

FutureLearn partner meeting, 26<sup>th</sup> Sept. 2013

# FLx: an experimental network using external analytics tools to explore FutureLearn data

Simon Buckingham Shum

<http://simon.buckinghamshum.net>

[s.buckingham.shum@gmail.com](mailto:s.buckingham.shum@gmail.com)

@sbskmi

# Quality analytics should...

improve the learning design

improve the user experience

improve the platform

advance learning sciences

# MOOCs are a major new platform for educational research

 Massachusetts Institute of Technology

• MITx



 H

## Watch the edX Press Conference

If stream does not start automatically, press play button below

Press conference: MIT, Harvard announce edX

Technology

UNIVE

Share More info



Rafael Reif

Anant Agarwal


“this is big data, giving us the chance to ask big questions about learning”

# MOOCs are a major new platform for educational research

**TALKS**

## Daphne Koller: What we're learning from online education

FILMED JUN 2012 • POSTED AUG 2012 • TEDGlobal 2012



**1,188,171 Views**

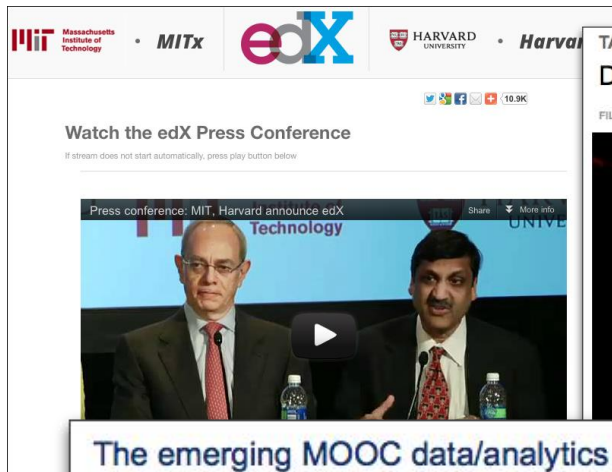
Daphne Koller is ent most intriguing course service, but as a way Coursera (cofounder quiz, peer-to-peer di assignment builds at how knowledge is pr

With Coursera, Daph Ng are bringing course for anyone who want

**RELATED PLAYLISTS**

24/7 A/B testing at scale becomes possible for educators, revealing effects that remain insignificant with conventional sample sizes

# MOOC analytics will be big



TALKS

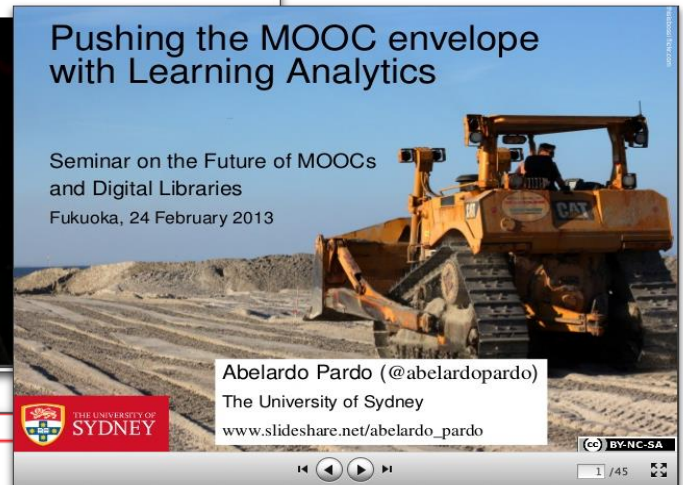
Daphne Koller: What we're learning from online education

FILMED JUN 2012 • POSTED AUG 2012 • TEDGlobal 2012



The emerging MOOC data/analytics ecosystem

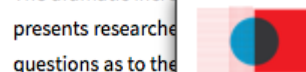
By sbs on January 31st, 2013



MOOC RESEARCH

Welcome to the **MOOCResearch Hub**

The dramatic increase in the number of MOOCs has been leading to a growing body of peer-reviewed research on the impact of MOOCs on education requires



MOOC RESEARCH

**Cost, Performance Metrics and Learner Analytics**

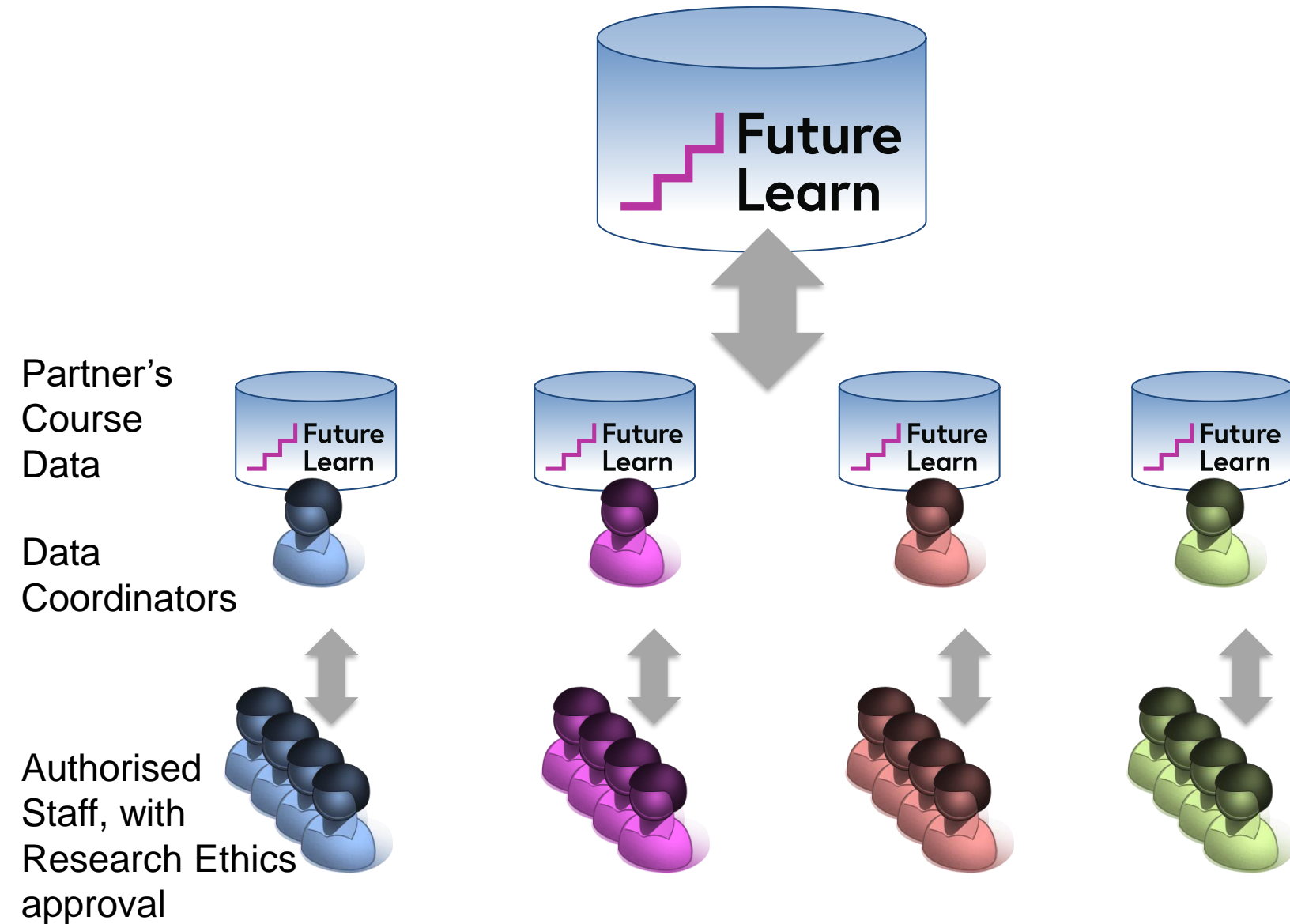
[http://www.ted.com/talks/daphne\\_koller\\_what\\_we\\_re\\_learning\\_from\\_online\\_education.html](http://www.ted.com/talks/daphne_koller_what_we_re_learning_from_online_education.html)

<http://people.kmi.open.ac.uk/sbs/2013/01/emerging-mooc-data-analytics-ecosystem>

[http://www.slideshare.net/abelardo\\_pardo/pushing-the-mooc-envelope-with-learning-analytics](http://www.slideshare.net/abelardo_pardo/pushing-the-mooc-envelope-with-learning-analytics)

<http://www.moocresearch.com/research-initiative/about#Cost,%20Performance%20Metrics%20and%20Learner%20Analytics>

# You will soon have rich datasets from your courses: are you ready?

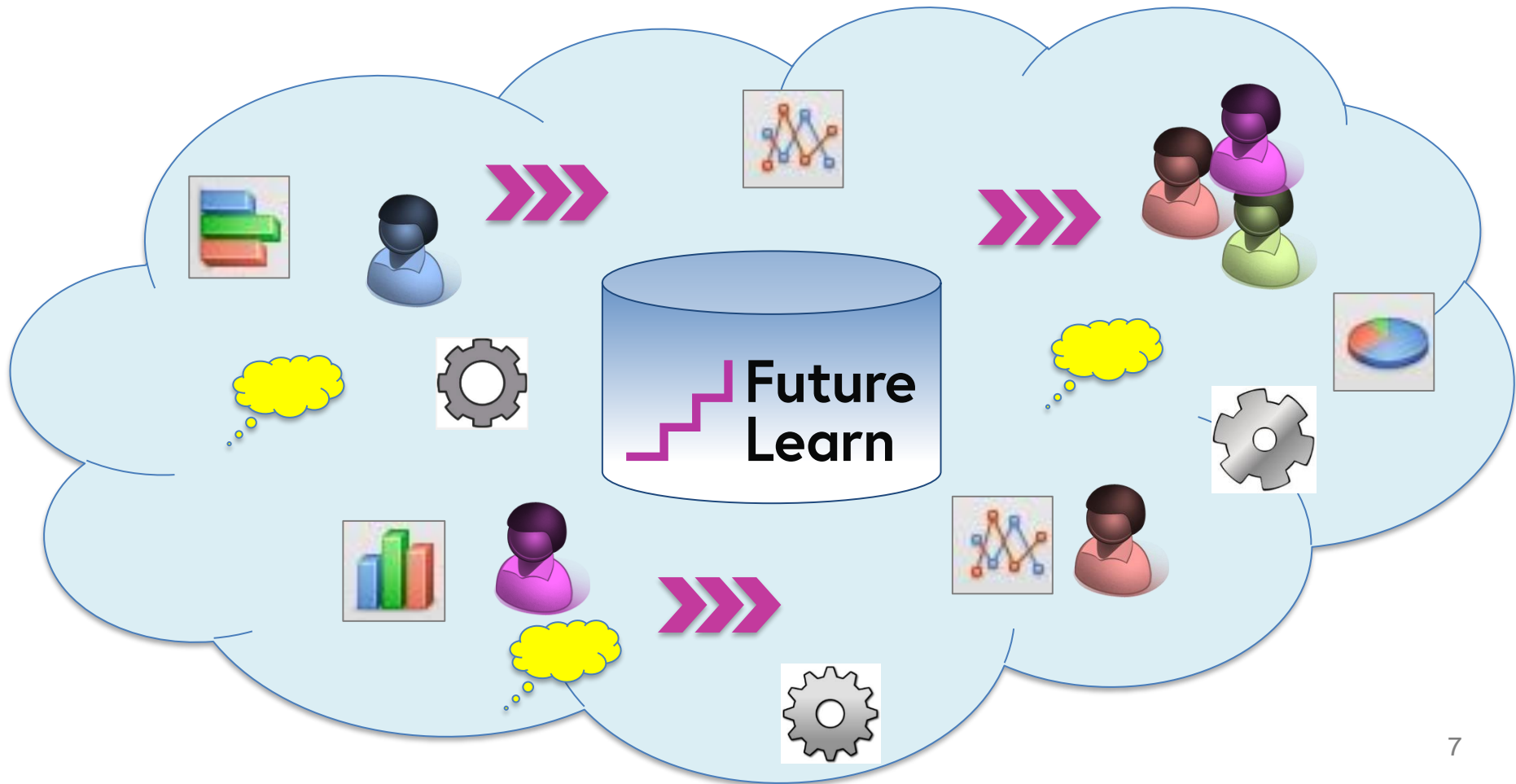




# FLx: experimental analytics ecosystem



Partners share ideas, workflows, analytics, and visualizations, collaborating around common interests.



an experimental network  
using external analytics tools  
to explore FutureLearn data

what might  
this look like?

*some examples of possible analytics workflows on FutureLearn data*



# SNA research reveals hidden social structure: SNAPP plugin -> Forum Subgraphs

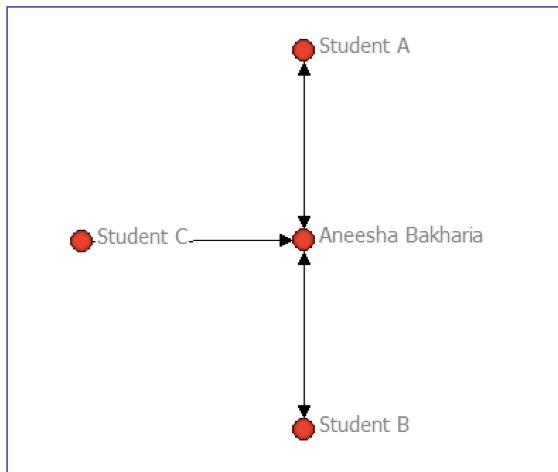
## FORUM A

Thread: Thread Start Reply

Total posts: 14 Unread posts: 0

- Thread Start Aneesha Bakharia
- SNA Tools Student A
  - RE:SNA Tools Aneesha Bakharia
    - RE:RE:SNA Tools Student B
    - RE:RE:SNA Tools Student C
    - RE:RE:RE:SNA Tools Aneesha Bakharia
- Centrality analysis Student A
  - RE:Centrality analysis Aneesha Bakharia

- No student interaction
- All interaction via Tutor/Lecturer



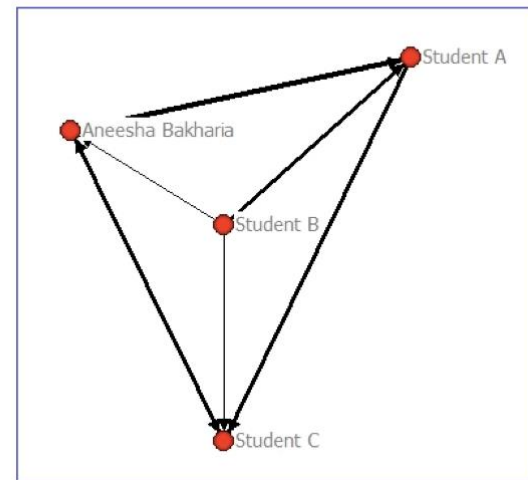
## FORUM B

Thread: RE:Cool SNA Visualisation Reply

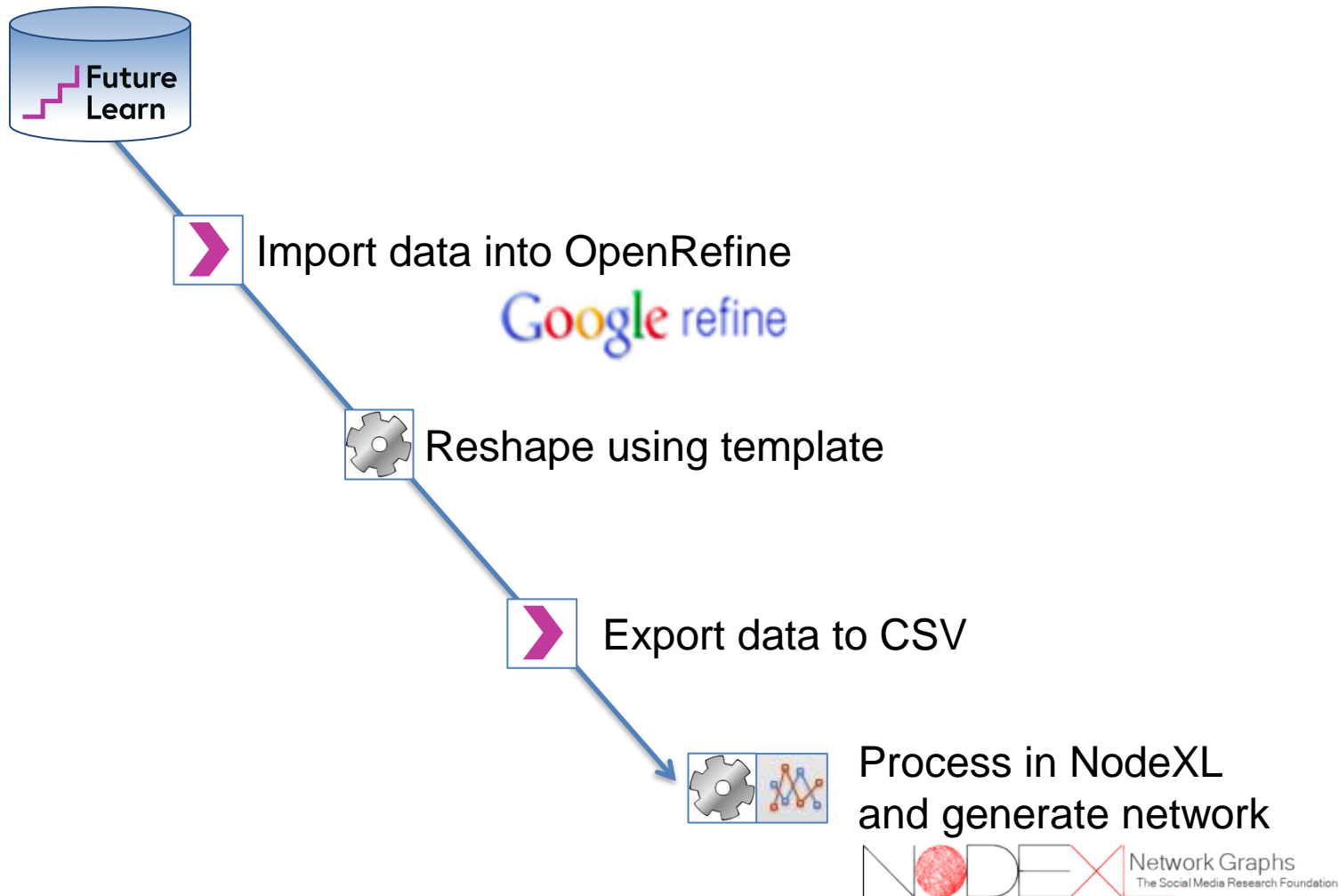
Total posts: 14 Unread posts: 0

- Thread Start Aneesha Bakharia
- Literture Review Student C
  - RE:Literture Review Student A
    - RE:RE:Literture Review Aneesha Bakharia
    - RE:RE:Literture Review Aneesha Bakharia
    - RE:RE:Literture Review Student B
- Sample Data for Analysis and Visualisation Student C
  - RE:Sample Data for Analysis and Visualisation Student A

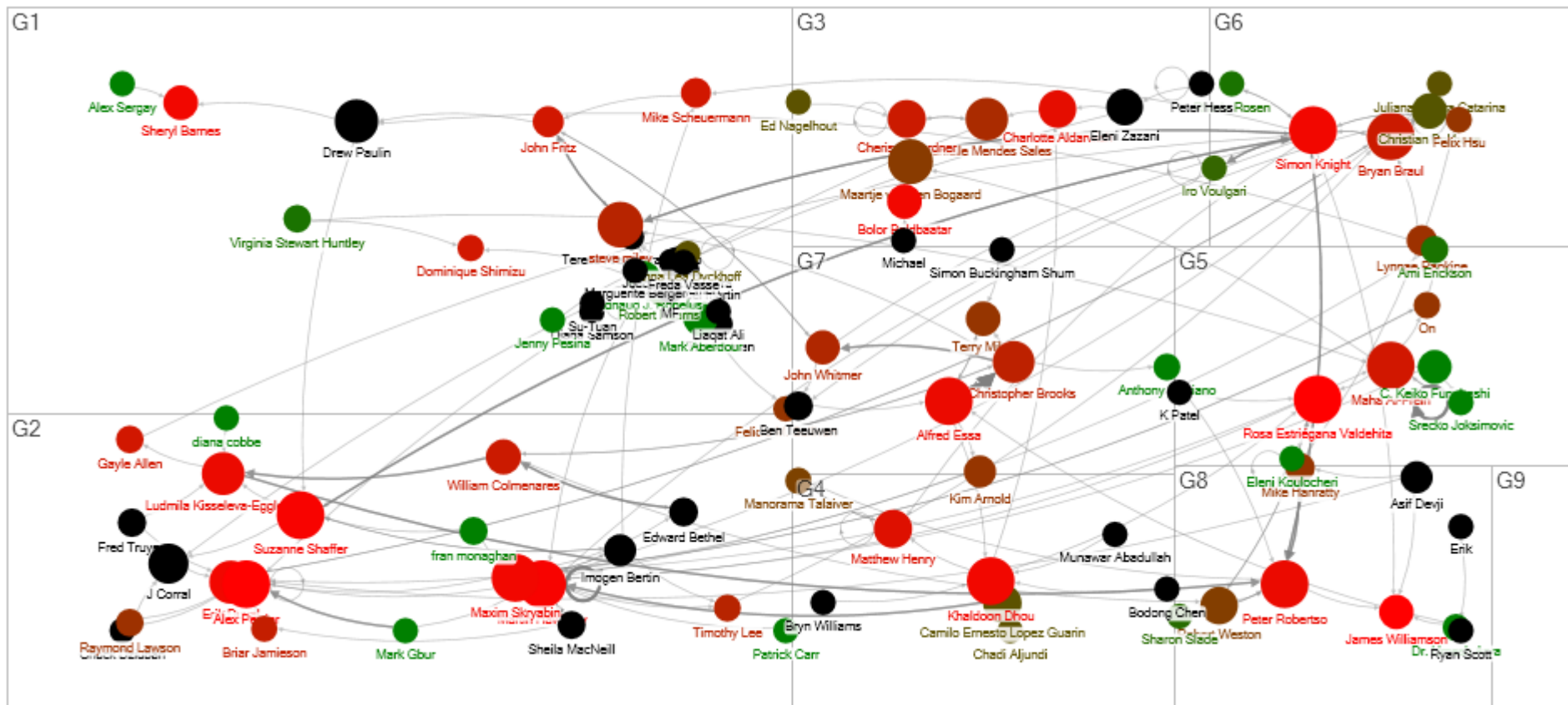
- Student to student social interaction beginning



# Workflow for social network analytics in NodeXL: *are learners forming effective peer-relationships?*

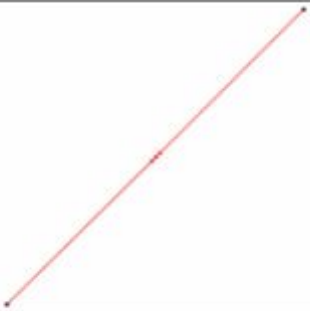
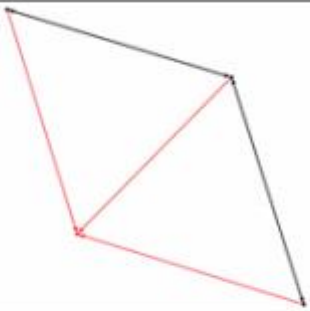
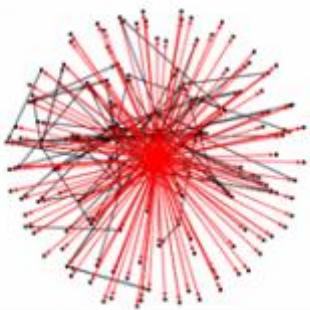

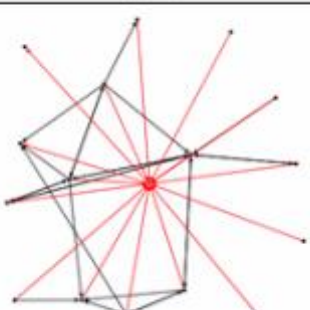
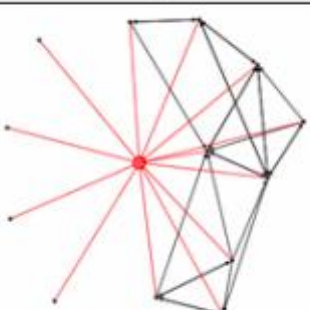


# Workflow for social network analytics in NodeXL: *are learners forming effective peer-relationships?*



# Example 2: NodeXL Reply Network Subgraphs

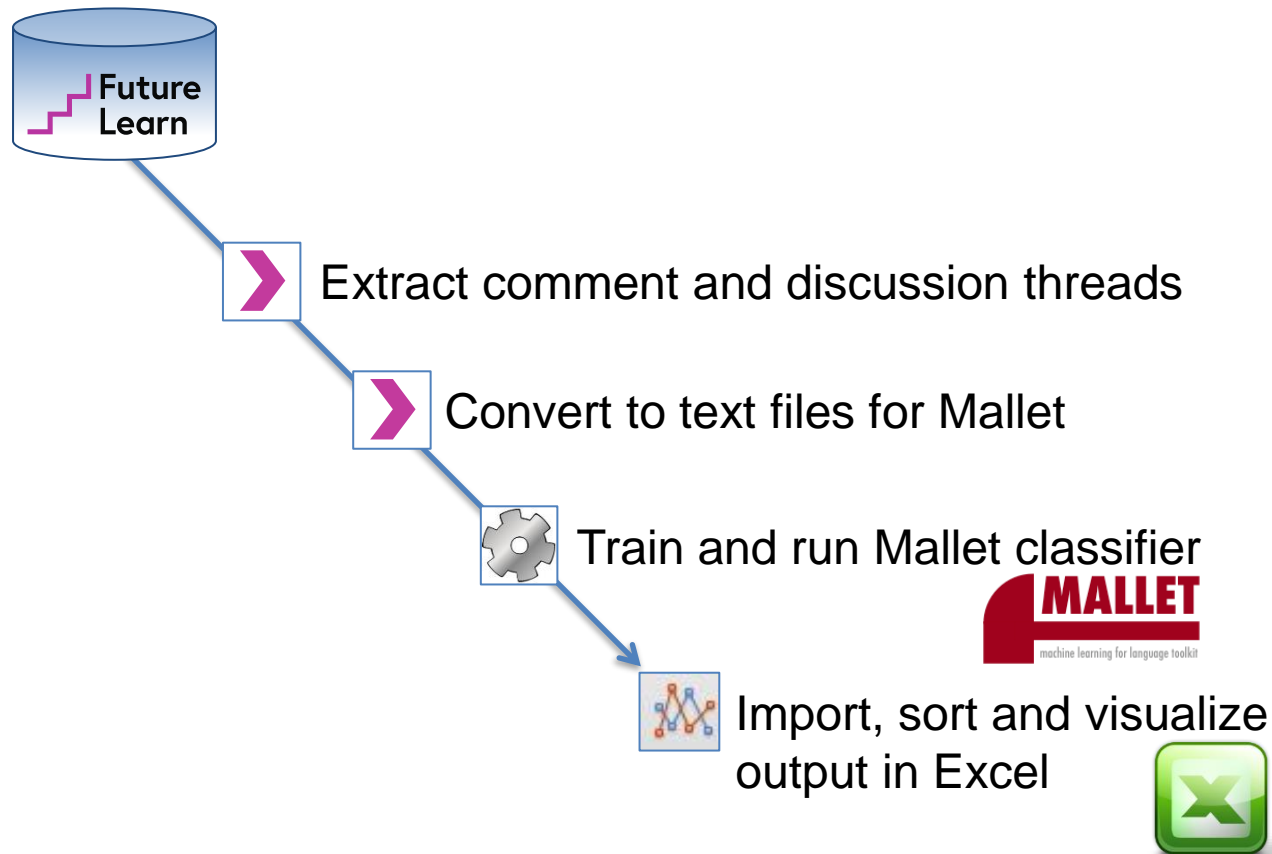
## Visualizing Threaded Conversation Networks: Mining Message Boards and Email Lists

<b>Question People</b> <ul style="list-style-type: none"><li>• Low In- and Out-Degree</li><li>• High Avg Degree of Neighbors</li></ul>		
<b>Answer People</b> <ul style="list-style-type: none"><li>• High % Out-Degree</li><li>• Low Clustering coefficient</li><li>• Low Avg Degree of Neighbors</li></ul>		
<b>Discussion Starters</b> <ul style="list-style-type: none"><li>• Low % Out-Degree</li><li>• High Clustering coefficient</li><li>• High Avg Degree of Neighbors</li></ul>		

SNA research seeks to quantify patterns for, e.g.

- isolated students
- dysfunctional groups
- users who are information brokers

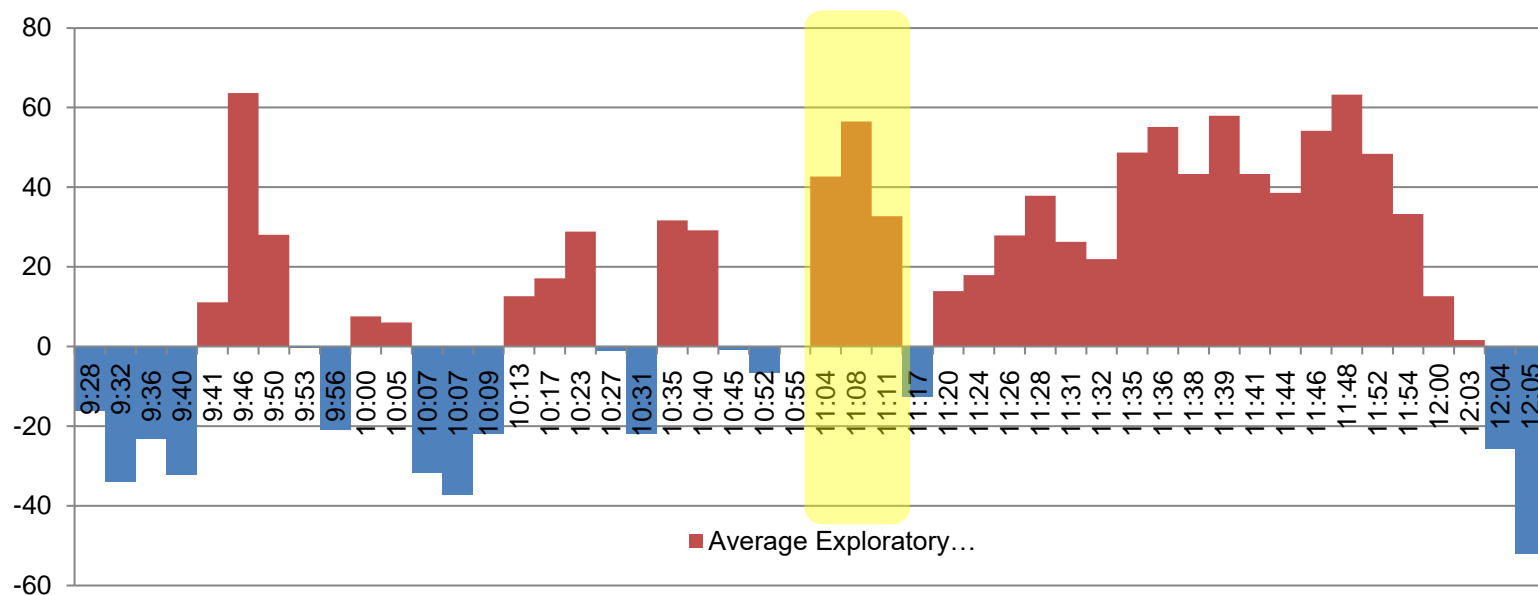
# Workflow for discourse analytics: *which sections of the course seem to provoke knowledge-building discourse?*



# Workflow for discourse analytics: *which sections of the course seem to provoke knowledge-building discourse?*

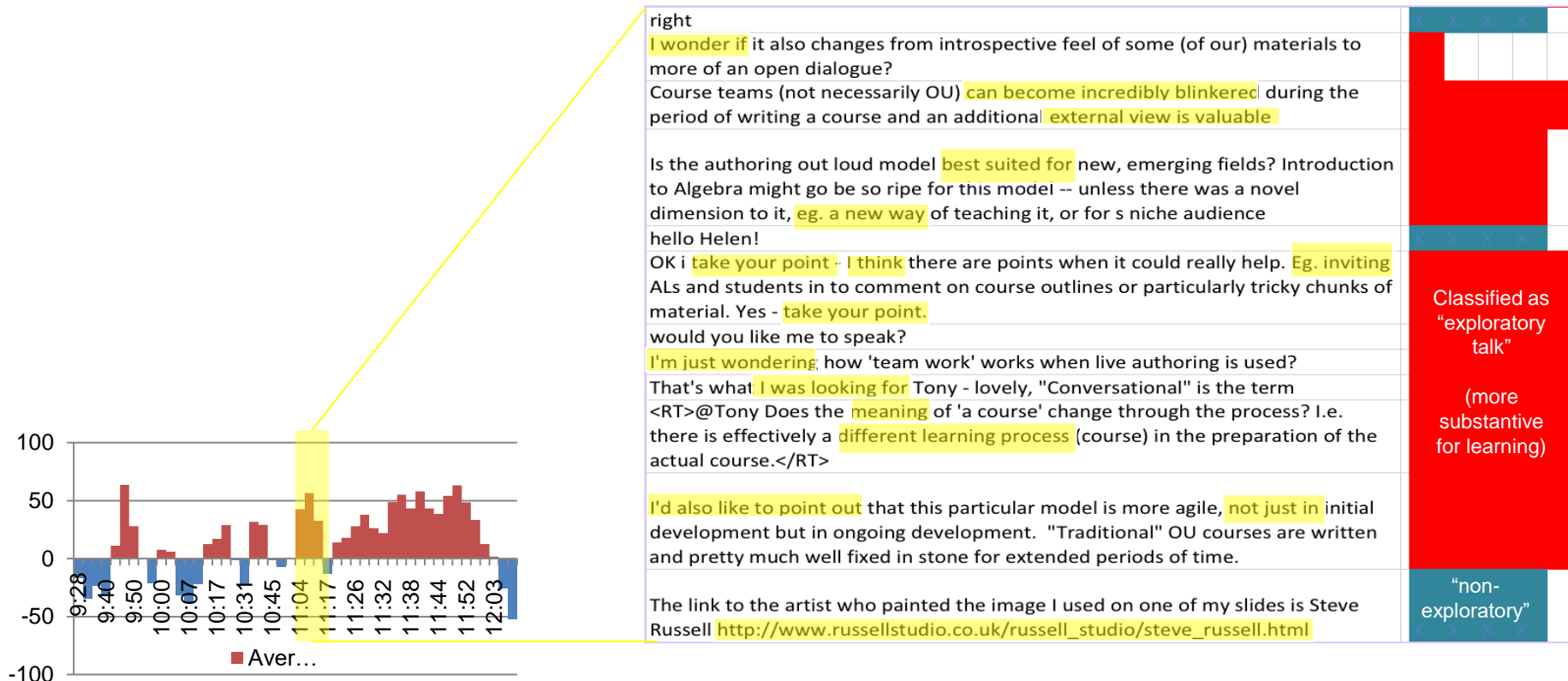
*Given a 2.5 hour textchat where were the most effective learning conversations?*

*Not at the start and end, but if we zoom in on a peak...*

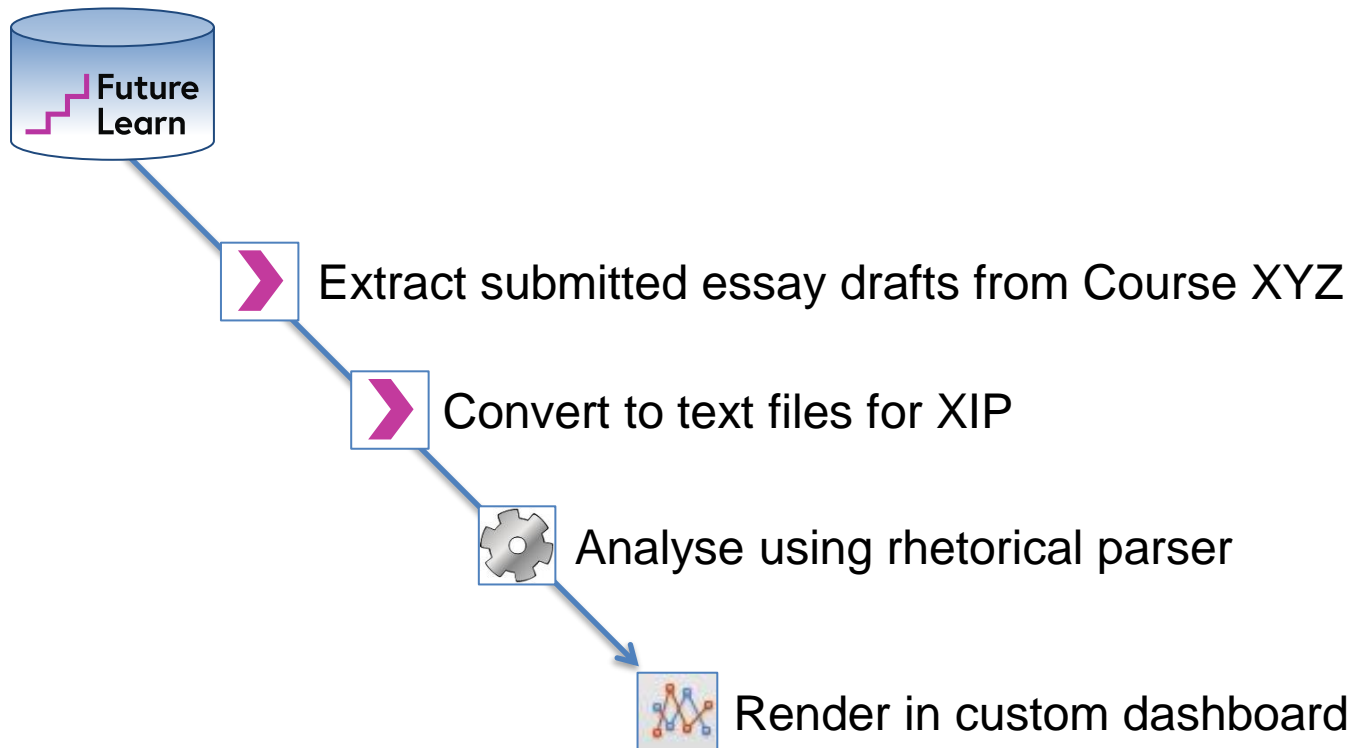




# Workflow for discourse analytics: *which sections of the course seem to provoke knowledge-building discourse?*



# Workflow for academic writing analytics: *to what extent does student writing display the hallmarks of scholarly argument?*



# Workflow for academic writing analytics: *to what extent does student writing display the hallmarks of scholarly argument?*

<body>SUMMARY Our work is a step towards building taxonomies based on both expert and data - driven approaches which we believe could lead to a needed trade-off between power and accuracy .

<body>EMPHASIS (The classic categorization of Austin [14] postulates five major speech act classes based on five categories of performative verbs :

<body>CONTRAST A drawback of the proposed model for representing dialogue utterances , i.e. the N leading tokens , is that the distance between two dialogue utterances is based on string operations rather than lexico - semantic distances which would be more meaningful for natural language dialogues .

**Figure 1: Raw XIP output showing sentence classification**

