UK Educational Evidence Portal Requirements for a new search engine using text mining

— 18th June 2008 —

James Thomas, Ruth Stewart, Sandy Oliver

EPPI-Centre, Social Science Research Unit,
Institute of Education, University of London

EPPI CENTRE
NaCTeM The National Centre for Text Mining
JISC
SSRU Social Science Research Unit
IOE LONDON Leading education and social research Institute of Education University of London

Overview

- What is eep?
- About our consultation
- A new search engine
  - Based on ASSERT
  - Developed by NaCTeM
- Next steps

UK Educational Evidence Portal

- eep – is an online resource dedicated to and for educational professionals, bringing together research and evidence from multiple sources
- An ‘all resources’ search enables users to search the contents of the websites of 28 organisations
- A ‘selected resources’ search enables users to search a much smaller pool of materials that have been assigned terms according to the British Education Thesaurus.

UK Educational Evidence Portal - Windows Internet Explorer
http://localhost:50302/NECE.SearchService.UI.Portal/Main/Default.aspx
Live Search
UK Educational Evidence Portal
Page Tools

educational evidence portal

Search eep

- Search all resources
- Search selected resources
- Desktop search tools

eep Information

- Key resources
- Background
- Quality assurance
- Making use of evidence
- Selection criteria

Welcome to eep

the UK Educational Evidence Portal

This portal helps you find educational evidence from a range of reputable sources using a single search. It is designed for both professional and lay people interested in education.

Search All Resources

Refine: Reset

Help | Save Preferences | Keywords Builder | Sources | Sort Order

Enter the keywords or phrase that interests you. The search will retrieve documents from the websites of all the organisations associated with eep that contain any of the words you have entered. For guidance on how to get the best from your search click **Help**

You can also **Search selected resources** from a database of resources selected by a smaller group of organisations to be particularly relevant to practitioners and researchers. These resources have been indexed to provide more focused searching.

This portal is being developed continuously and new content added from time to time. The database of selected resources is currently limited in size but will become increasingly valuable as more resources are added.

Give us your views!
An online survey of users' opinions about eep is being carried out right now, by the Higher Eduction Academy, which will influence the future development of eep. Take a minute to feed back your thoughts.

Next steps for eep
The eep consortium is developing the portal to suit the needs of specific groups. Through workshops with subject associations, children's' services and policymakers it is finding out what is needed. Contact eep if you would like to suggest another group at **eep@cfbt.com**.

ESRC now part of eep
The Economic and Social Research Council has made the education documents from its Society Today database available

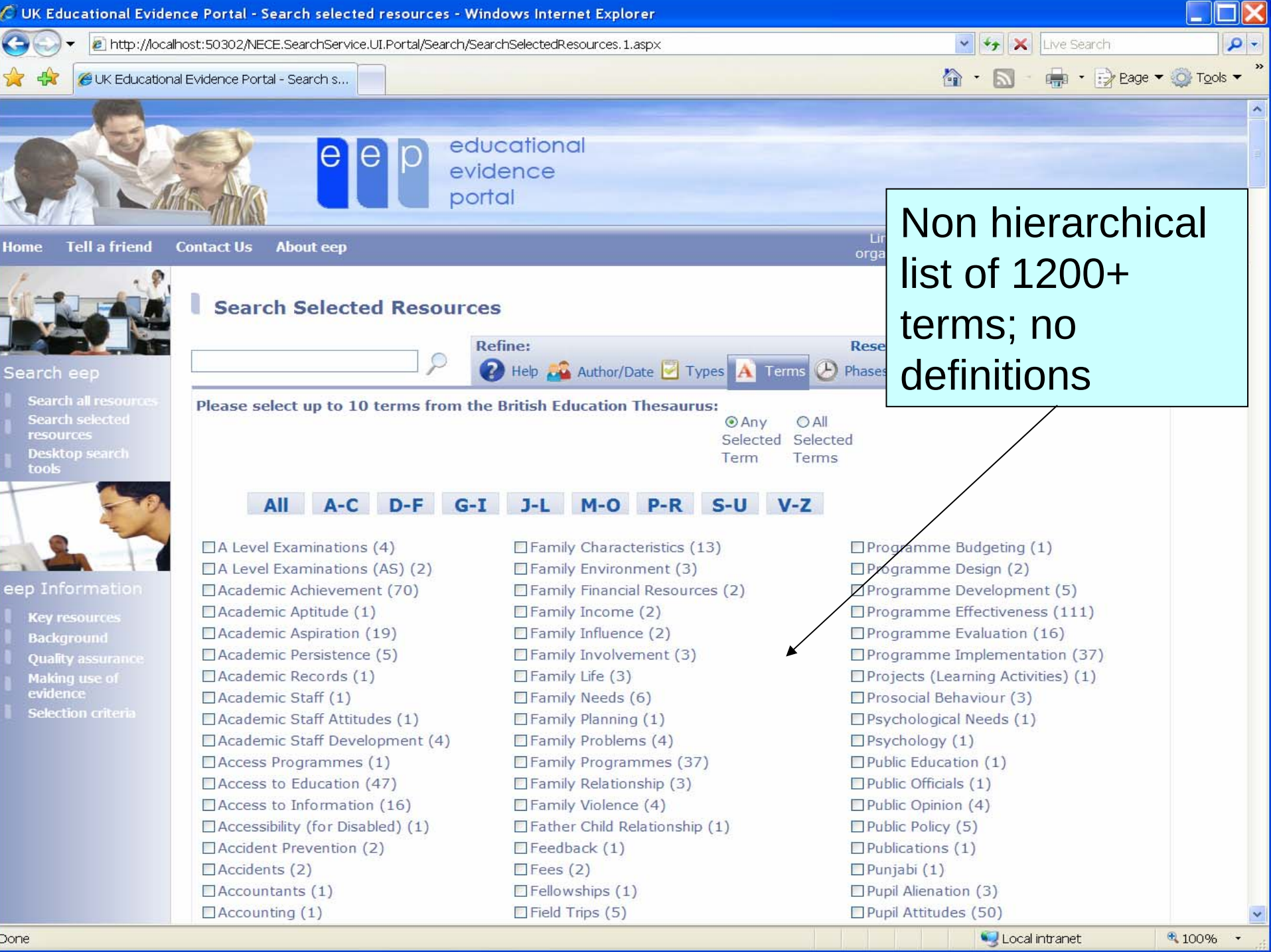
Done
Local intranet
100%

The ‘all resources’ search

- Uses the Microsoft ‘Windows Live’ web service
- Searches the websites of 28 organisations
- Results ranked according to Windows Live algorithm (maximum 250 hits) – certain sites always end up on top
- Sometimes large numbers of irrelevant hits

The ‘selected resources’ search

- A much smaller pool of documents: approximately 1300
- Documents have been identified as relevant by eep consortium and assigned British Education Thesaurus Terms by the BEI
- Records are in the form of citations with URLs
- Stored and searched in an SQL Server database


UK Educational Evidence Portal - Search selected resources - Windows Internet Explorer
http://localhost:50302/NECE.SearchService.UI.Portal/Search/SearchSelectedResources.1.aspx
UK Educational Evidence Portal - Search s...
educational evidence portal
Home
Tell a friend
Contact Us
About eep
Search Selected Resources
Refine:
Help
Author/Date
Types
Terms
Phases
Please select up to 10 terms from the British Education Thesaurus:
Any Selected Term
All Selected Terms
All
A-C
D-F
G-I
J-L
M-O
P-R
S-U
V-Z
A Level Examinations (4)
A Level Examinations (AS) (2)
Academic Achievement (70)
Academic Aptitude (1)
Academic Aspiration (19)
Academic Persistence (5)
Academic Records (1)
Academic Staff (1)
Academic Staff Attitudes (1)
Academic Staff Development (4)
Access Programmes (1)
Access to Education (47)
Access to Information (16)
Accessibility (for Disabled) (1)
Accident Prevention (2)
Accidents (2)
Accountants (1)
Accounting (1)
Family Characteristics (13)
Family Environment (3)
Family Financial Resources (2)
Family Income (2)
Family Influence (2)
Family Involvement (3)
Family Life (3)
Family Needs (6)
Family Planning (1)
Family Problems (4)
Family Programmes (37)
Family Relationship (3)
Family Violence (4)
Father Child Relationship (1)
Feedback (1)
Fees (2)
Fellowships (1)
Field Trips (5)
Programme Budgeting (1)
Programme Design (2)
Programme Development (5)
Programme Effectiveness (111)
Programme Evaluation (16)
Programme Implementation (37)
Projects (Learning Activities) (1)
Prosocial Behaviour (3)
Psychological Needs (1)
Psychology (1)
Public Education (1)
Public Officials (1)
Public Opinion (4)
Public Policy (5)
Publications (1)
Punjabi (1)
Pupil Alienation (3)
Pupil Attitudes (50)
Search eep
Search all resources
Search selected resources
Desktop search tools
eep Information
Key resources
Background
Quality assurance
Making use of evidence
Selection criteria
Non hierarchical list of 1200+ terms; no definitions
Done
Local intranet
100%

Our consultation

- Aimed to speak to users and potential users of eep in order to:
  - Understand what they search for, how and where
  - Favourite features, ‘wish lists’ and ‘show stoppers’
- Interviews and focus groups
- Recorded and made detailed notes
- Identified key priorities for new search engine

Who we spoke to

- 47 people working in education
  - 5 focus groups and
  - 13 interviews

- The focus groups were with:
  - National Foundation for Educational Research (NFER)
  - Training and Development Agency (TDA)
  - CfBT Education Trust
  - 2 groups from the Department for Children Schools and Families (DCSF) (analysts and researchers).

Who we spoke to - interviews

- Four teachers / people preparing teaching resources
  - a primary school deputy head
  - a primary school teacher
  - a secondary school head of subject
  - a field studies centre trainer
  - an education information officer within an NGO
- Information specialists from the Higher Education Academy and Learning and Skills Council, the DCSF, Institute of Education, London (IoE)
- Leaders of
  - Best practice and research for a county council
  - Research and development at a higher education college
  - 2 subject associations

What people told us – what they DON'T want

- **Show stoppers**
  - Site registration required
  - Payment required
- **(Dis)functional**
  - Tick boxes or specific pre-selected terms of limited value
  - Small search boxes
  - Confusion over 'select all button'
  - Similar/ duplicate results
  - Many citations of a report, but not the report itself
  - Broken links
  - Jargon
- **Cosmetic**
  - Adverts
  - Popups

What people told us – what they DO want

- Expected / standard
  - If not available considered ‘show stoppers’
- Desirable criteria
  - Favourite features of sites they use
- Competitive edge
  - Wish list

What they DO want – **Expected / standard**

- Fast searches
- Clear information/ guidance
- Clear instructions for searching
- Definitions available
- Clarity over which sites are being searched

What they DO want – **Desirable criteria**

- Can do a useful task in 10 minutes
- Can save and return
- Simplicity of site design (clean page/ three clicks)
- Intuitive design
- “Trusted” sources
- Display search history
- Display 5-20 hits initially??
- Display: details of source, keywords, date created, type of source, pictures
- Clear citations / attributions
- Able to navigate to keywords in found doc
- Markers of ‘bias’: credible organisations, well designed sites, quality assured sites, star rating
- Presenting results in terms of their % match to the search terms used
- Present results in different formats
- Useful pointers of where to look next

What they DO want - **Competitive edge**

- Search for complex and new concepts, changing policies/ curricula/ terminology
- Search for document types (policy, research, resources, news, pictures, blogs, curricula) not just content
- Search for UK/ date in search (not just search output)
- Search for higher education/ schools / certain organisations
- Search for author, title
- Filter by concepts
- Natural language searches
- A “search engine that interrogates you”
- Access to full text documents for free

Text mining search for eep

eep consortium websites

webcrawler

eep database

ASSERT

Next steps...

1. Develop three options for searching eep using text mining
2. Evaluate these alternatives
3. Demonstrate and deploy

Next steps (1): search engines

- Free-text search - lists of documents
  - Search returns documents which contain the selected term(s) AND related documents
- Free-text search - clusters of documents
  - Rather than a simple list, this search returns a list of clusters which can then be explored
- Taxonomy search - either clusters or list of documents
  - Returns results in either format based on user selection of term(s) from a hierarchical taxonomy

JISC NaCTeM The National Centre for Text Mining ePPI CENTRE

Next steps (2): evaluations

- Internal evaluation
  - Sensitivity, specificity and efficiency
  - Research team and eep development group
- External evaluation
  - User testing
  - Interviews and observations

Next steps (3): demonstrate / deploy

- Demonstration site on the eep portal
- If evaluations are positive -
  - The preferred search(es) will be used on the live portal

educational evidence portal

Search eep
- Search all resources
- Search selected resources
- Desktop search tools

eep Information
- Key resources
- Background
- Quality assurance
- Making use of evidence
- Selection criteria

Seed text mining search

depression

- All Documents (159)
- IPS Model Of Supported Employment (22)
- For People With Psychiatric Disabilities (14)
- Patients Are (15)
- Cognitive Therapy (13)
- Clients In The Program (11)
- Mental Health Care (11)
- For Persons With Mental Illness (10)
- Of Job (7)
- The Community (10)
- Mentally Ill (8)
- As An Evidence-based Practice (8)
- Scores And (6)
- Of Sheltered Workshop Workers Was (5)
- People With Dual Diagnoses (4)
- And Social (4)
- Day Treatment Program (4)
- Levels Of Service (4)
- Long-term Employment (4)
- Goals To (3)

- VOCATIONAL-REHABILITATION IN SCHIZOPHRENIA - 1995
- Vocational rehabilitation in psychiatry: A re-evaluation - 2004
- The connection between mental illness, rehabilitation and work (ZERA) - A training programme for medical-vocational rehabilitation of persons with mental illness ZUSAMMENHANG ZWISCHEN ERKRANKUNG, REHABILITATION UND ARBEIT (ZERA) - EIN SCHULUNGSPROGRAMM FUR DIE MEDIZINISCH-BERUFLICHE REHABILITATION PSYCHISCH KRANKER MENSCHEN - 2001
- Long-term psychiatric patients in vocational rehabilitation programmes: A naturalistic follow-up study over 3 years - 2000
- - 1974
- Recent trends in vocational rehabilitation for people with psychiatric disability - 1994
- THE SELF-HELP-EMPLOYMENT-CENTER - SOME OUTCOMES FROM THE FIRST YEAR - 1995
- Adding a vocational focus to mental health rehabilitation. - 1996
- Vocational decision making in a psychiatric outpatient program - 1998
- Correlational designs in rehabilitation - 2004
- An individualized job engagement approach for persons with severe mental illness - 1999
- Vocational rehabilitation for persons with schizophrenia: Recent research and implications for practice - 2000
- Assistive technology: Choosing the right tool for the right job - 2006
- Development of the DPA fidelity scale: Using fidelity to define an existing vocational model - 2004
- Patterns in Referral and Admission to Vocational Rehabilitation Associated with Coexisting Psychiatric and Substance-Use Disorders - 2003
- Strategies to Promote Successful Employment of People with Psychiatric Disabilities - 2005
- Organising for work: a job clinic for people with mental health needs - 1997
- -
- The Clubhouse model and employment following brain injury - 1996
- The consumer-run business: People with psychiatric disabilities as entrepreneurs - 1998
- Person-centered planning for the millennium: We're old enough to remember when PCP was just a drug - 1997

UK Educational Evidence Portal - Search all resources - Windows Internet Explorer

http://localhost:50302/NECE.SearchService.UI.Portal/Search/SearchAllResources.1.aspx

UK Educational Evidence Portal - Search al...

eep educational evidence portal

Home | Tell a friend | Contact Us | About eep

Links to eep organisations: Select...

Search eep

- Search all resources
- Search selected resources
- Desktop search tools

eep Information

- Key resources
- Background
- Quality assurance
- Making use of evidence
- Selection criteria

Seed text mining search

depression

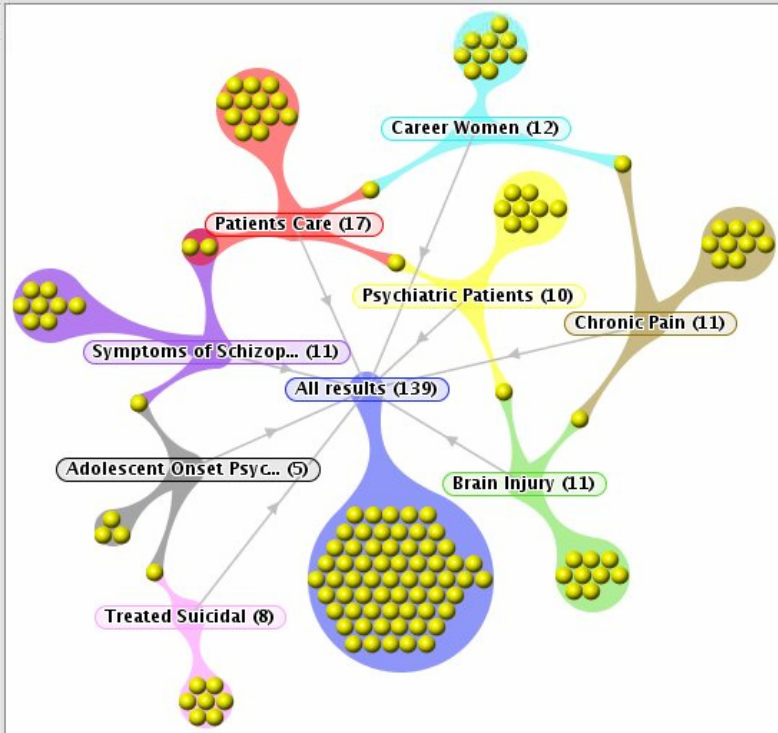
Visualisation

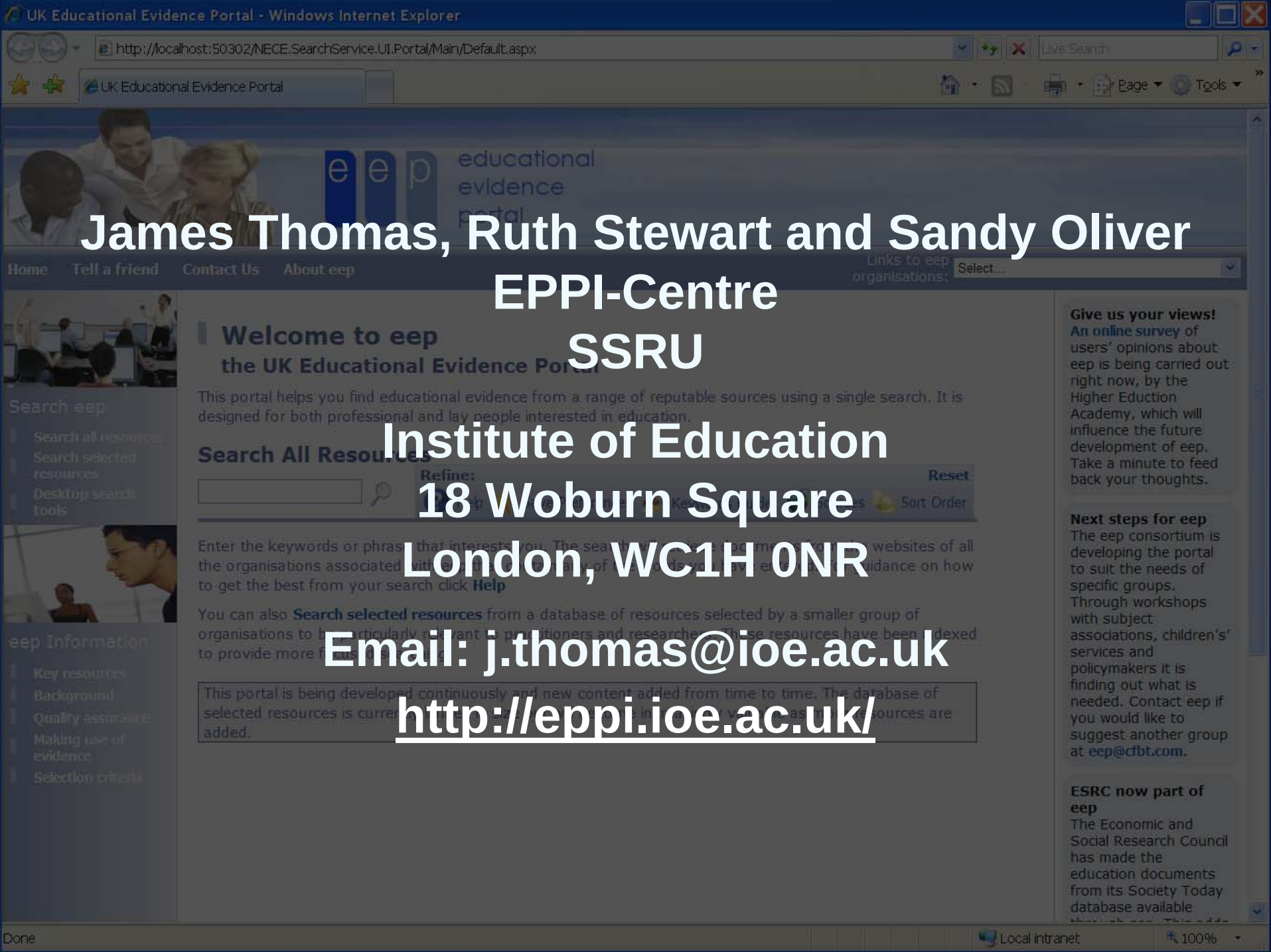
Cluster map for query: depression

Scale: scale to fit

Categories

All results	139
Patients Care	17
Job Club Interventi...	19
Substance Abuse ...	14
Hidden Costs of M...	13
Career Women	12
Brain Injury	11
Chronic Pain	11
Developmental Dis...	10
Symptoms of Sch...	11
Quality of Life	9
Treated Suicidal	8
Psychiatric Patien...	10
Psychological Well-...	7
Group Willing Recei...	6
PORT Clients Survey	7
Domestic Violence	5
Perceived Stress	7
Sexual Harassment ...	5
Workers Male	6
Adolescent Onset ...	5
Differences betwee...	6





James Thomas, Ruth Stewart and Sandy Oliver

EPPI-Centre

SSRU

Institute of Education

18 Woburn Square

London, WC1H 0NR

Email: j.thomas@ioe.ac.uk

http://eppi.ioe.ac.uk/