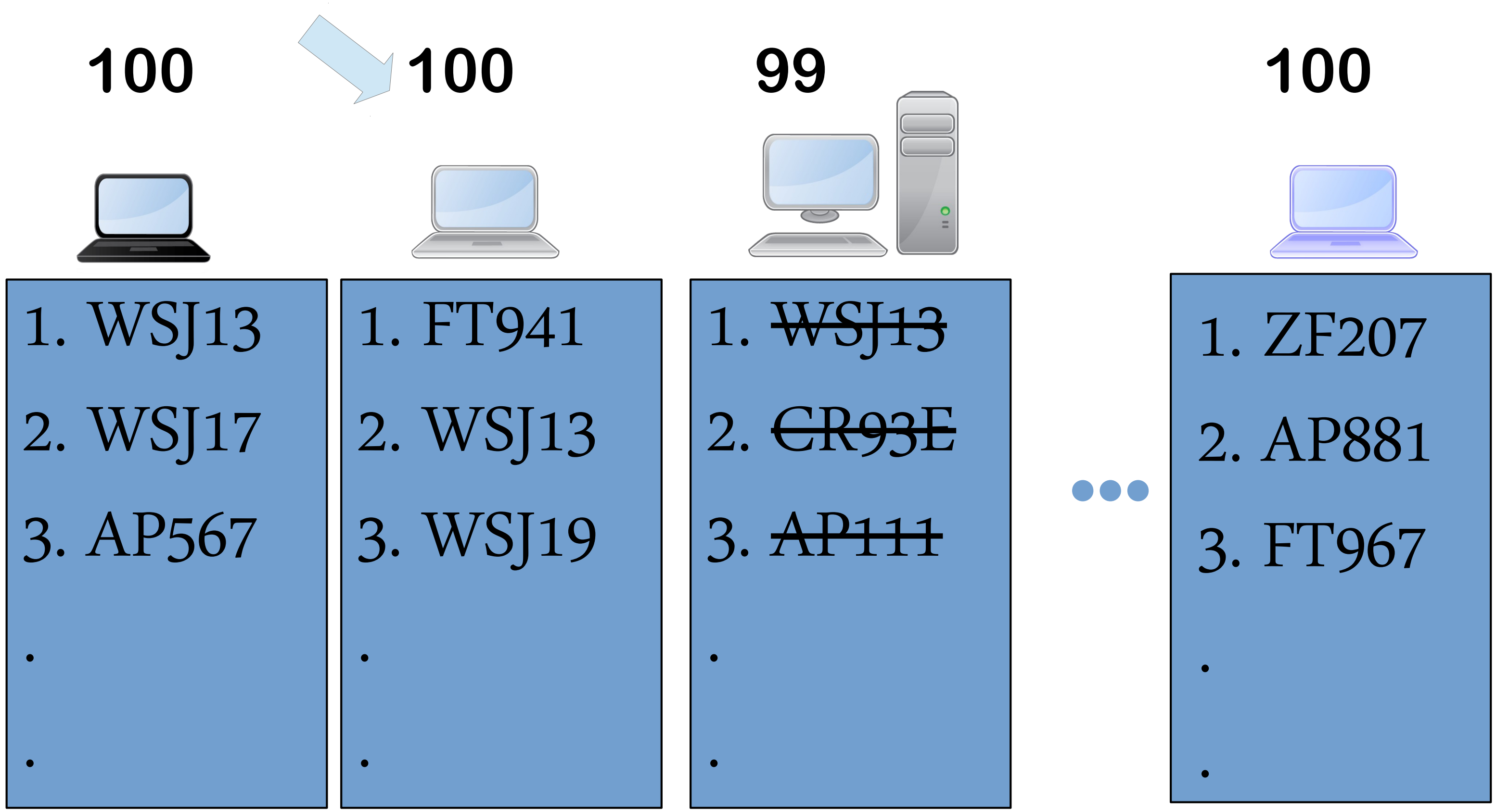
.

.

experiments: baselines

MoveToFront (MTF) (Cormack et al 98)

and we jump again to another max priority run



experiments: baselines

Moffat et al.'s method (A) (Moffat et al 2007)

based on rank-biased precision (RBP)
sums a rank-dependent score for each doc

score					
0.20	1. WSJ13	1. FT941	1. WSJ13	•••	1. ZF207
0.16	2. WSJ17	2. WSJ13	2. CR93E		2. AP881
0.13	3. AP567	3. WSJ19	3. AP111		3. FT967
.	.	.	.		.

experiments: baselines

Moffat et al.'s method (A) (Moffat et al 2007)

based on rank-biased precision (RBP)
sums a rank-dependent score for each doc

score					
0.20	1. **WSJ13**	1. FT941	1. **WSJ13**		1. ZF207
0.16	2. WSJ17	2. **WSJ13**	2. CR93E	•••	2. AP881
0.13	3. AP567	3. WSJ19	3. AP111		3. FT967
.	.	.	.		.

WSJ13: 0.20+0.16+0.20+...

all docs are ranked by decreasing accummulated score
and the ranked list defines the order in which docs are judged

experiments: baselines

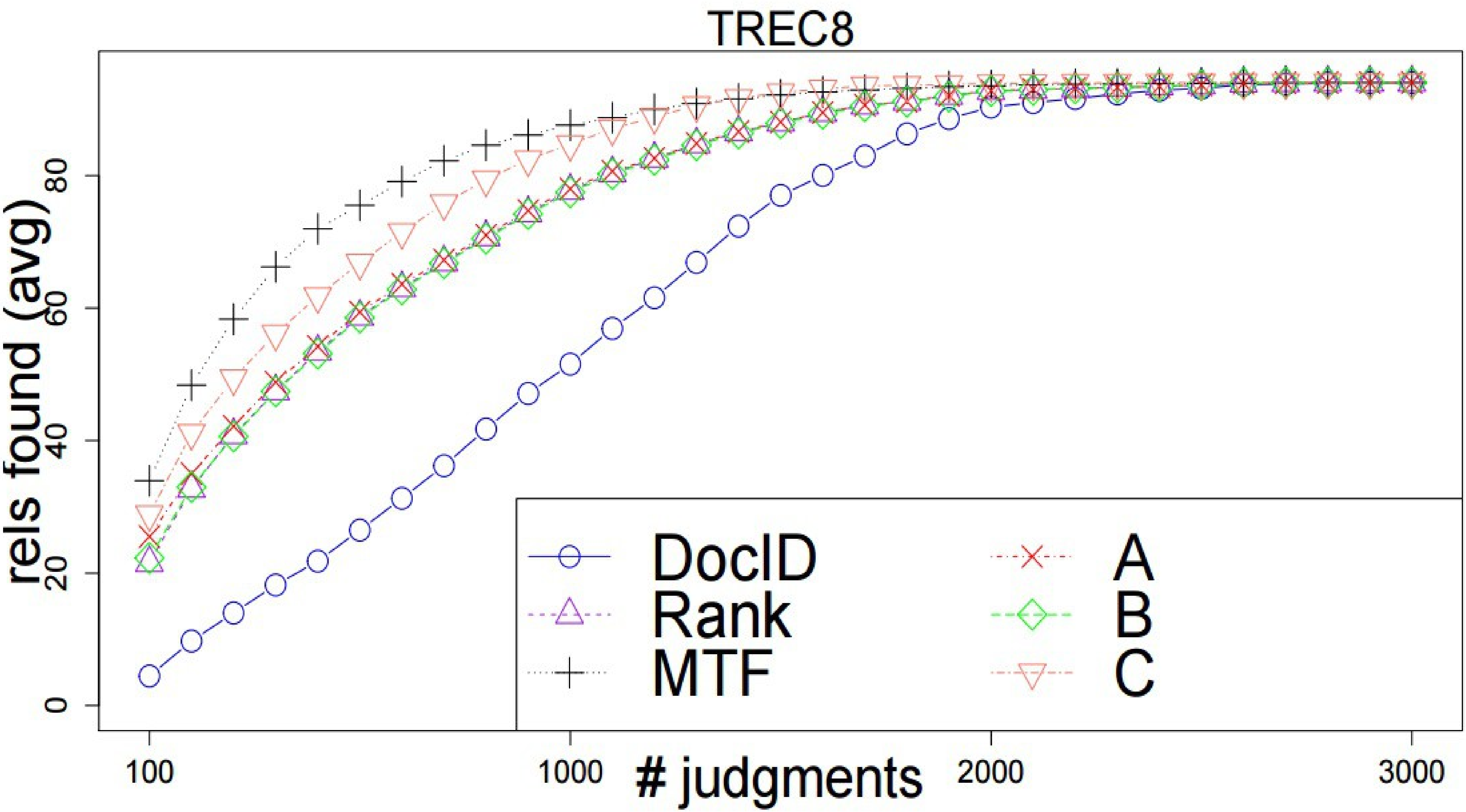
Moffat et al.'s method (B) (Moffat et al 2007)

evolution over A's method
considers not only the rank-dependent doc's contributions but also the runs' *residuals*
promotes the selection of docs from runs with many unjudged docs

Moffat et al.'s method (C) (Moffat et al 2007)

evolution over B's method
considers not only the rank-dependent doc's and the residuals
promotes the selection of docs from effective runs

experiments: baselines



MTF: best performing baseline

experiments: MTF vs bandit-based models

Method	Number of judgements 100	300	500	700	900	1100	2000	all
				TREC5				
MTF	27.66	49.5	63.54	72.04	78.64	84.58	99.58	**109.9**
BLA	23.46	45.1	59.36	69.82	77.34	82.7	97.58	**109.9**
MM	**27.76**	**53.54**	**68.18**	**77.96**	**84.18**	**88.74**	**101.42**	**109.9**
RANDOM	20.94	41.48	53.38	63.42	71.28	76.36	93.92	**109.9**
UCB	20.82	46.42	58.56	67.5	74.44	79.8	96.4	**109.9**
ϵ_n-GREEDY	21.1	46.7	59.96	69.48	76.82	82.26	96.88	**109.9**
				TREC6				
MTF	31.96	55.7	66.68	75.82	82.04	86.28	**91.8**	
BLA	24.94	46.42	60.62	70.4	78.84	84.56	**91.8**	
MM	**32.12**	**56.1**	**68.98**	**78.06**	**83.16**	**87.24**	**91.8**	
RANDOM	25.56	46.9	59.98	69.7	77.4	83.82	**91.8**	
UCB	27.86	48.96	62.36	71.38	79.14	84.9	**91.8**	
ϵ_n-GREEDY	27.6	50.8	63.12	71.84	78.8	84.44	**91.8**	
				TREC7				
MTF	**35**	**58.04**	**70.58**	**78.52**	**83.48**	**86.94**	**92.7**	**92.84**
BLA	27.64	49.8	62.3	71.42	78.3	83.16	91.58	**92.84**
MM	34.62	56.4	70	78.18	83	86.36	92.4	**92.84**
RANDOM	27.48	50.74	62.86	71.86	78.44	83.44	91.3	**92.84**
UCB	30.32	52.4	64.44	72.44	79.32	83.68	91.06	**92.84**
ϵ_n-GREEDY	28	53.8	65.06	73.54	79.22	83.02	91.2	**92.84**
				TREC8				
MTF	34.06	58.48	71.78	79.22	84.5	87.58	93.22	**94.04**
BLA	27.14	50.42	64.9	73.96	80.36	84.96	93.36	**94.04**
MM	**34.4**	**59.34**	**72.9**	**80.82**	**85.56**	**88.8**	**93.54**	**94.04**
RANDOM	26.94	50.58	63.9	72.58	79.28	83.48	92.58	**94.04**
UCB	29.86	52.9	65.92	74.06	80.52	85.12	93.4	**94.04**
ϵ_n-GREEDY	28.16	53.66	66.76	74.88	80.6	84.88	93.2	**94.04**

experiments: MTF vs bandit-based models

	Number of judgements							
Method	100	300	500	700	900	1100	2000	all
				TREC5				
MTF	27.66	49.5	63.54	72.04	78.64	84.58	99.58	**109.9**
BLA	23.46	45.1	59.36	69.82	77.34	82.7	97.58	**109.9**
MM	**27.76**					**88.74**	**101.42**	**109.9**
RANDOM							93.92	**109.9**
UCB							96.4	**109.9**
ϵ_n-GREE								**109.9**
MT								
B								
M								**.84**
BL								**2.84**
MM								**92.84**
RAND								**92.84**
UCB							.06	**92.84**
ϵ_n-GREEDY							91.2	**92.84**
MTF	34.06	58.48				87.58	93.22	**94.04**
BLA	27.14	50.42	64.9	73.96	80.36	84.96	93.36	**94.04**
MM	**34.4**	**59.34**	**72.9**	**80.82**	**85.56**	**88.8**	**93.54**	**94.04**
RANDOM	26.94	50.58	63.9	72.58	79.28	83.48	92.58	**94.04**
UCB	29.86	52.9	65.92	74.06	80.52	85.12	93.4	**94.04**
ϵ_n-GREEDY	28.16	53.66	66.76	74.88	80.6	84.88	93.2	**94.04**

Random: **weakest** approach

BLA/UCB/ϵ$_n$-greedy are **suboptimal**
(sophisticated exploration/exploitation trading not needed)

MTF and MM: best performing methods

improved bandit-based models

MTF: forgets quickly about past rewards
(a single non-relevance doc triggers a jump)

non-stationary

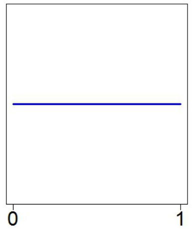
bandit-based
solutions:

not all historical rewards count the same

MM-NS and BLA-NS non-stationary variants of MM and BLA

stationary bandits

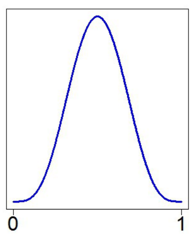
$\mathrm{Beta}(\alpha,\beta)$, $\alpha,\beta=1$



rel docs add 1 to α
non-rel docs add 1 to β

(after n iterations)

$\mathrm{Beta}(\alpha_n,\beta_n)$



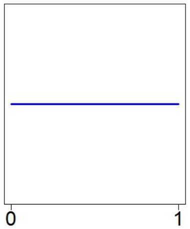
$$\alpha_n = 1 + jrel_s$$
$$\beta_n = 1 + jret_s - jrel_s$$

$jrel_s$: # judged relevant docs (retrieved by s)
$jret_s$: # judged docs (retrieved by s)

all judged docs count the same

non-stationary bandits

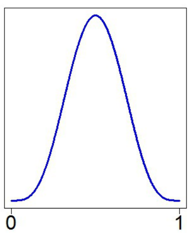
$\mathrm{Beta}(\alpha,\beta)$, $\alpha,\beta=1$



$jrel_s = rate * jrel_s + rel_d$
$jret_s = rate * jret_s + 1$

(after n iterations)

$\mathrm{Beta}(\alpha_n,\beta_n)$



$$\alpha_n = 1 + jrel_s$$
$$\beta_n = 1 + jret_s - jrel_s$$

rate>1: weights more early relevant docs
rate<1: weights more late relevant docs
rate=0: only the last judged doc counts
(BLA-NS, MM-NS)
rate=1: stationary version

experiments: improved bandit-based models

	Number of judgements							
Method	100	300	500	700	900	1100	2000	all
TREC5								
MTF	27.66	49.5	63.54	72.04	78.64	84.58	99.58	**109.9**
BLA	23.46	45.1	59.36	69.82	77.34	82.7	97.58	**109.9**
BLA-NS	22.8	44.8	58.74	68.78	76.04	81.44	97.4	**109.9**
MM	27.76	53.54	68.18	77.96	84.18	88.74	**101.42**	**109.9**
MM-NS	**30**	**56.76**	**70.96**	**79.06**	**85.18**	**89.42**	101.32	**109.9**
TREC6								
MTF	31.96	55.7	66.68	75.82	82.04	86.28	**91.8**	
BLA	24.94	46.42	60.62	70.4	78.84	84.56	**91.8**	
BLA-NS	25.58	46.68	60.44	70	76.7	83.16	**91.8**	
MM	32.12	56.1	68.98	**78.06**	83.16	87.24	**91.8**	
MM-NS	**33.5**	**58.2**	**69.72**	77.9	**83.48**	**87.44**	**91.8**	
TREC7								
MTF	35	58.04	70.58	78.52	83.48	86.94	**92.7**	**92.84**
BLA	27.64	49.8	62.3	71.42	78.3	83.16	91.58	**92.84**
BLA-NS	27.8	50.14	62.54	70.34	76.92	81.56	90.28	**92.84**
MM	34.62	56.4	70	78.18	83	86.36	92.4	**92.84**
MM-NS	**36.8**	**62.42**	**74.42**	**81.74**	**85.82**	**88.32**	92.58	**92.84**
TREC8								
MTF	34.06	58.48	71.78	79.22	84.5	87.58	93.22	**94.04**
BLA	27.14	50.42	64.9	73.96	80.36	84.96	93.36	**94.04**
BLA-NS	27.12	49.5	63.68	72.06	77.56	82.54	92.42	**94.04**
MM	34.4	59.34	72.9	80.82	85.56	88.8	93.54	**94.04**
MM-NS	**36.96**	**64.62**	**77.3**	**82.5**	**86.34**	**89.2**	**93.6**	**94.04**

conclusions

multi-arm bandits: **formal** & **effective** framework for doc adjudication in a pooling-based evaluation

it's good to react quickly to non-relevant docs
(non-stationary variants)

it's not good to increasingly reduce exploration
(UCB, ϵ_n-greedy)

future work

query-related variabilities

metasearch

hierarchical bandits

stopping criteria

reproduce our experiments & test new ideas!

http://tec.citius.usc.es/ir/code/pooling_bandits.html

(our R code, instructions, etc)

Pooling - Bandits

This document provides all details needed to reproduce the experiments reported in the paper **D. Losada, J. Parapar, A. Barreiro. "Feeling Lucky? Multi-armed Bandits for Ordering Judgements in Pooling-based Evaluation". ACM Symposium on Applied Computing, 2016**.

Any scientific publication derived from the use of this software should explicitly refer to this ACM SAC paper.

Next, we explain the data used for experimentation and **provide our R code**, which implement all pooling strategies.

Data

We used four TREC collections (http://trec.nist.gov): TREC5, TREC6, TREC7 and TREC8.

NIST kindly provided the runs that contributed to the pools of the adhoc tasks of TREC5, TREC6, TREC7 and TREC8 (http://trec.nist.gov/data/intro_eng.html).

The pooled runs are archived by NIST within a password protected area. If you want to reproduce our experiments you need to request access to the protected area (follow the instructions given at http://trec.nist.gov/results.html).

- **TREC5**
  - 101 runs in the pool (77 adhoc + 24 other).

    The **77 adhoc runs** are: input.anu5aut1 input.anu5aut2 input.anu5man4 input.anu5man6 input.brkly15 input.brkly16 input.brkly17 input.brkly18 input.city96a1 input.city96a2 input.CLCLUS input.CLTHES input.colm1 input.colm4 input.Cor5A1se input.Cor5A2cr input.Cor5M1le input.Cor5M2rf input.Ctifr1 input.Ctifr2 input.DCU961 input.DCU962 input.DCU963 input.DCU964 input.DCU969 input.DCU96C input.DCU96D input.erliA1 input.ETHal1 input.ETHas1 input.ETHme1 input.fsclt3 input.fsclt4 input.genrl1 input.genrl2 input.genrl3 input.genrl4 input.glair4 input.gmu96au1 input.gmu96au2 input.gmu96ma1 input.gmu96ma2 input.ibmgd1 input.ibmgd2 input.ibmge1 input.ibmge2 input.ibms96a input.ibms96b input.INQ301 input.INQ302 input.KUSG2 input.KUSG3 input.LNaDesc1 input.LNaDesc2 input.LNmFull1 input.LNmFull2 input.mds001 input.mds002 input.mds003 input.Mercure-al input.Mercure-as input.MONASH input.pircsAAL input.pircsAAS input.pircsAM1 input.pircsAM2 input.sdmix1 input.sdmix2 input.umcpa1 input.uncis1 input.uncis2 input.UniNE7 input.UniNE8 input.uwgcx0 input.uwgcx1 input.vtwnA1 input.vtwnB1

    The **other 24 runs** are: input.anu5mrg0 input.anu5mrg1 input.anu5mrg7 input.CLATMC input.CLATMN input.CLPHR0 input.CLPHR1 input.CLPHR2 input.fsclt3m input.genlp1 input.genlp2 input.genlp3 input.genlp4 input.MTRa961 input.sbase1 input.sbase2 input.UniNE0 input.UniNE9 input.xerox_nlp1 input.xerox_nlp2 input.xerox_nlp3 input.xerox_nlp4 input.xerox_nlp5 input.xerox_nlp6

Feeling Lucky? Multi-armed bandits for Ordering Judgements in Pooling-based Evaluation

David E. Losada

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Centro Singular de Investigación en Tecnoloxías da Información

Javier Parapar, Álvaro Barreiro

UNIVERSIDADE DA CORUÑA

Acknowledgements:

MsSaraKelly. picture pg 1 (modified). CC BY 2.0.
Sanofi Pasteur. picture pg 2 (modified). CC BY-NC-ND 2.0.
pedrik. picture pgs 3-5. CC BY 2.0.
Christa Lohman. picture pg 3 (left). CC BY-NC-ND 2.0.
Chris. picture pg 4 (tag cloud). CC BY 2.0.
Daniel Horacio Agostini. picture pg 5 (right). CC BY-NC-ND 2.0.
ScaarAT. picture pg 14. CC BY-NC-ND 2.0.
Sebastien Wiertz. picture pg 15 (modified). CC BY 2.0.
Willard. picture pg 16 (modified). CC BY-NC-ND 2.0.
Jose Luis Cernadas Iglesias. picture pg 17 (modified). CC BY 2.0.
Michelle Bender. picture pg 25 (left). CC BY-NC-ND 2.0.
Robert Levy. picture pg 25 (right). CC BY-NC-ND 2.0.
Simply Swim UK. picture pg 37. CC BY-SA 2.0.
Sarah J. Poe. picture pg 55. CC BY-ND 2.0.
Kate Brady. picture pg 58. CC BY 2.0.
August Brill. picture pg 59. CC BY 2.0.

This work was supported by the "Ministerio de Economía y Competitividad" of the Goverment of Spain and FEDER Funds under research projects TIN2012-33867 and TIN2015-64282-R.

GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

Unión Europea

Fondo Europeo de Desarrollo Regional
"Una manera de hacer Europa"