## CREST Open Workshop: Code Provenance and Clone Detection

23 November 2010



## Cloned COW: COWs are Type 3 Clones





## Cloned COW: COWs are Type 3 Clones

# CREST CREST







admin





- admin
- 4 faculty







- admin
- 4 faculty
- 8 post docs







- l admin
- 4 faculty
- 8 post docs
- PhD students







- admin
- 4 faculty
- 8 post docs
- PhD students
- I-4 resident visiting scholars



CREST

#### Research



#### Research

Service Oriented Computing Search Based Software Engineering Clone Detection Dependence Analysis Quantitative Information Flow Digital Humanities Testing

Requirement Engineering

Slicing





#### Research

Service Oriented Computing Search Based Software Engineering Cone Detection Dependence Analysis Quantitative Information Flow

**Digital Humanities** 

Testing

**Requirement Engineering** 

Slicing





## Funding



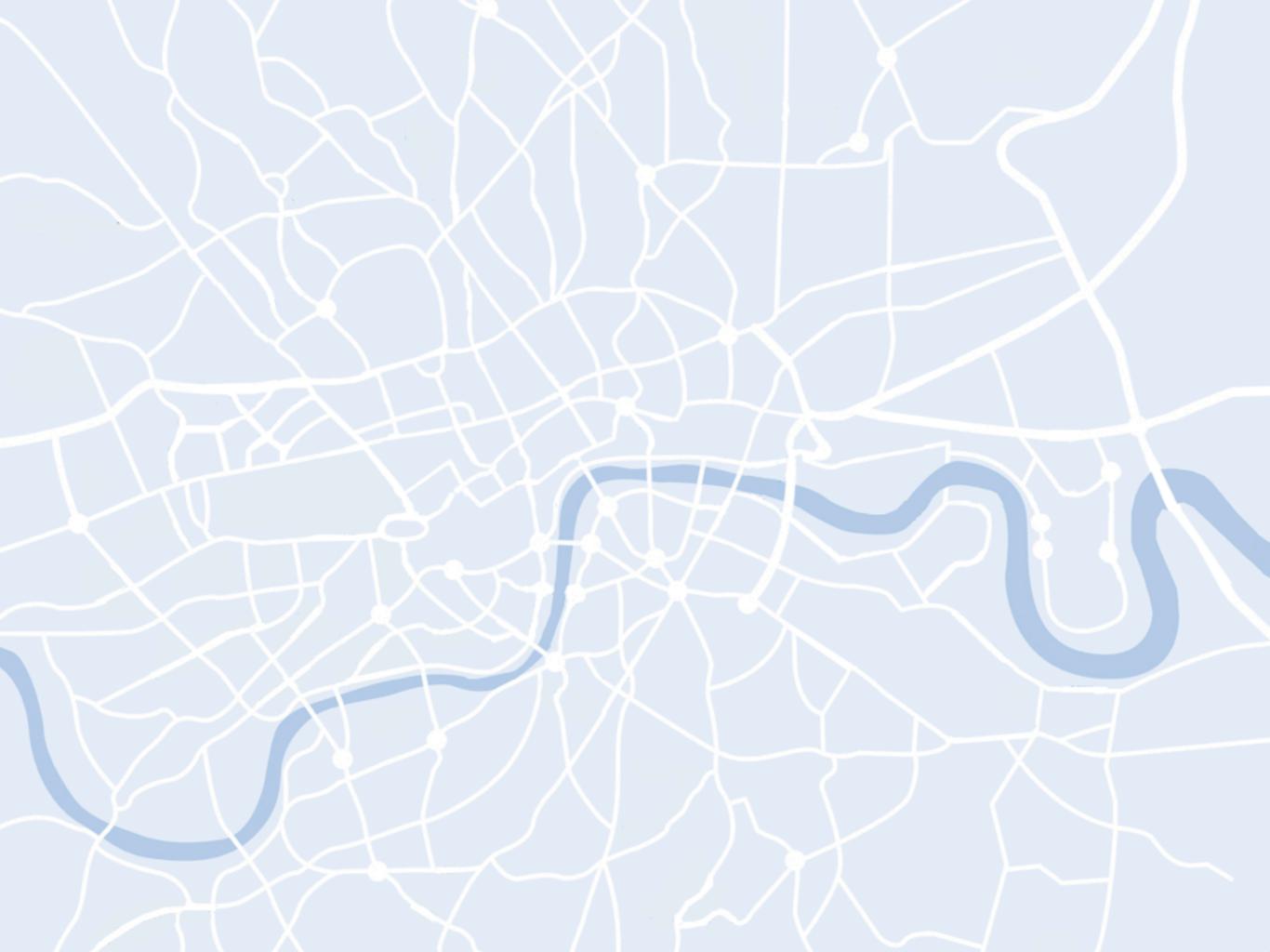
Engineering and Physical Sciences Research Council











#### University College London

ŵ

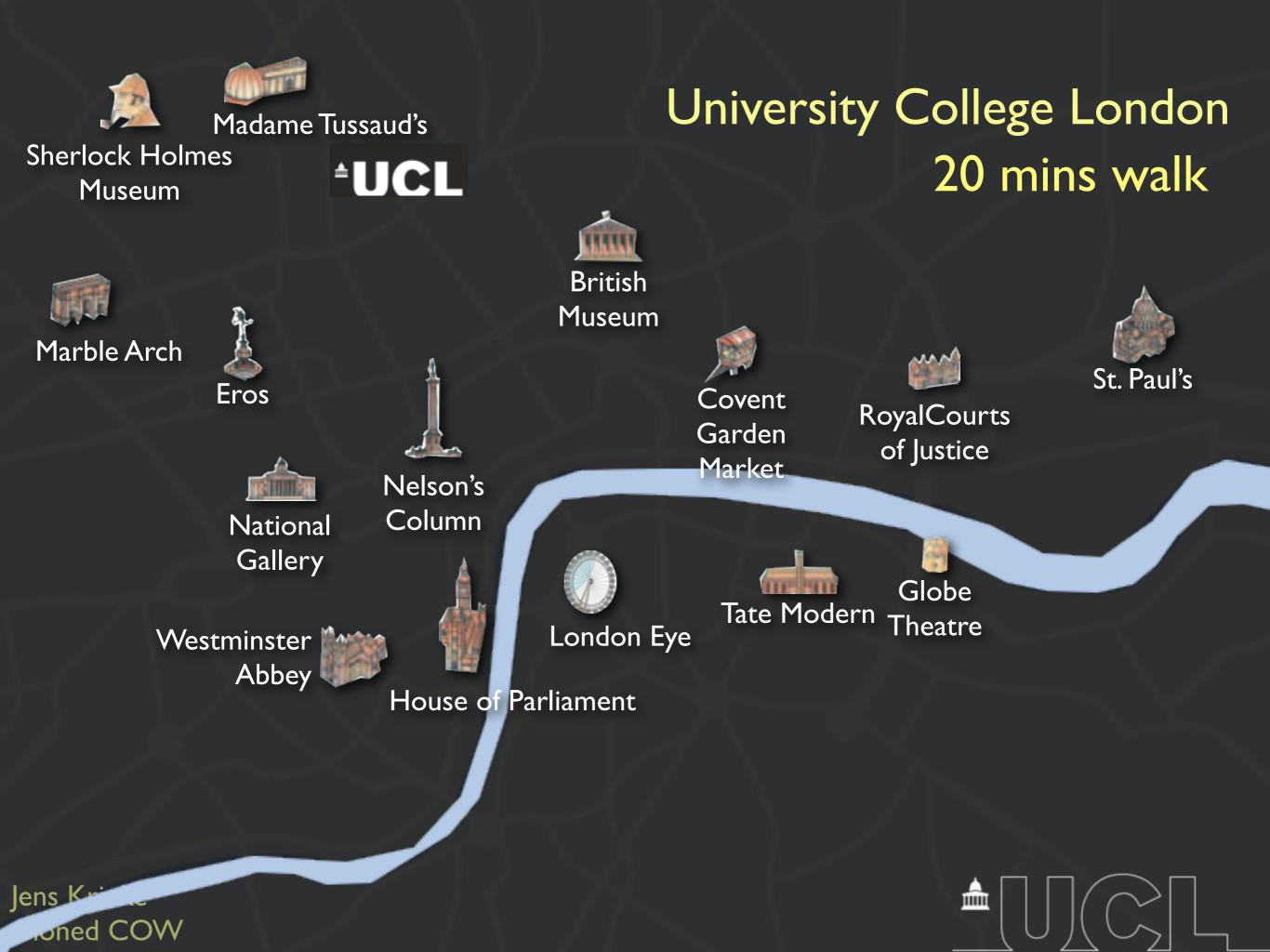
#### <sup>≜</sup>UCL

#### University College London 20 mins walk









## Search Based Software Engineering

Apply search techniques to Software Engineering

#### Multi objective optimisation:

- cost and value
- faults, coverage, time
- metric sets
- risk and reward

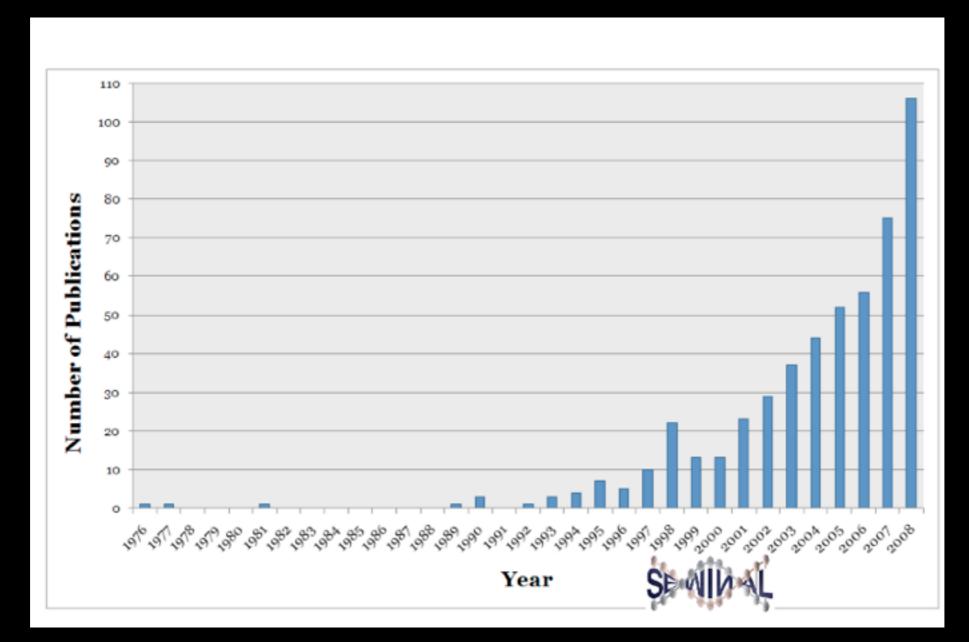
Centre for Research in Evolution, Search &

#### Software Engineering Application areas:

- Requirements
- Testing
- Project Management
- Refactoring
- Source Code Analysis

Ev TeS

### Search Based Software Engineering





## Testing, Testing, Testing



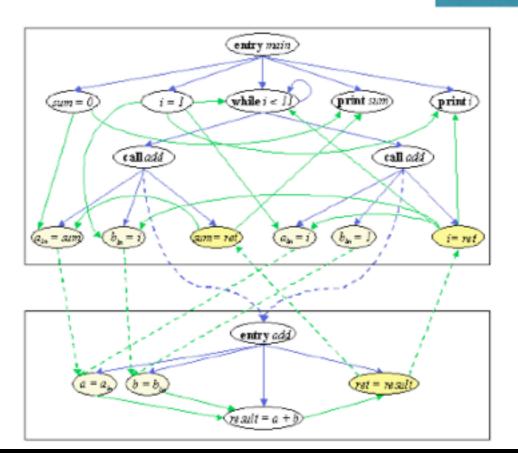
## Testing, Testing, Testing

Mutation testing Model based testing Automated test data generation Web application testing Service oriented testing Fuzz testing Security testing **Regression testing** 



### Dependence Analysis

dependence clusters impact analysis program and model based slicing assessing third party code





#### Quantitative Information Flow

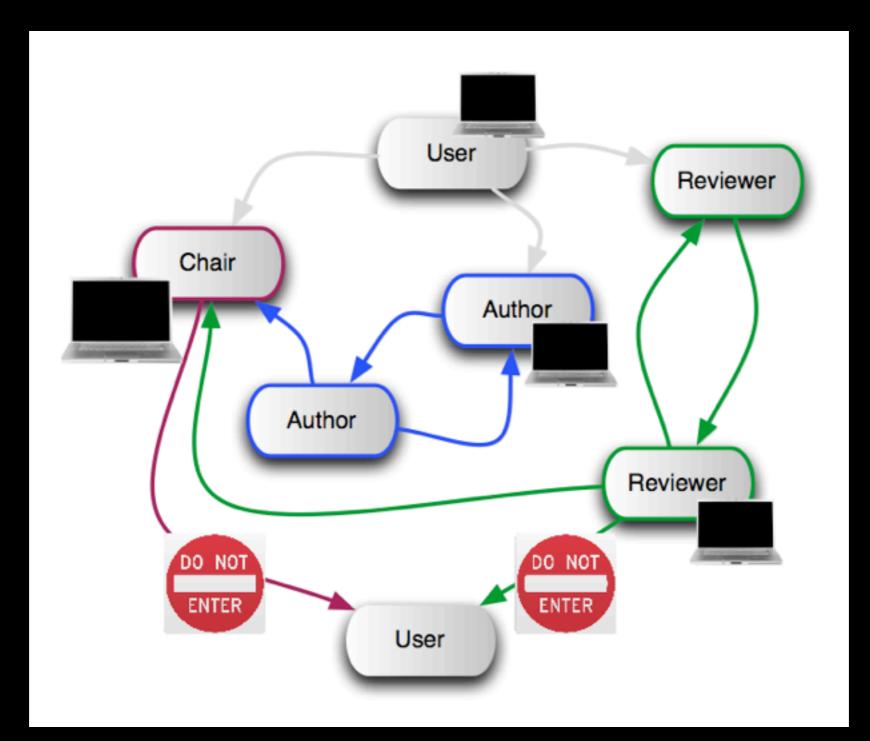
#### calculate amount of information flowing find security leaks Quantitative Information Flow

An EPSRC funded research project based at three sites in London:





#### Information Flow Control





### Clone Detection

using static analysis applied to code provenance applied to graphical languages



North R						
	ox 10.3.5.0					100 (B) - D-
file I	tilt Scope Metrics Preferences	View Help				
2	File Table Clone-Set Table	Statler Plot Source	e Text	Scrapbook		)
-	H:Wkamilya¥prog¥smith2008¥Cem>	1 25	H:Wan	wyeWprogWarnith2008#GemXWgemxW3	8 Hi¥kan	iya¥prog¥smith2005¥GemX¥gema¥
	File ID Path		110	TableItem item = table.( -	91	buffer.append(comr +
	E 25 gemckCloneSalaction		111	for (int $j = 0; j < column$	92	buffer.append(filetw
			112	if (] >- 1) (	93	) (£ -< () N eals (
	25 genut/ClonesetTable;	and the second se	112	int metricindex = 1-	- 94	Int metricIndex = 1
	27 geroofFleSelectionLe		114	String text = fieldFr	95	String text = fieldF-
	🗄 28 gemokikietable java	1 A A A A A A A A A A A A A A A A A A A	115	buffor.appored(text)	96	ba/flor, appond(text)
	E 29 gento/GroupExplores		116	} else {	97	} else {
			117	buffer.expend(item.	-90	buffer.append[term
	30 genoxilSourceTextRul ;;		118	}	99	}
	31 gemokiVain.jeva		115	If (] < columns.Jongth	100	If () < columns.length
	E 32 gemok/vain/window.te		120	<pre>buffer.append(Wt');</pre>	101	buffer.append(Wt');
	33 gemoWVessages.java		121	2	302	2
			122	}	103	>
	34 genodiMetricModelsHi		123	buffer.append(StringUtil.)	104	buffer.append(StringUtil.)
	25 gomemultipleSource		124	)	105	)
	T 35 genod/ScatterPlotPan		125	} else {	106	} else {
	<ul> <li>37 gemol/ScopeHistoryIt</li> </ul>		128	TableColumn]] columns = b	107	tabletolumn]] columns = 6
			127	for (int ) = 0; ) < $2 - 1; +4$	108	for $(int ) = 0; 1 < 4 - 1; +4$
	38 genovsauretextext		128	buffer.append[columns[]]	109	buffer.append(columns())
	39 genod/SourceTextPan		129	buffer.append('%t');	110	buffer.append(W);
	40 genuckSourceTextRule		130	in the second	111	had a second and a second a second
	<ul> <li>[1] 41 genol(TableWithChec</li> </ul>		131	buffer.sopend(columns[2 - 1 buffer.sopend(StringUtil.Ne	112	buffer.append(columns(4 - ) buffer.append(StringUtiLNe
			133	for (int i = 0; i < selected)	114	for fint i = 0; i < selected)
0	1 41 marriel Michaels Carthon		1	In the selected		IN PARTY OF THE SELECTED
teady.	PNI 3	ava etdes: 89 (	(2)	#Clone Netsc 21.7 (d) Rec	t	• 🗶 🕇 💺 🖂 Signore case

n i

е

## Digital Humanities

Mapping outwards from software to systems Software Engineering Analysis translates to

- musicology
- archaeology





#### PERSONAL-TOUCH



## SBSE Repository

Raject Community	ware Engineering By Automated SEarch
This page collects the work which techniques (i. e. Genetic Algorithm	REPOSITORY address the software engineering problems using metaheuristic search optimisation is) into the Repository of Publications on Search Based Software Engineering
SBSE	<ul> <li>SBSE repository is maintained by Yuanyuan Zhang</li> <li>616 relevant publications are included</li> <li>Last updated on the 1 December 2009 (Update logs)</li> </ul>
	The number of publications in the year from 1976 to 2009.
<b>7</b>	The ratio of SE research fields that involoved SBSE.
	The ratio of publications number in the world countries.



## GP Bibliography



#### The Genetic Programming Bibliography

[Search It | Most recent | Top 10 Cited | Most Downloaded | Add to It ]



bibliography It is

The bibliography is part of the <u>Collection of Computer Science Bibliographies</u> It is maintained and managed by <u>William Langdon</u>, <u>Steven Gustafson</u>, and <u>John Koza</u>.

Add single entries or emai1 William Langdon or firstname D0T lastname AT-SYMBOL research D0T ge D0T com a bibtex file containing the entries.

Bibtex file gp-bibliography.bib (Compressed) Over 6000 GP references

Other formats refer format (Compressed) text file (Compressed)

Author format (split by first author)

#### Search Interfaces

The GP bibliography is one of the many online computer science bibliographies There are several on line search tools for these bibliographies. For example: o <u>WWW form</u> based search

o The coauthor graphs can also be used for searching by author name.

o You can find all entries for an author using a URL like this one: <u>http://www.cs.bham.ac.uk/~wbl/biblio/gp-html/WilliamBLangdon.html</u> Naturaly you must replace WilliamBLangdon.html by the author's name.

#### Other Resources

o Bibliographies of evolutionary computation conferences.

o Hints on using bibliographies with PC based word processing packages

o Links to other information on bibliographies.

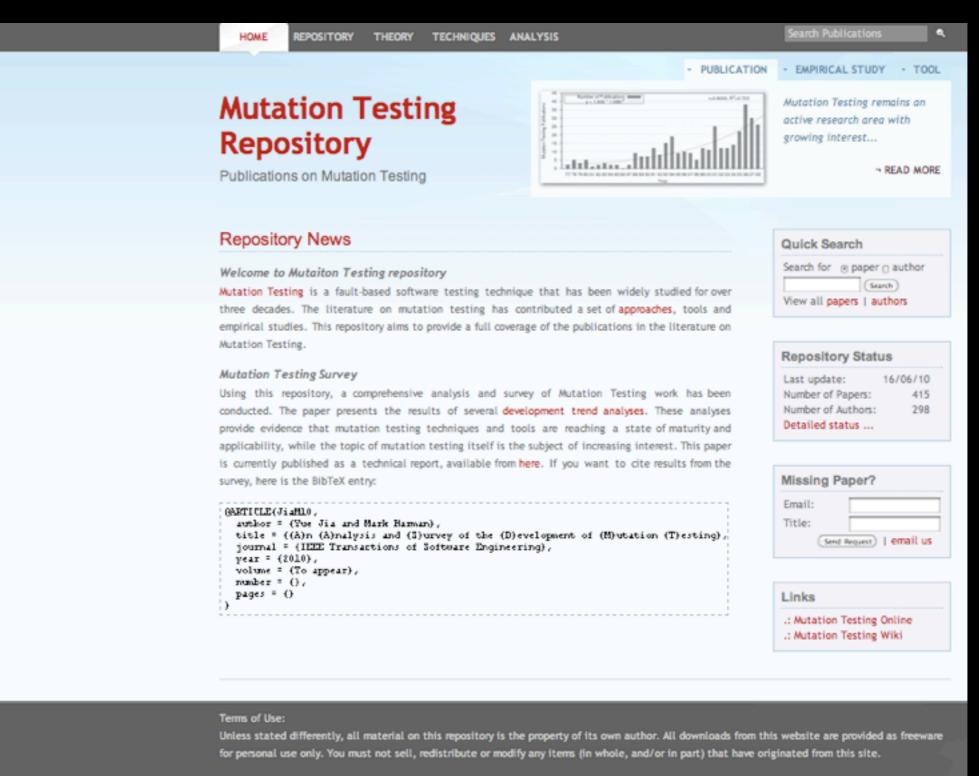
o Subdirectory containing some tools and templates for maintaining bibliographies.

o Links to other bibliographies.

o Home pages of some GP researchers



#### Mutation Testing Repository



Copyright © 2007 - 2010 Yue Jia and Mark Harman all rights reserved.



### COWs

CREST Open Workshop

Typically last week of the month

Aim for last Wednesday (where rooms available)

Discussion based

Recorded and archived



### COWs

CREST Open Workshop Typically last week of the month Aim for last Wednesday (where rooms available) Discussion based Recorded and archived

Recorded and archived



### COWs

CREST Open Workshop Typically last week of the month Aim for last Wednesday (where rooms available) Discussion based Recorded and archived

#### Previous COWs



### Previous COWs

- Nov 09: Search Based Software Engineering
- **Dec 09: Software Testing**
- **Jan 10: Using Static Analysis for Fault Prediction**
- Feb 10: Operational Research for Software Engineering Methods
- Mar 10: Information Theory for Search Based software Engineering
- Apr 10: Dependence Analysis and Slicing for Programs and Models
- May 10: Information Flow and Security
- **Oct 10: Advanced Mootation Testing**



## Happy Birthday!

First COW: 24/25 November 2009



## Upcoming COWs

#### **10th CREST Open Workshop**

#### **Program Analysis and Slicing**

Date: 24 - 25 Jan, 2011

Venue: William Goodenough Room, Goodenough Club

#### 11th CREST Open Workshop

#### SBSE for Early Lifecyle Software Engineering: Requirements and Management

Date: Wed 23 Feb, 2011

Venue: Engineering Front Executive Suite, Roberts Building



Jian Ren, COW Programme manager Lena Hierl, Administrative arrangements Yue Jia, visitor liaison EPSRC, HP, British Council, funding Attendees and speakers



Jian Ren, COW Programme manager Lena Hierl, Administrative arrangements Yue Jia, visitor liaison EPSRC, HP, British Council, funding Attendees and speakers



Jian Ren, COW Programme manager Lena Hierl, Administrative arrangements Yue Jia, visitor liaison EPSRC, HP, British Council, funding Attendees and speakers



Jian Ren, COW Programme manager Lena Hierl, Administrative arrangements Yue Jia, visitor liaison EPSRC, HP, British Council, funding Attendees and speakers



Jian Ren, COW Programme manager Lena Hierl, Administrative arrangements Yue Jia, visitor liaison EPSRC, HP, British Council, funding Attendees and speakers



Jian Ren, COW Programme manager Lena Hierl, Administrative arrangements Yue Jia, visitor liaison EPSRC, HP, British Council, funding Attendees and speakers



#### Domestics

Fire or other alarm

Toilets

Discussions



#### Domestics

Fire or other alarm

Toilets

Discussions



## Mary Augusta Ward

- British novelist (Mrs Humphry Ward)
- II June 1851 24 March 1920
- one of the founders of the Women's National Anti-Suffrage League
- helped establish an organization
   for working and teaching among the poor
- The house acted as center of training and community meeting and socialising point

#### Introductions ...

