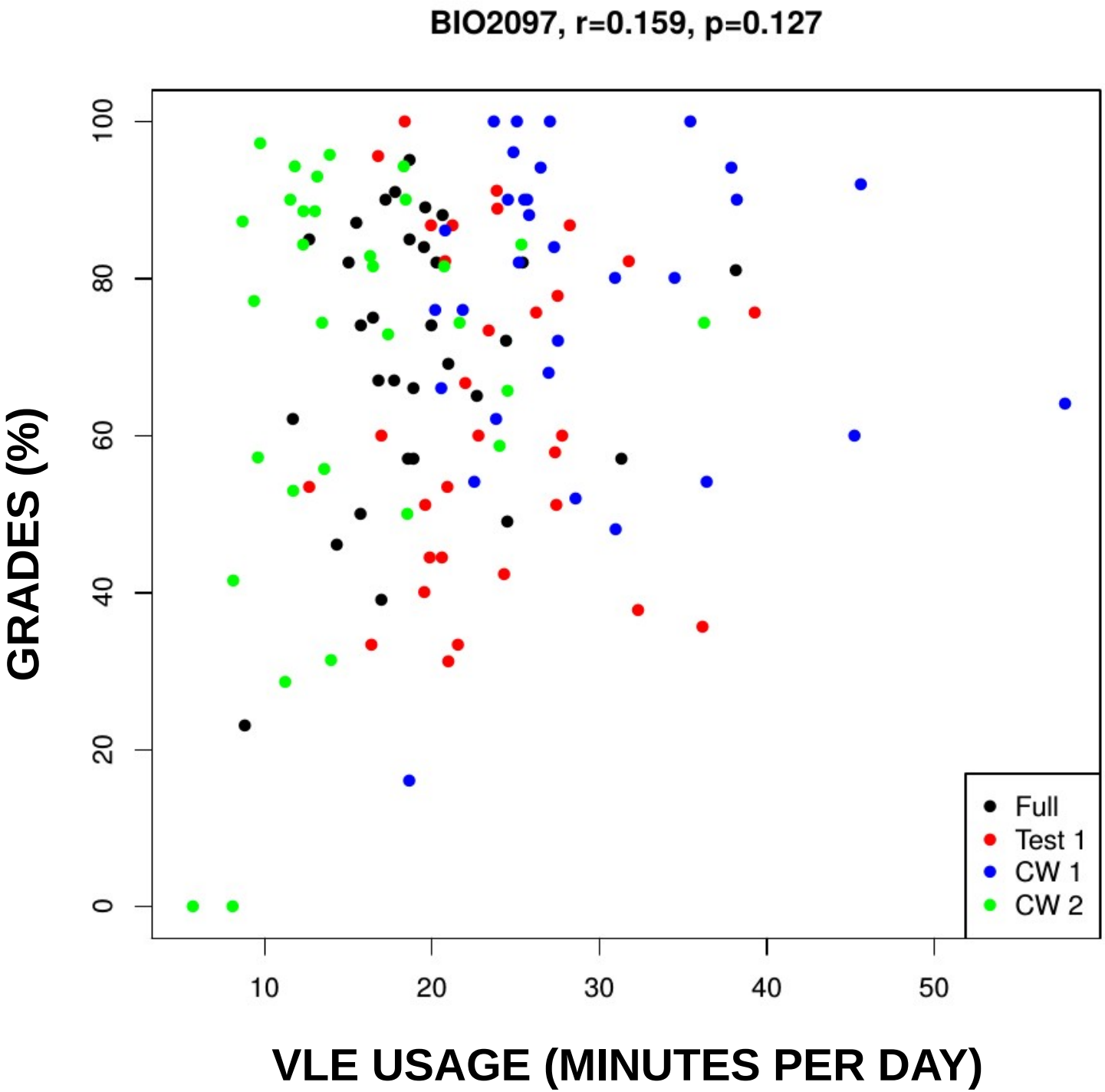
Engagement and success at a bricks-and-mortar University

Chris Boulton, Carmel Kent, Joanne Smith, Hywel Williams
+ the ELA team
University of Exeter

22nd February 2017
Ninth UK Learning Analytics Network meeting
@ University of Exeter



Engagement at BaM universities cannot be reduced to VLE use.

Is engagement at BaMs still predictive of student success?

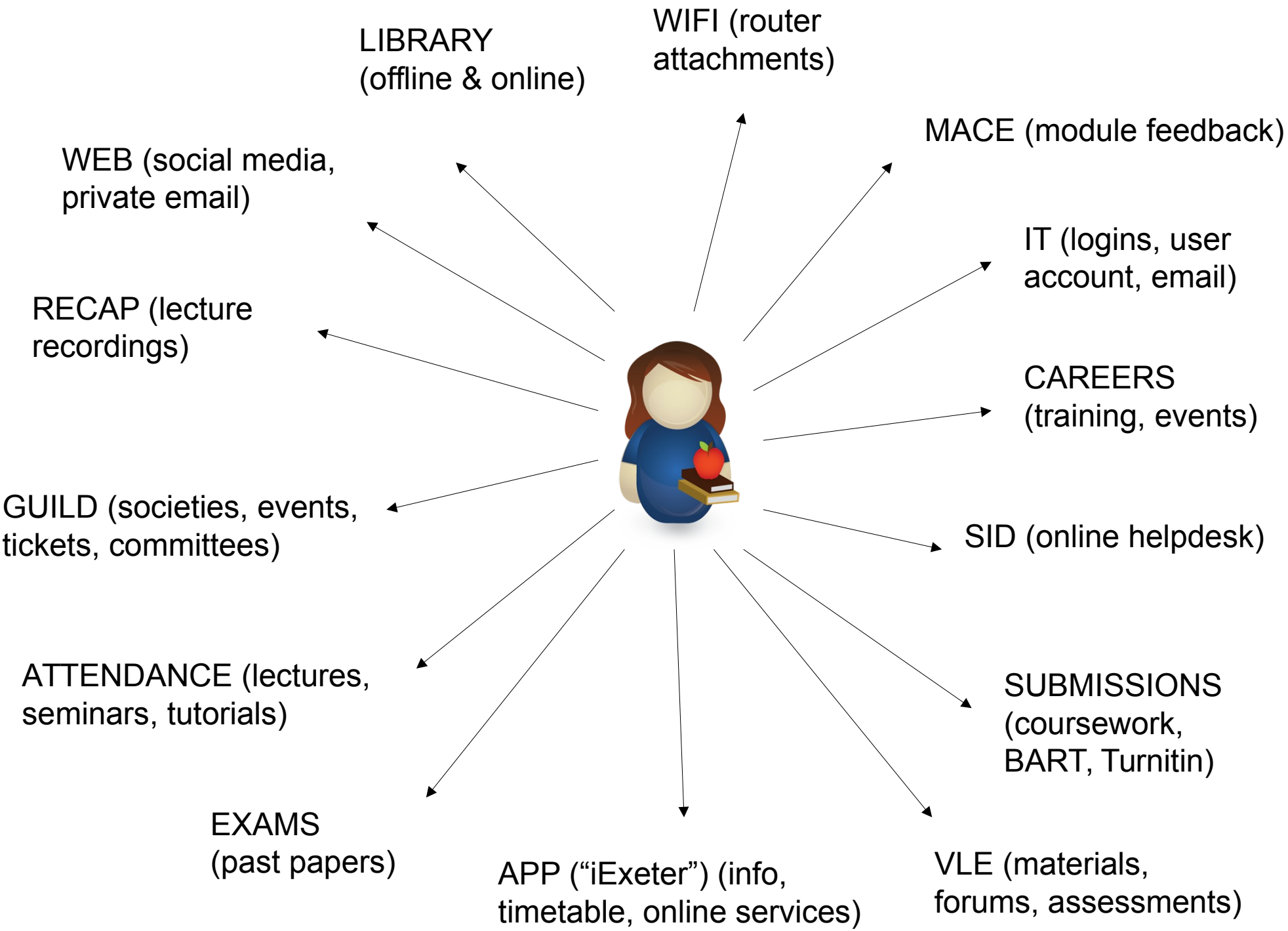
RQ1) How to measure engagement at a bricks-and-mortar University?

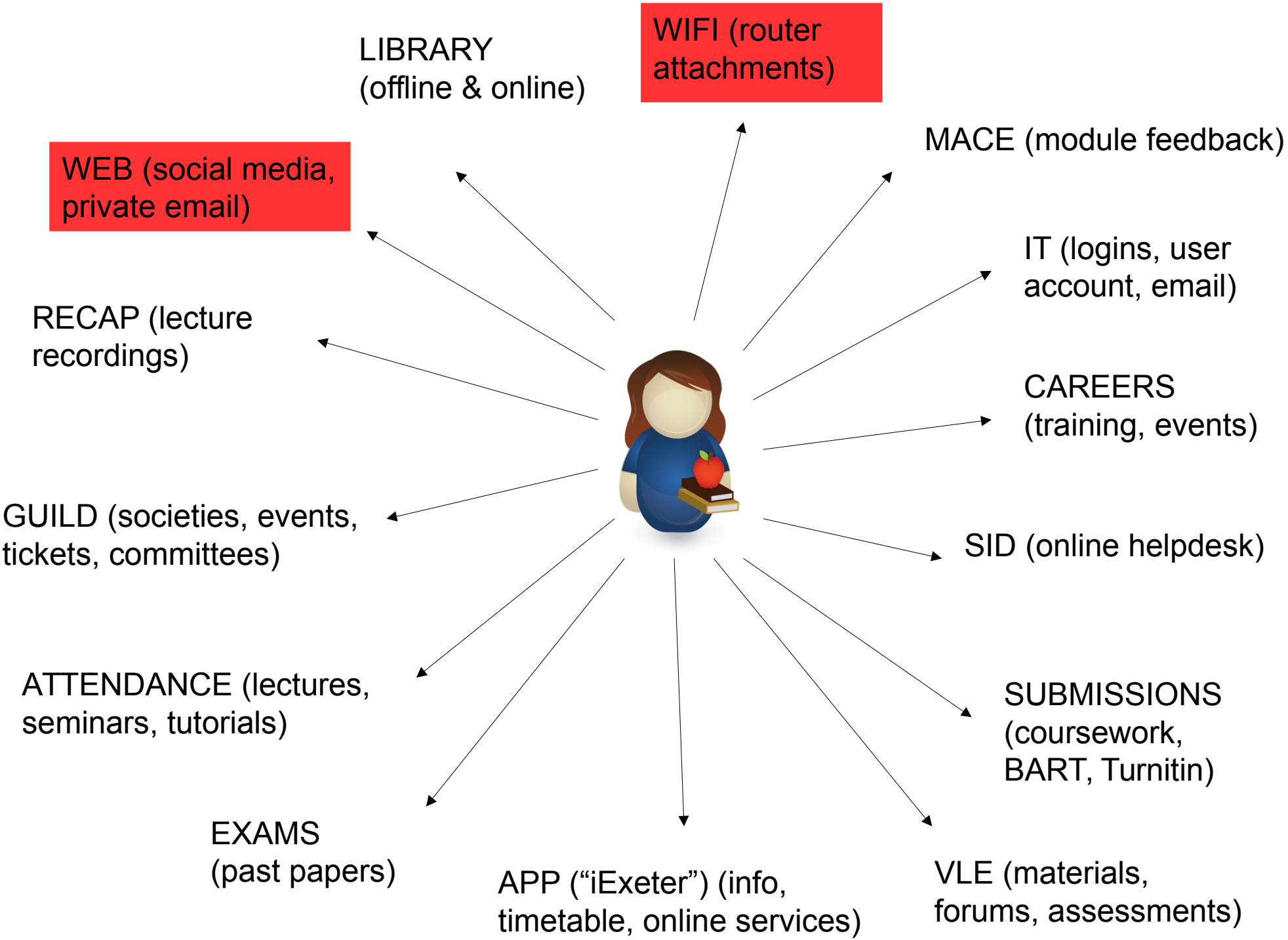
RQ2) Does engagement predict success?

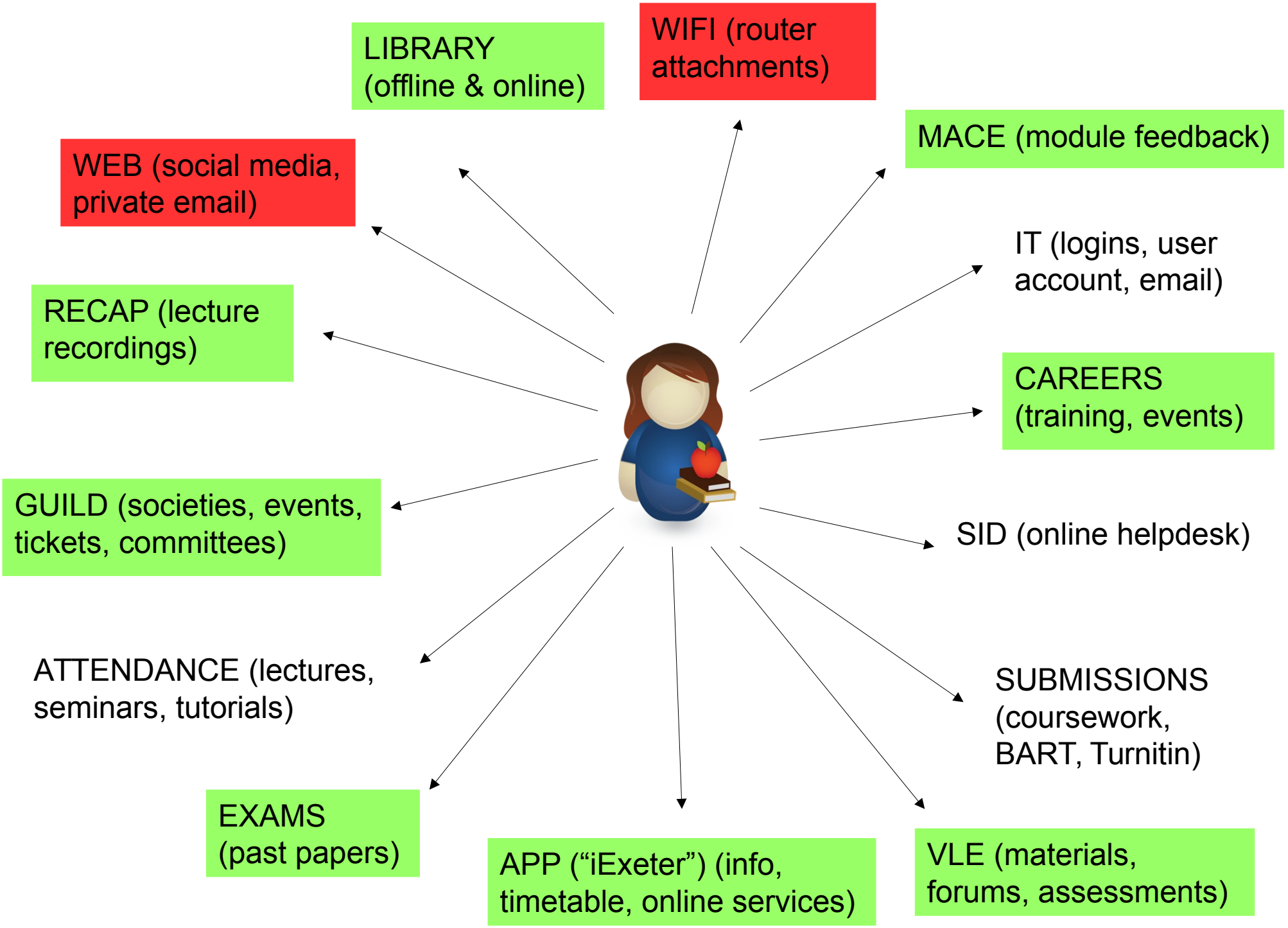
RQ3) Can we identify students in need of extra support?

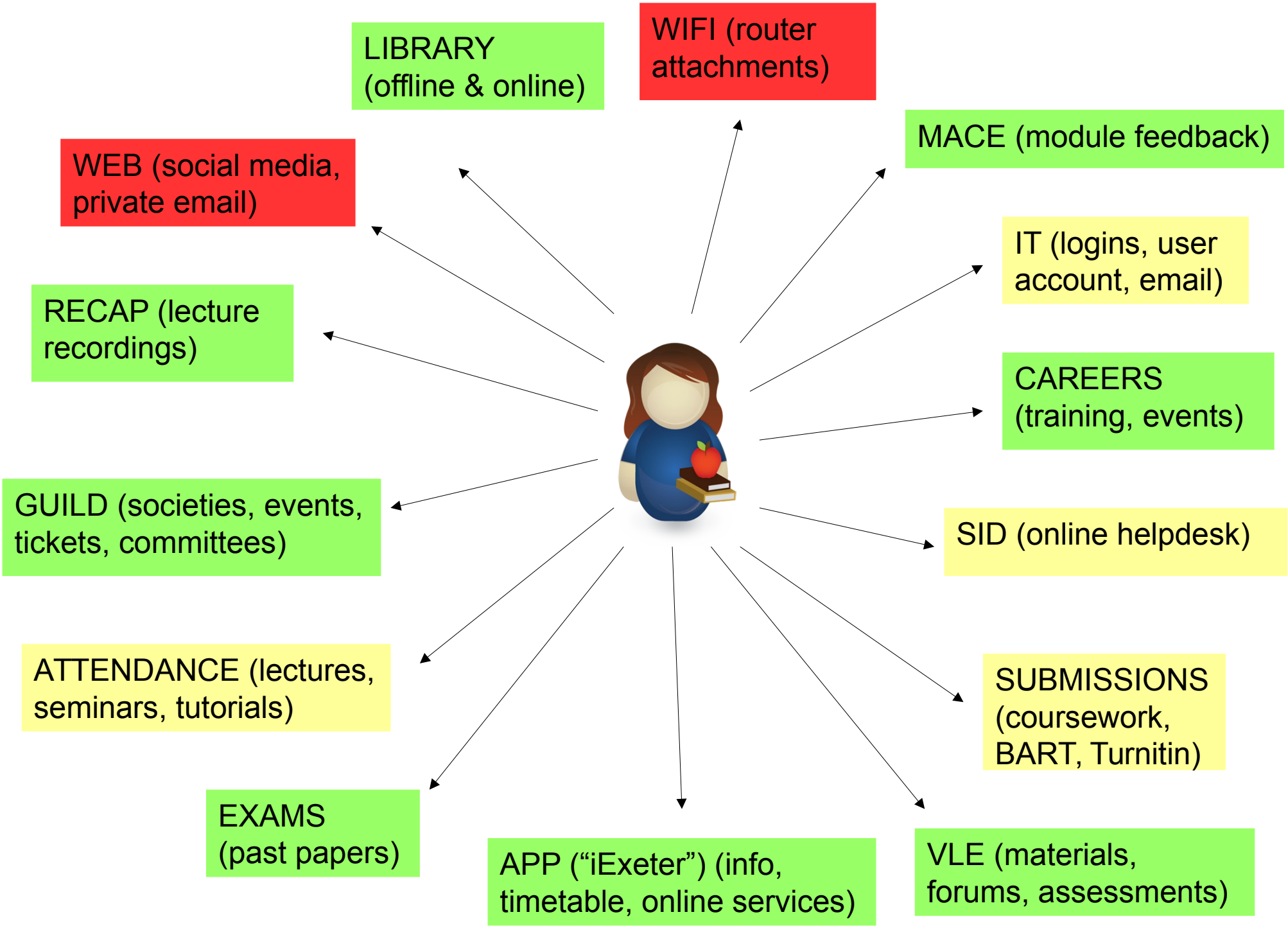
Measuring engagement

- Engagement has **many dimensions**...
  - Physical attendance at lectures
  - Interacting with staff
  - Being on campus
  - Student societies / sports / hobbies
  - Using digital resources (e.g. VLE, library)
  - Using external digital tools (e.g. social media)
  - ...
- Many of these leave **digital traces**
- **Pragmatic approach**: Focus on digital data that is routinely collected.
  - (Working closely with data warehouse project and IT managers.)

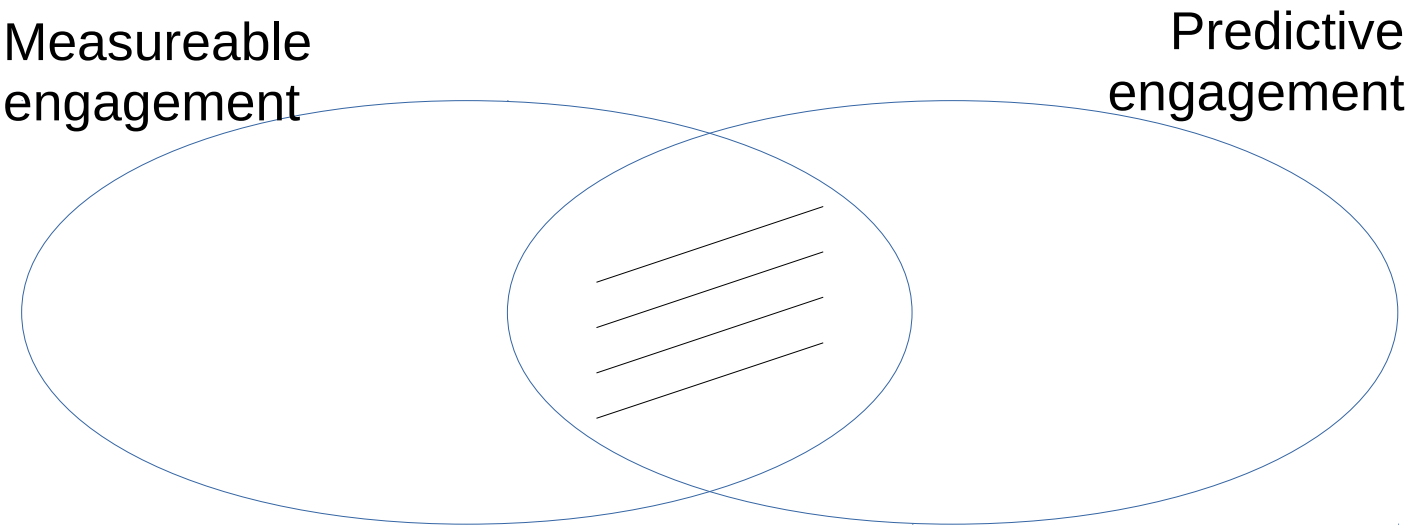

LIBRARY
(offline & online)
WIFI (router
attachments)
MACE (module feedback)
WEB (social media,
private email)
IT (logins, user
account, email)
RECAP (lecture
recordings)
CAREERS
(training, events)
GUILD (societies, events,
tickets, committees)
SID (online helpdesk)
ATTENDANCE (lectures,
seminars, tutorials)
SUBMISSIONS
(coursework,
BART, Turnitin)
EXAMS
(past papers)
APP (“iExeter”) (info,
timetable, online services)
VLE (materials,
forums, assessments)


LIBRARY
(offline & online)
WIFI (router attachments)
MACE (module feedback)
WEB (social media, private email)
IT (logins, user account, email)
RECAP (lecture recordings)
CAREERS
(training, events)
GUILD (societies, events, tickets, committees)
SID (online helpdesk)
ATTENDANCE (lectures, seminars, tutorials)
SUBMISSIONS
(coursework, BART, Turnitin)
EXAMS
(past papers)
APP ("iExeter") (info, timetable, online services)
VLE (materials, forums, assessments)


LIBRARY (offline & online)
WIFI (router attachments)
MACE (module feedback)
WEB (social media, private email)
IT (logins, user account, email)
RECAP (lecture recordings)
CAREERS (training, events)
GUILD (societies, events, tickets, committees)
SID (online helpdesk)
ATTENDANCE (lectures, seminars, tutorials)
SUBMISSIONS (coursework, BART, Turnitin)
EXAMS (past papers)
APP ("iExeter") (info, timetable, online services)
VLE (materials, forums, assessments)


LIBRARY (offline & online)
WIFI (router attachments)
MACE (module feedback)
WEB (social media, private email)
IT (logins, user account, email)
RECAP (lecture recordings)
CAREERS (training, events)
GUILD (societies, events, tickets, committees)
SID (online helpdesk)
ATTENDANCE (lectures, seminars, tutorials)
SUBMISSIONS (coursework, BART, Turnitin)
EXAMS (past papers)
APP ("iExeter") (info, timetable, online services)
VLE (materials, forums, assessments)

Measuring engagement



- Unclear which forms of engagement will be useful
  - → need to be both **measurable** and **predictive**
- Focus on digital data may introduce bias
  - → running a complementary engagement survey (ongoing)

RQ1) How to measure engagement at a bricks-and-mortar University?

RQ2) Does engagement predict success?

RQ3) Can we identify students in need of extra support?

Operationalising engagement

Variable	System
Event sign-ups	Careers Events System
Event attendance	Careers Events System
Event follow-through (attendance/sign-ups)	Careers Events System
Logins	Virtual Learning Environment (VLE)
Logins	Inter Library Loans (Library ILL)
Logins	Library fines system.
Number of fines paid	Library fines system
Logins	MACE (Module and Course Evaluation) system.
Evaluations submitted	MACE
Logins	Exam archive (past papers)
Paper views	Exam archive
...	...

Defining “success”

- Many kinds of student, many student goals:
  - Completion, grades, employment, awards, training/skills, sports, social life, wellbeing...
- Some variables can be measured, others can not
- So far: grades and completion (others to follow)

Variable	System
Average number of attempts per module	SITS
Average results for all modules (weighted by credit value)	SITS
Module pass rate	SITS
Average deviation from module average	SITS

Entire cohort – predictors of average credit-weighted module grade

Demographic	**Engagement**
Gender [U = 231120953.50**]	MACE evaluations [r = 0.250**]
Away from home [U= 152140073.00**]	MACE logins [r = 0.262**]
Disability type [H(10) =168.02**]	
Disability [H(3)=73.89**]	
Country of domicile [H(140)=1,554.98**]	
Ethnicity [H(18)=627.97**]	
National identity [H(7)=360.69**]	
Nationality [H(187)=1,880.03**]	
Parents' occupational [H(326)=869.74**]	

Statistics: r – Spearman's, U – Mann-Whitney, H – Kruskal-Wallis
** - significant at p<0.01
Sample: n=30,781 students in three years

Entire cohort – predictors of average credit-weighted module grade

Coefficients	Estimate	Std. Error	t value	p-value
(Intercept)	52.446	0.823	63.751	$<2*10^{-16}$ **
Gender (Male)	-0.239	0.19286	-1.242	0.2143
Age at beginning of year	-0.328	0.035	-9.399	$<2*10^{-16}$ **
Away from home	5.973	0.224	26.616	$<2*10^{-16}$ **
Disability Type (Unknown)	2.107	1.936	1.088	0.2764
Disabled (Yes)	-2.834	0.284	-9.978	$<2*10^{-16}$ **
log(events attended + 1)	2.832	0.132	21.495	$<2*10^{-16}$ **
Committee interactions	8.914	0.489	18.215	$<2*10^{-16}$ **
log(VLE + 1)	-1.944	0.085	-22.870	$<2*10^{-16}$ **
log(Past exams + 1)	0.130	0.067	1.937	0.0528 *
log(Library logins + 1)	4.842	0.302	16.009	$<2*10^{-16}$ **
log(MACE + 1)	9.224	0.187	49.364	$<2*10^{-16}$ **

Multiple regression model: $F(11, 44425) = 512.7$, $p<2.2e\ -16$, $R^2 = 0.1126$, Adjusted $R^2 = 0.1124$, residual standard error=20.08 .
Sample: n=30,781 students in three years 2013-2014-2015

Biosciences discipline – Different predictors for high/low performing students

	Demographics	Engagement
Low performers (*)	Age at start of year debt Gender Distant_home Disability Ethnicity Parents occupation	exFactor attended exFactor signups exfactor_attend_percentage
High performers	debt Gender Overseas National identity Ethnicity Parents occupation	grand_challenge_signups grand_challenge_attended grand_challenge_percentage exfactor_attend_percentage exFactor signups

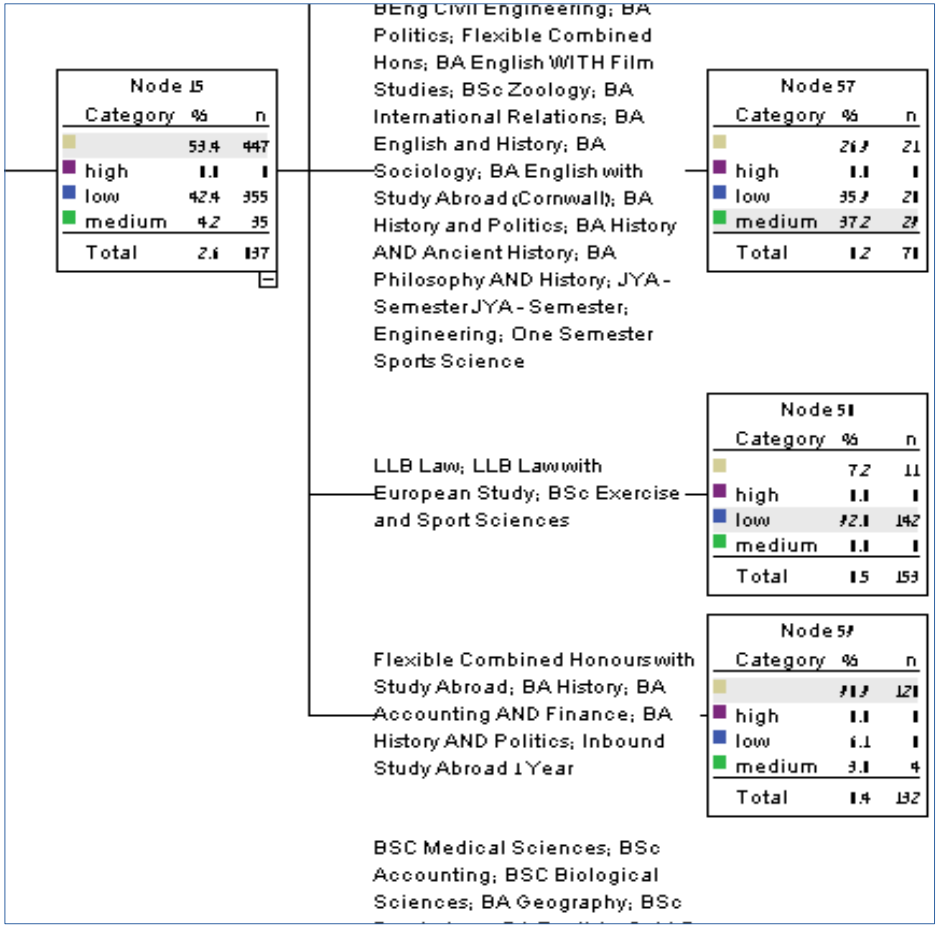
* Low performers also predicted by: #module assessment types, total interruption duration, avg class size.

Biosciences cohort: n=3355 (student-year records, 2270 students), 27 programmes, years 2014/2015.

Different predictors for different disciplines

	Demographics	**Engagement**
BIOSCIENCES	Is_debt Gender Distant_home Ethnicity	MACE- logins MACE- evaluations ticket membership grand_challenge_signups grand_challenge_attended grand_challenge_percentage leadership_management_percentage exeter_leader_award_attend_perc exfactor_attend_percentage exFactor signups
ENGLISH	Is distant Gender National identity Nationality Country of domicile Ethnicity Disability Parent's occupation Spoken language	MACE- logins MACE- evaluations

Decision trees: Predicting grades



Aim: Use all variables to predict high/low/medium grades

Promising outputs, e.g. more than 75% of “low” grade...

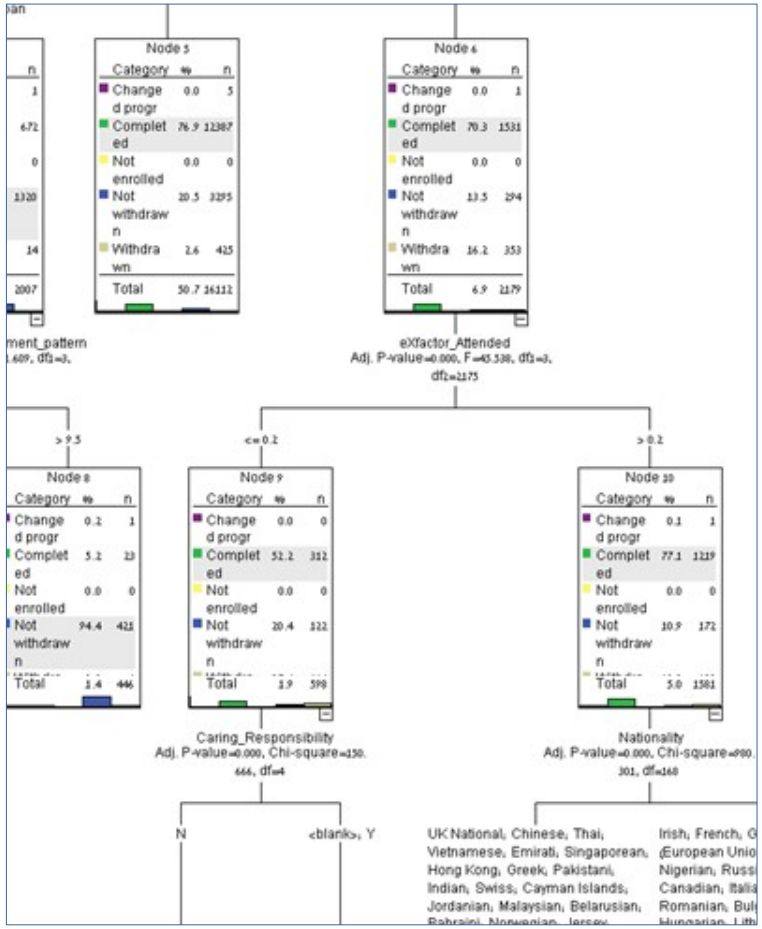
> “[iExeter views in term 1 <= 5.727273]
>
> AND
>
> [number of module assessment types are 6 or 7]
>
> AND [campus is St. Lukes OR Streatham]”

RQ1) How to measure engagement at a bricks-and-mortar University?

RQ2) Does engagement predict success?

RQ3) Can we identify students in need of extra support?

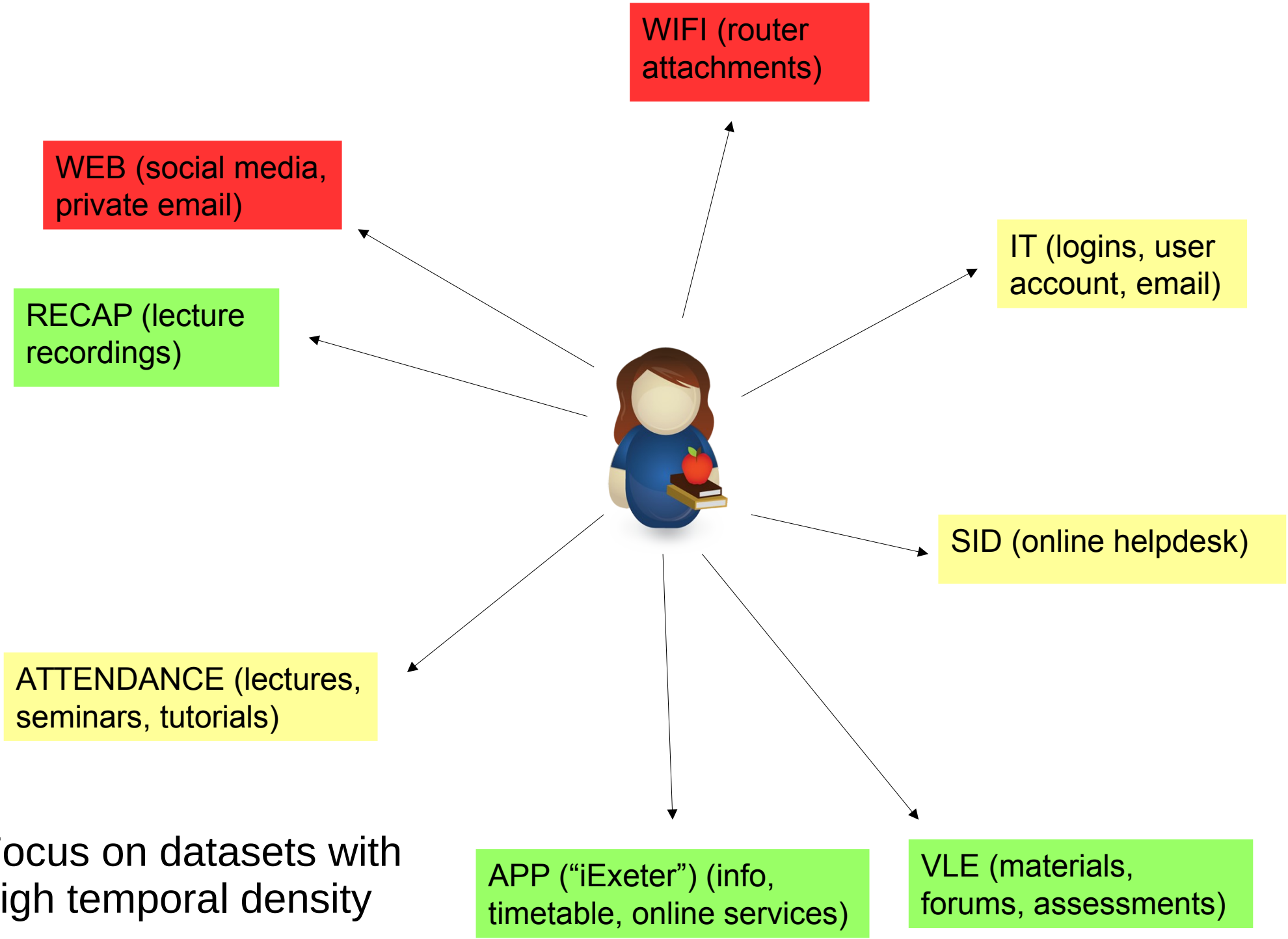
Decision trees: Predicting withdrawal



Aim: Use all variables to predict withdrawals vs completition, identify at-risk groups.

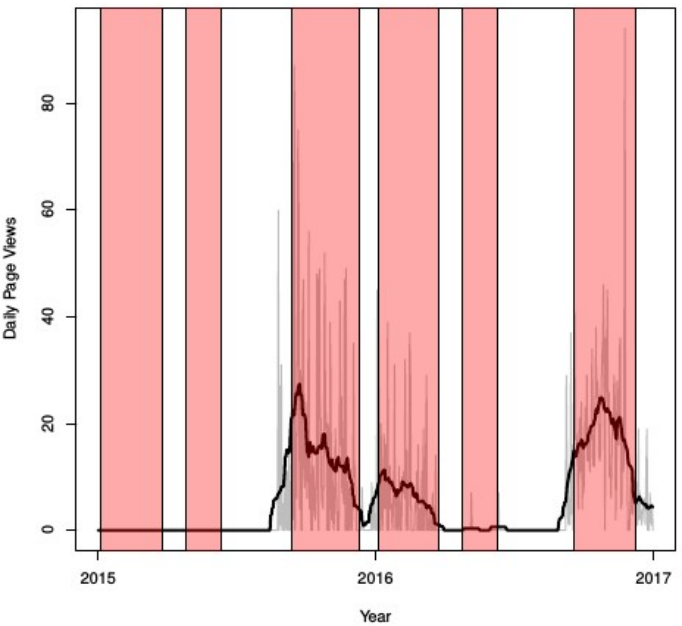
Promising initial results:

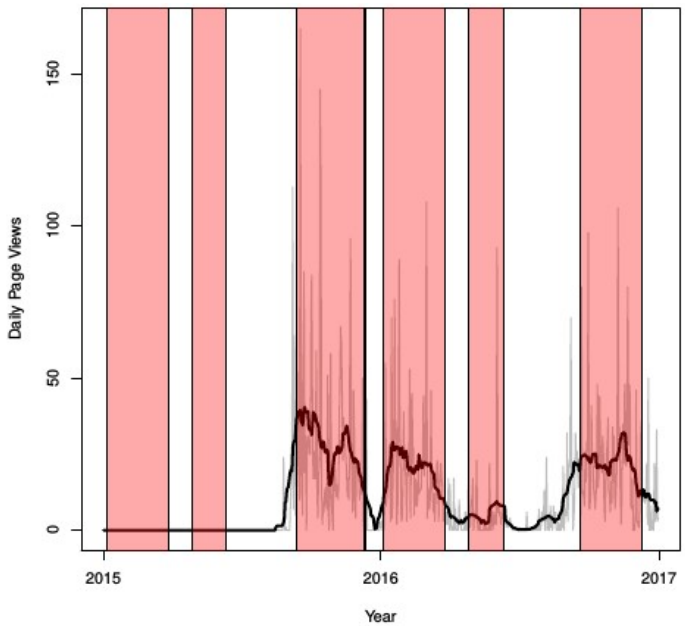
- E.g. 50% likelihood of withdrawal for students who sign up to exFactor events but do not attend, and who have caring responsibilities.

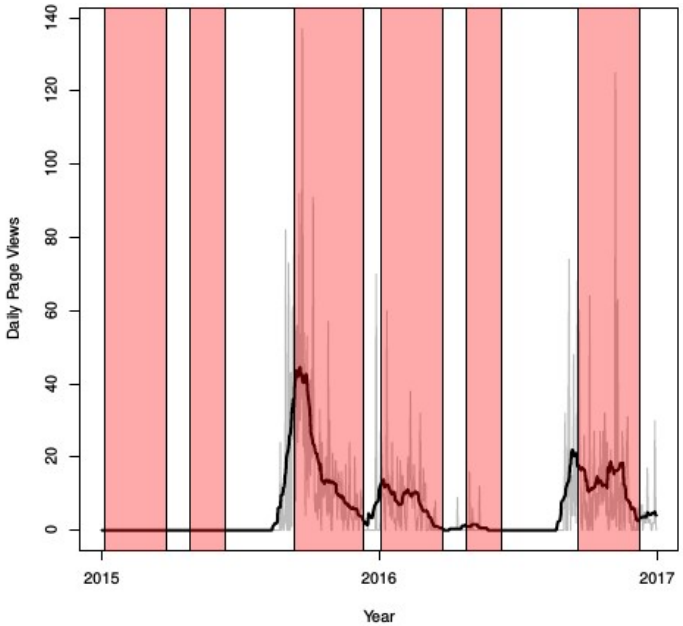


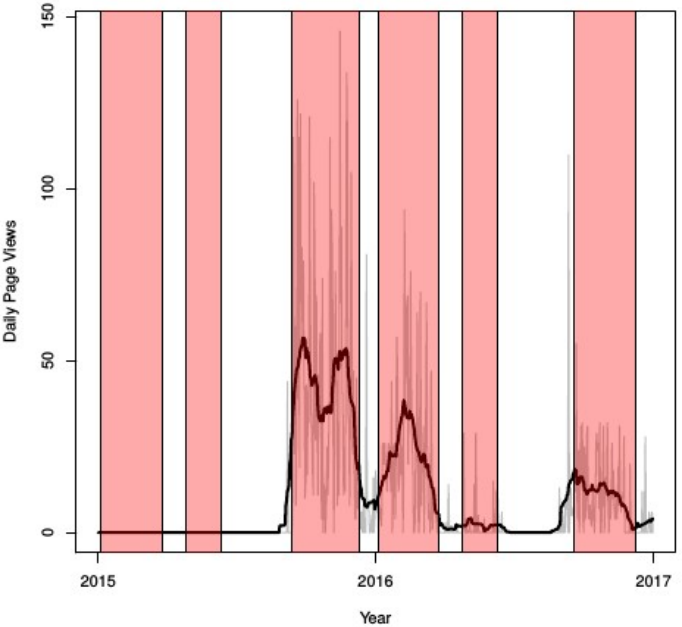
Focus on datasets with high temporal density

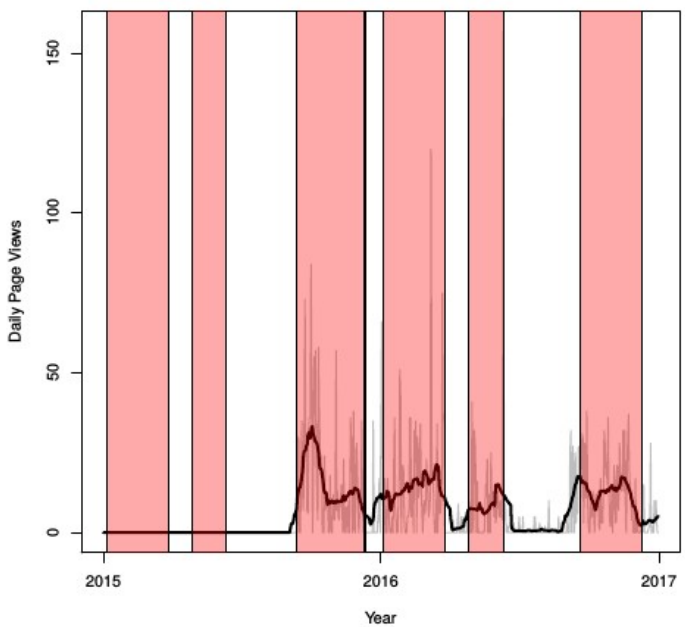
iExeter

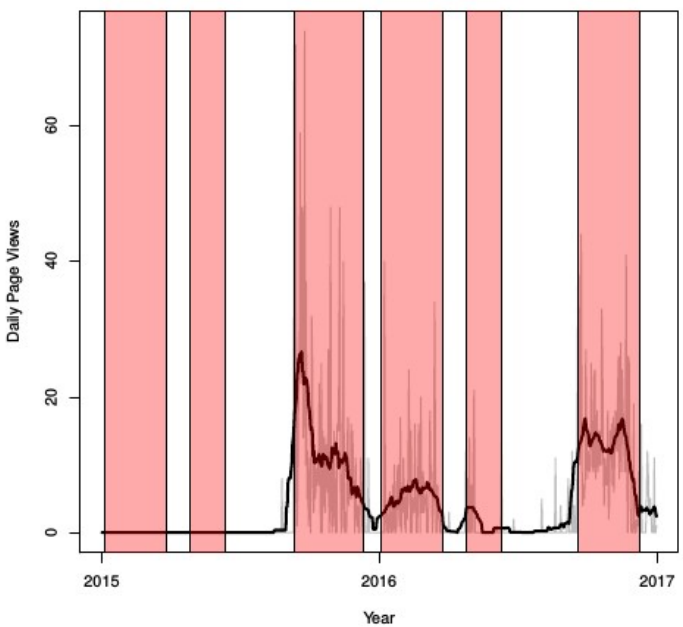

Daily Page Views
0
20
40
60
80
2015
2016
2017
Year


Daily Page Views
0
50
100
150
2015
2016
2017
Year

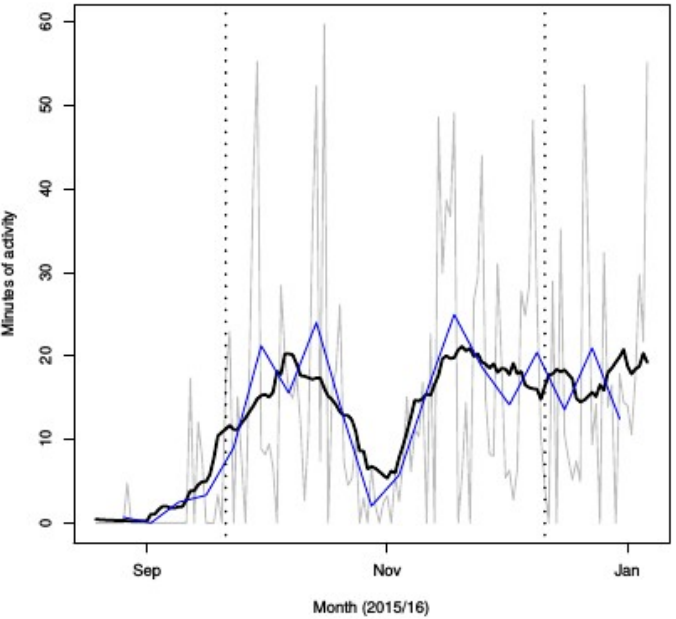

Daily Page Views
0
20
40
60
80
100
120
140
2015
2016
2017
Year

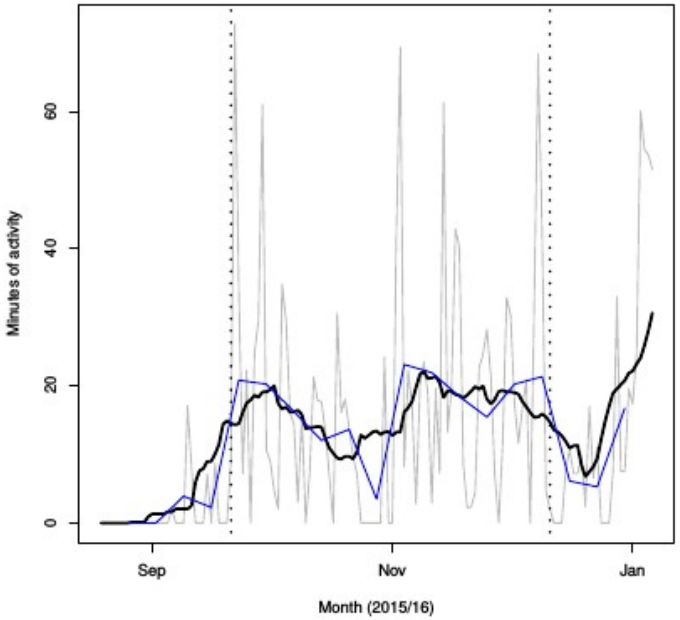

Daily Page Views
0
50
100
150
2015
2016
2017
Year

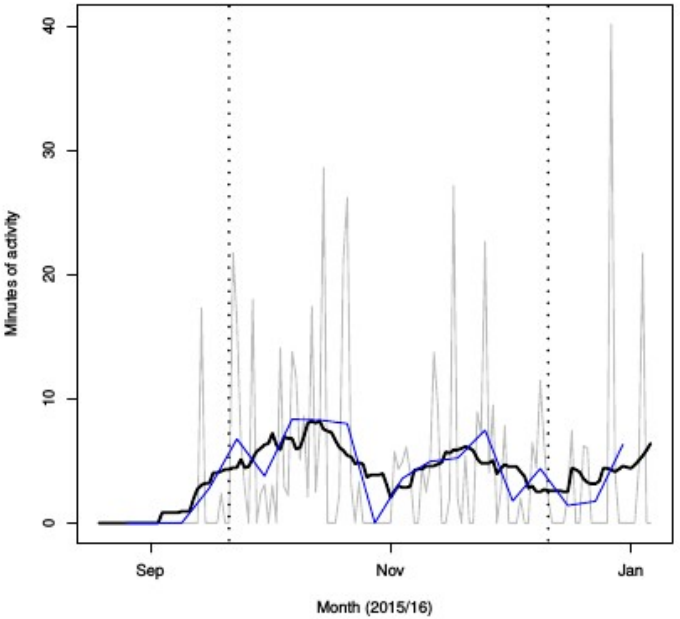

Daily Page Views
0
50
100
150
2015
2016
2017
Year

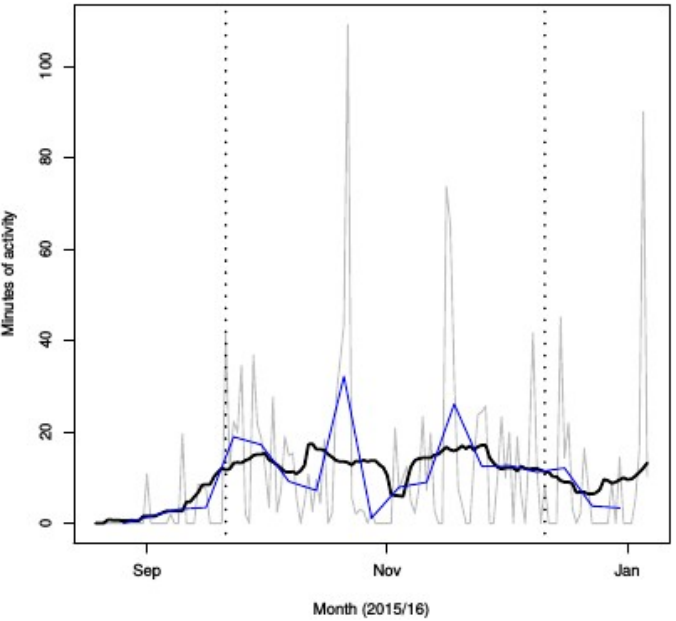

Daily Page Views
0
20
40
60
2015
2016
2017
Year

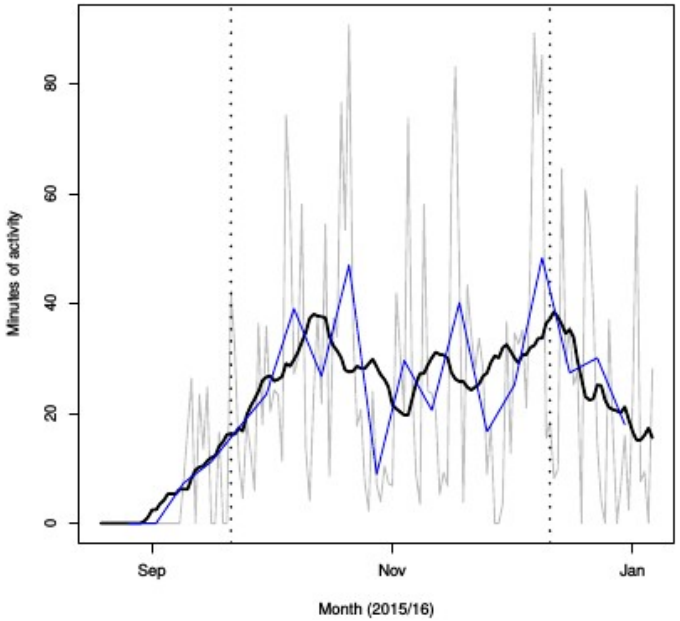
ELE

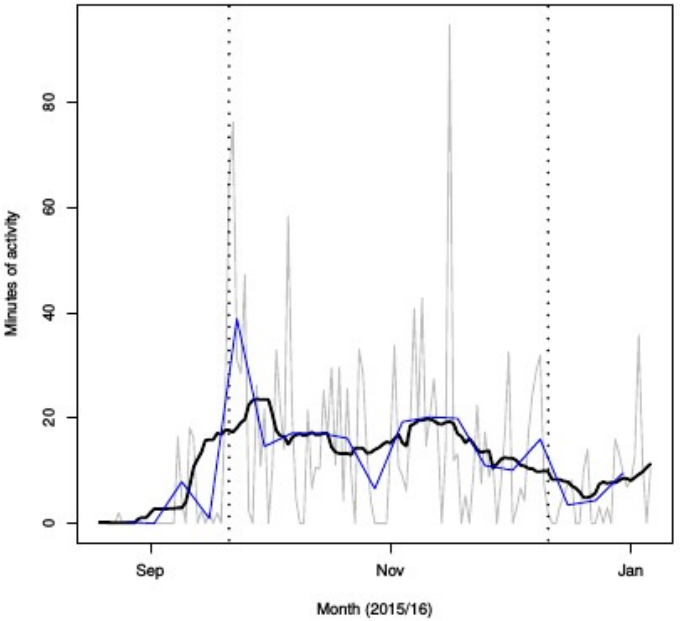

Minutes of activity
0
10
20
30
40
50
60
Sep
Nov
Jan
Month (2015/16)


Minutes of activity
0
20
40
60
Sep
Nov
Jan
Month (2015/16)

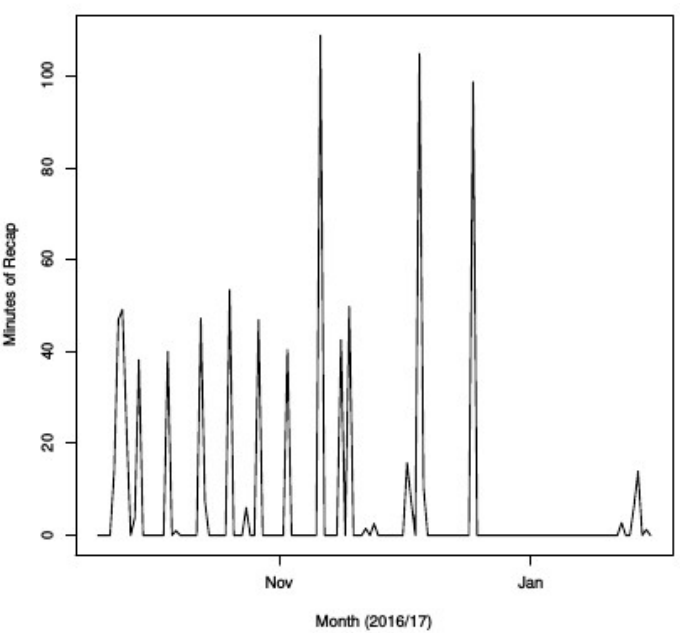

Minutes of activity
0
10
20
30
40
Sep
Nov
Jan
Month (2015/16)

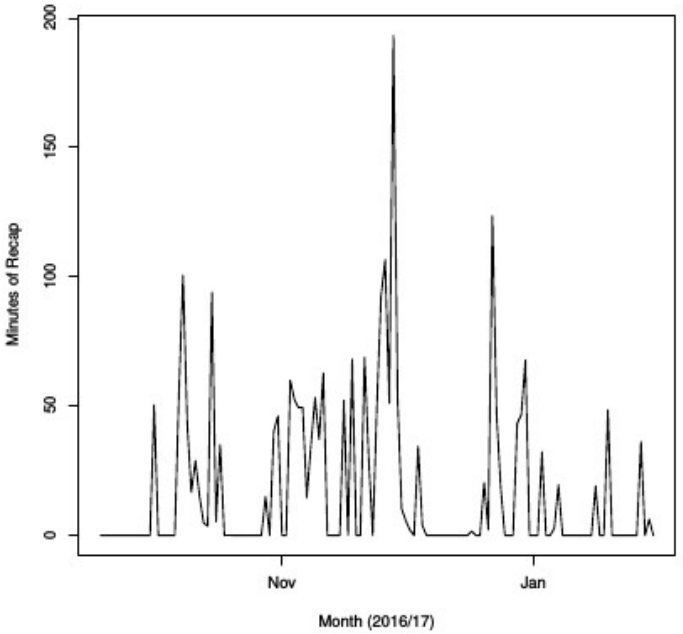

Minutes of activity
0
20
40
60
80
100
Sep
Nov
Jan
Month (2015/16)

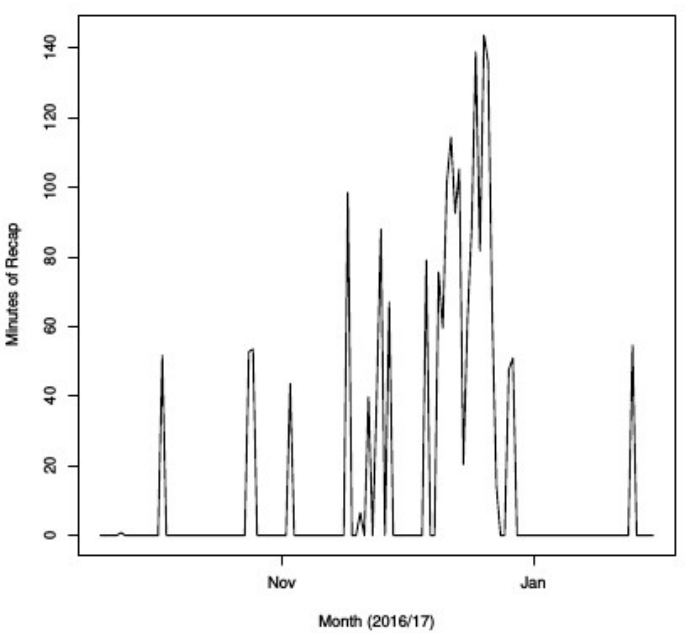

Minutes of activity
0
20
40
60
80
Sep
Nov
Jan
Month (2015/16)

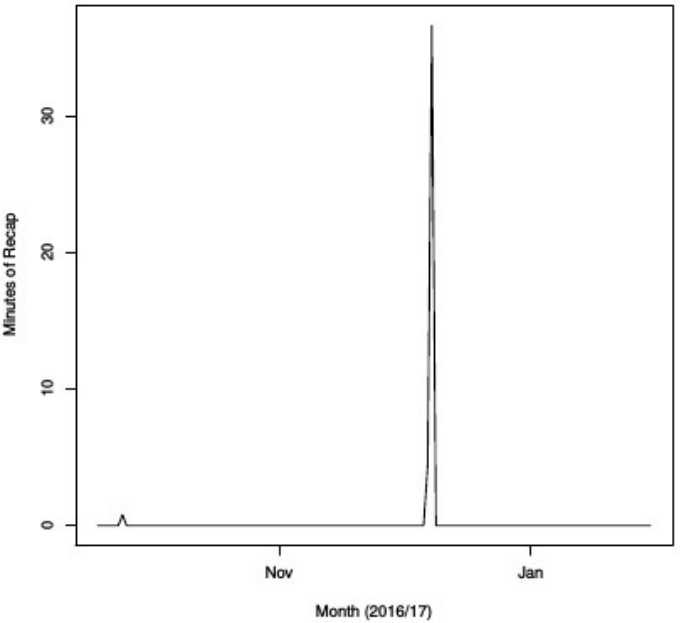

Minutes of activity
0
20
40
60
80
Sep
Nov
Jan
Month (2015/16)

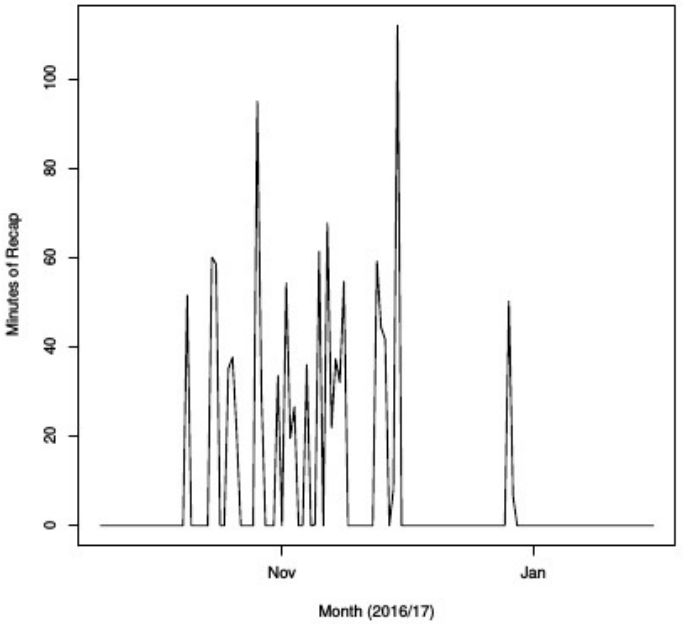
Recap

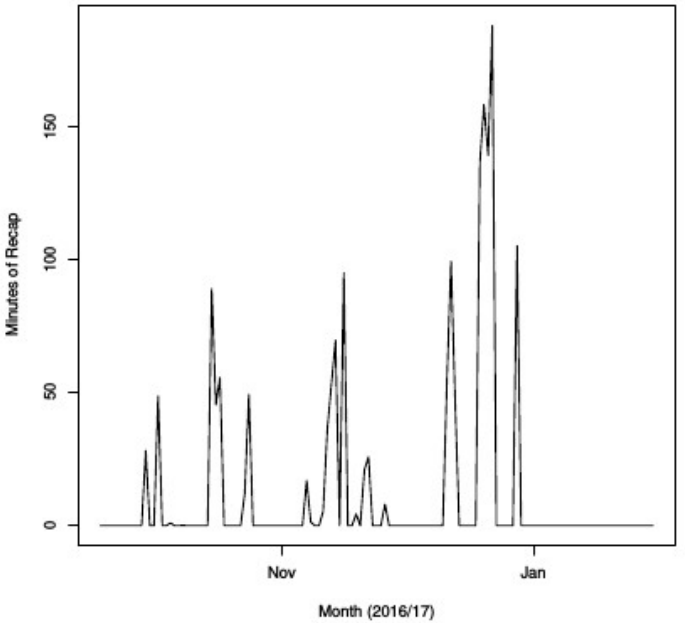

Minutes of Recap
0
20
40
60
80
100
Nov
Jan
Month (2016/17)


Minutes of Recap
0
50
100
150
200
Nov
Jan
Month (2016/17)

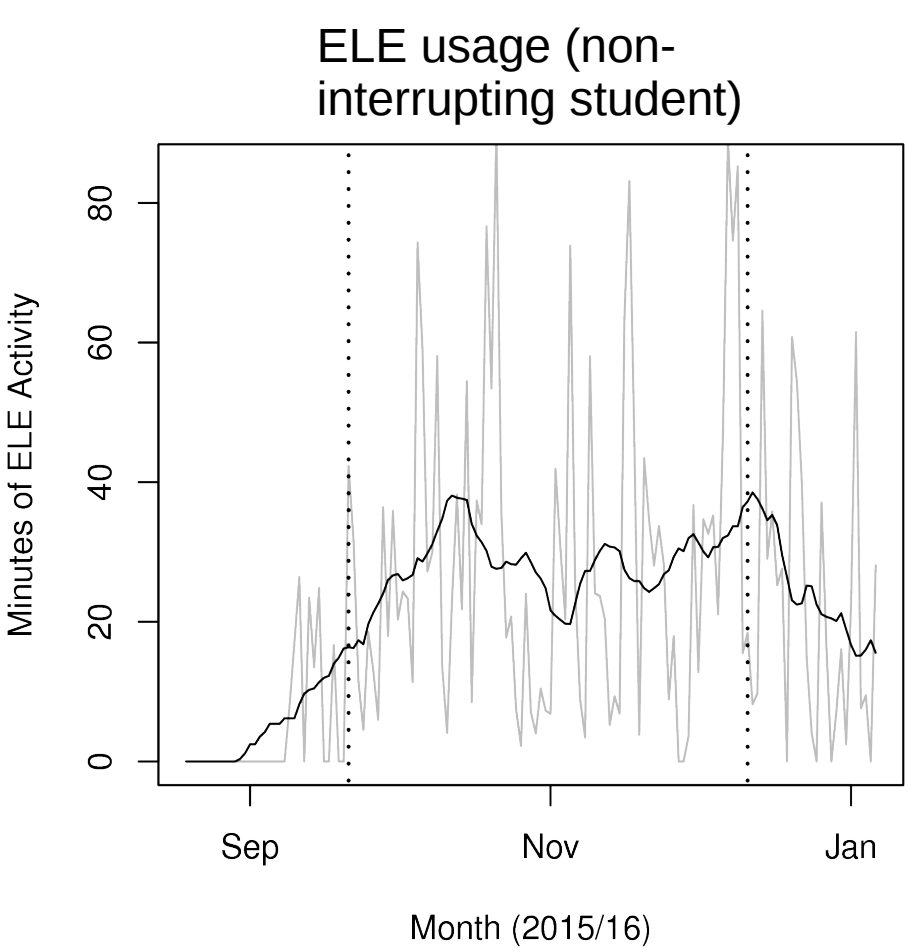

Minutes of Recap
0
20
40
60
80
100
120
140
Nov
Jan
Month (2016/17)

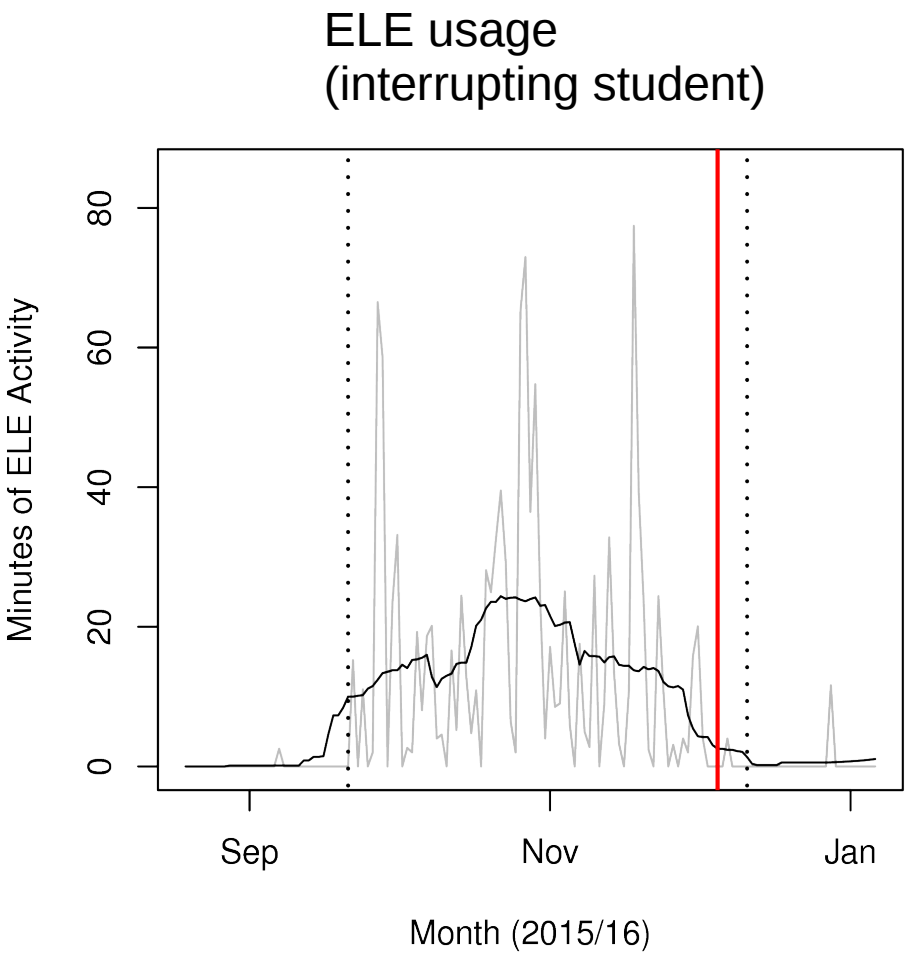

Minutes of Recap
0
10
20
30
Nov
Jan
Month (2016/17)


Minutes of Recap
0
20
40
60
80
100
Nov
Jan
Month (2016/17)


Minutes of Recap
0
50
100
150
Nov
Jan
Month (2016/17)

Does system usage change prior to interruption events?





Discussion

- **Caveats**: Early stages of analysis. Several datasets not examined. Findings not yet confirmed. Too early to draw robust conclusions.
- **Methods** are evolving:
  - A lot of data-wrangling needed before analysis could begin
  - Different modes of analysis: Aggregated (post hoc, population-level) vs Dynamic (real-time, individual-level)
- **Ethics**: How should this data be used?
  - “Knowledge is power” vs “first do no harm”
  - How robust do models need to be to (e.g.) guide interventions?
- **Next steps:** Complete survey. Continue analysis. Import more datasets. Inform learning analytics strategy.

Acknowledgements

- Thanks to the Business Intelligence project team and IT department for lots of help extracting data!

Dr Carmel Kent

Dr Chris Boulton

Demographic variables

Variable	System
Gender	Registration
Disability (type of disability)	Registration
National identity	Registration
Nationality	Registration
Country of domicile	Registration
Ethnicity	Registration
Age at enrollment	Registration
Age at start of year	Registration
Living away from home	Registration
Parents' occupational background	Registration

Temporal correlation between systems