

KVASIR Project – Knowledge ViA Semantic Information Retrieval

Liang Wang, Jussi Kangasharju

Department of Computer Science, liang.wang@cs.helsinki.fi

Future Content-Aware Network



- Motivation Content consumers only care about "what they need".
- **Question** How to provide smart content provision and recommendation?
- **Objective** Content are interconnected semantically in the network.

Information-Centric Architecture

- Access content by name instead of location.
- Push popular content to network edge.
- Open network interface to applications.

Research Challenges

- How to efficiently reduce the dimensionality?
- How to perform fast search in high-dimensional space?
- How to effectively minimize the index file size?
- How to smartly recommend based on one-class SVM?

Big Data Framework

- Fast proccess large text corpus.
- Accelerate design \rightarrow develop \rightarrow deploy cycle.
- Cope with fast rate of user-generated content.

Machine Learning Techniques

- Latent Semantic Analysis \rightarrow Content patterns.
- Random projection \rightarrow Accuracy vs. Efficiency.
- Approx. nearest neighbor \rightarrow Fast search.

Flexible Frontend Design

WIKIPEDIA	Dinosaur
The Free Encyclopedia	From Wikipedia, the free encyclopedia
Main page Contents	For other uses, see Dinosaur (disambiguation).
Featured content	Dinosaurs are a diverse group of animals of the clade Dinosauria. They fir
	, from the beginning of
WIKIPEDIA	ene extinction event le
	od dinosaurs during the
93 Therop	oda W sutingting sugget that as

Join Us! If You Want To ...

- Develop a high performance server backend ...
- Contribute to the frontend enhancement ...
- Learn how to build a realistic Internet service ...
- See how to apply machine learning in the wild ...
- Write an interesting master thesis ...

IN CONTRACTOR

Image: With the second seco
Image: Second state Image: Second state
Image: Second state of the second s
w or crests are common to that egg laying and nest light, many prehistoric d neters (30 feet 4 inches) sturdy bones are more
that egg laying and nest light, many prehistoric d neters (30 feet 4 inches) sturdy bones are more
hers around the dinosaur tree
he name is somewhat m
os discovered in Venezuela 🚺 it characteristics traditio
ops Discovered in Venezuela
babysitter' discovered I and cold-blooded. Mos numerous adaptations for

50	moorue	in son	r dat	a se	
88	Quicks	ort	W ted p	ositi	
81	Sorting	algorithm	W es a	cons	
80	Bubble	sort	W tas	it rec	
78	Self-or	ganizing list	W	ethin	
77	Best, w	vorst and average case	W		
77	Double	Double-ended priority queue			
75	Pairwis	e summation	W		
74	Cycle s	sort	W	I C	
NEWS					
JUL 16	Take m	ne off the terrorist watch li	st! <mark>CNN</mark>		
iguages	92	6	lming	g one	
العر بي		from the input data,	tinds the loca	ation	
Azərbaycanca	Sorting is typically done in-place, by i				

Insertion sort

From Wikipedia, the free encyclopedia

Insertion sort is a simple sorting algo

advanced algorithms such as quicksc

W

Bentley s

data set

than mo:

The Free

Main page Contents Featured

WIKIPEDIA The Free Encyclopedia

Main page

Contents

Featured conten

Et property

WIKIPEDIA

Insertion sort

KVASIR

		 Please 	se co	ontact	us →		
IPEDI Encyclope	A	Lamport's	bake	ry algorit	hm		
e content				It has been sugg	ested that Lampor	rt's Distribu	ted Mutual Exclusion Alg
PEDIA					too o lint of rotorou	ooo_roloto	citations. (December 201
90	Lampor	rt's bakery algorithm				W	lust Leslie Lamport, whi
65	Suzuki-Kasami algorithm				W	the same resources. Da	
65	Peterson's algorithm				W	kery algorithm is one of	