GEORG ECKERT INSTITUT
Leibniz-Institut für internationale Schulbuchforschung

# DATA AND EDUCATION

Felicitas Macgilchrist

Georg-Eckert-Institut for International Textbook Research
and Georg-August University Göttingen
macgilchrist@leibniz-gei.de
@discoursology

Data in Discourse Analysis | TU Darmstadt | 19 February 2020


CasCard

LEARNING ECONOMY

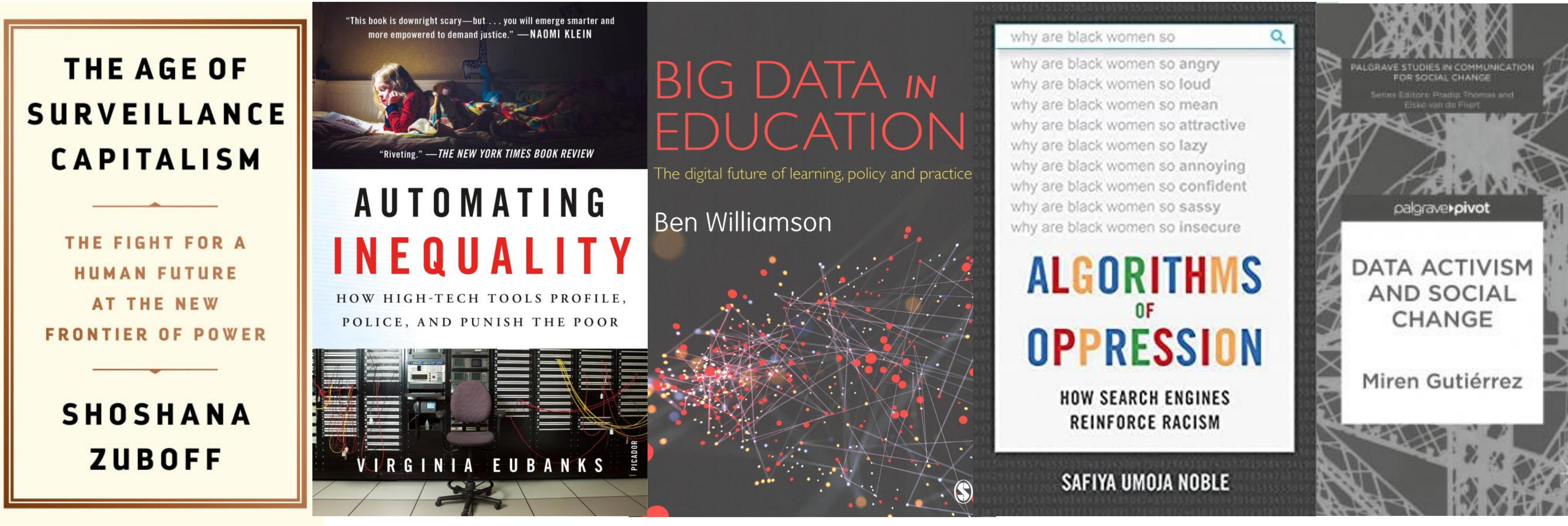
THREAD THE NEEDLE

# Quantify Outcomes

By measuring the **value chain of education** we can stand up a 22nd century-ready learning ecosystem that quantifies impact and lets us solve many of our toughest challenges.

DOWNLOAD CASE STATEMENT

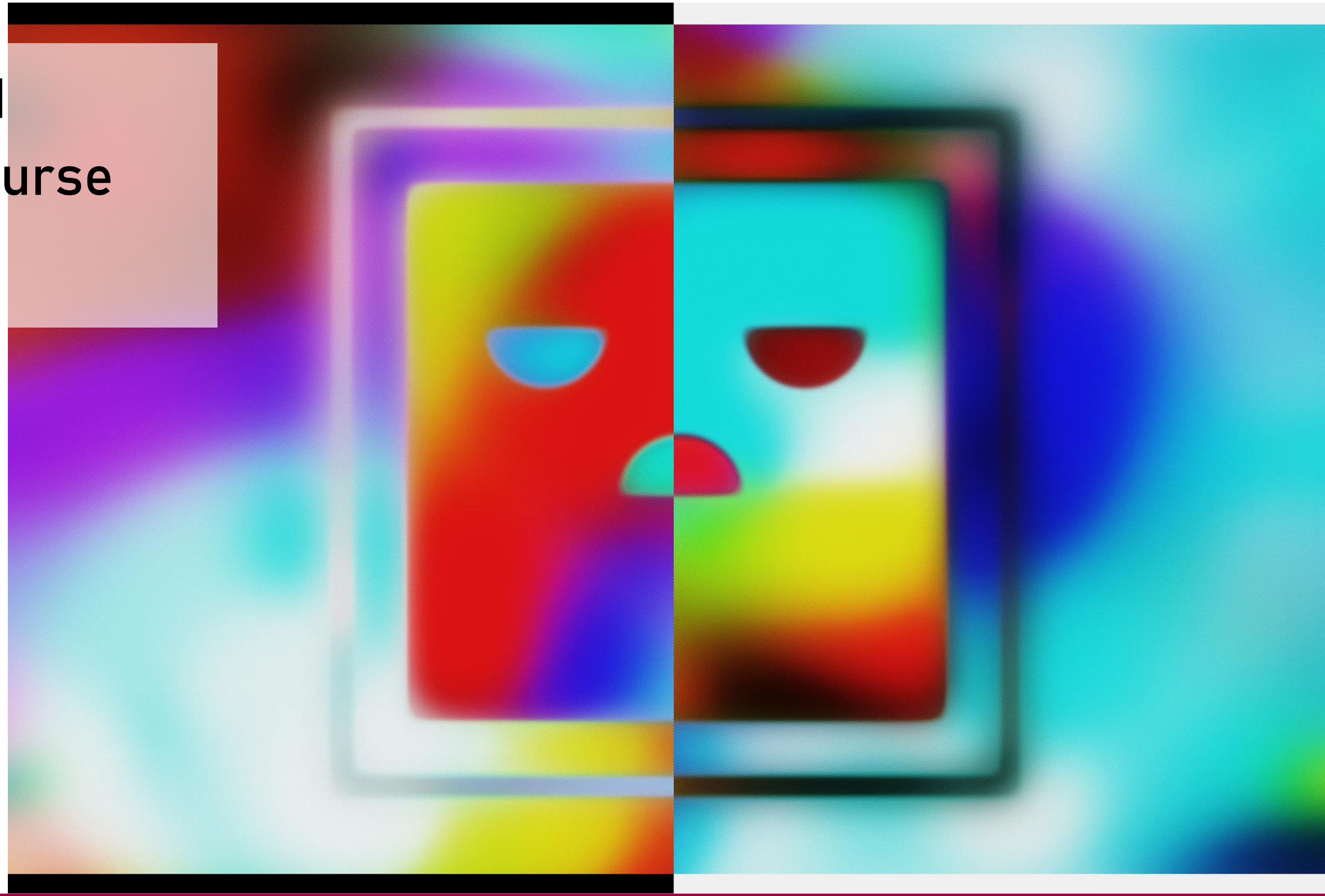
# DATAFIED TIMES






GEORG ECKERT INSTITUT
Leibniz-Institut für internationale Schulbuchforschung

“How has data, its encoding and markup, changed the way discourse studies perceive their subject?”



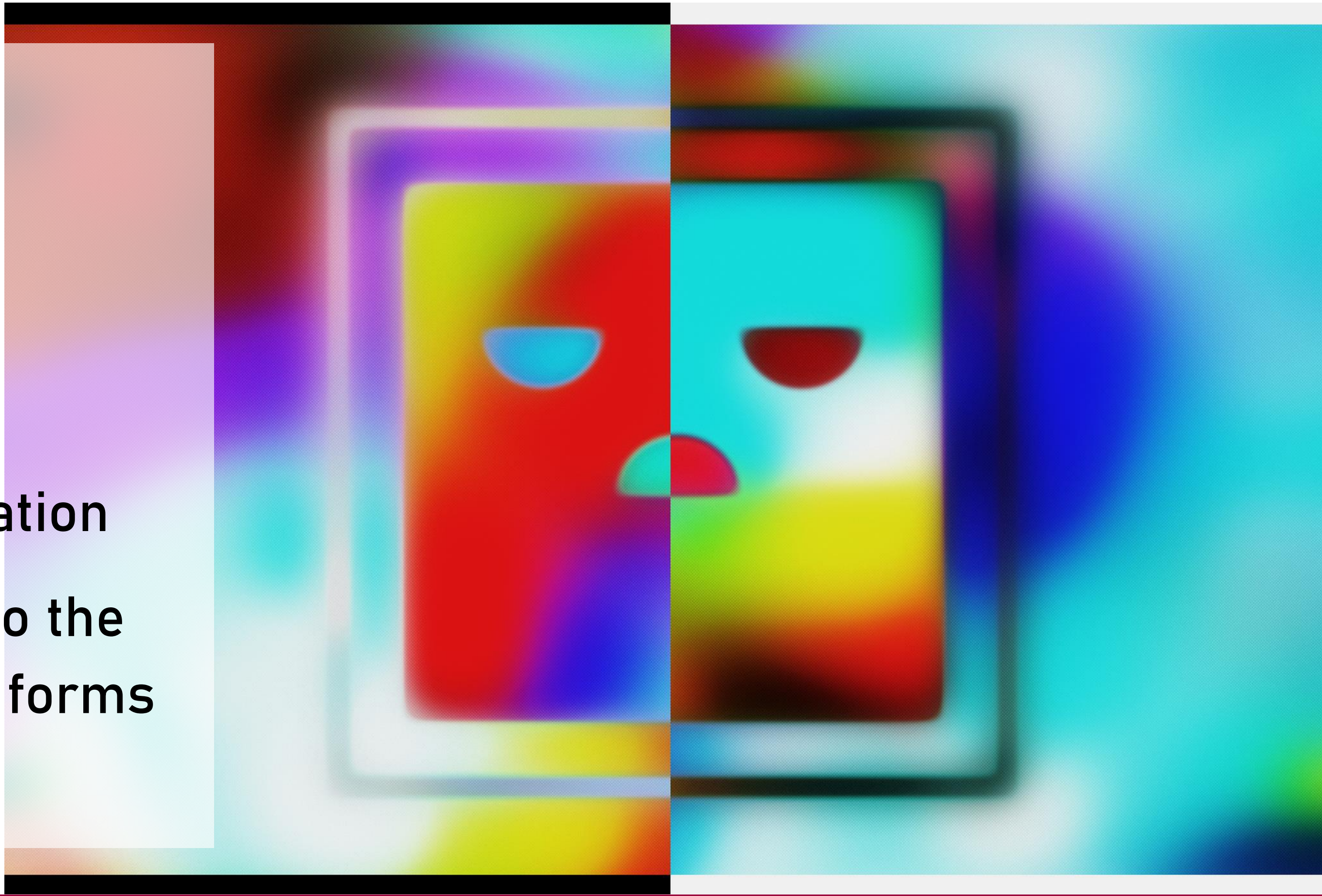
Augmented Nonsense CC BY 2.0 Surian Susay

Data changes:

0) Corpora gei.digital The Digital Textbook Library

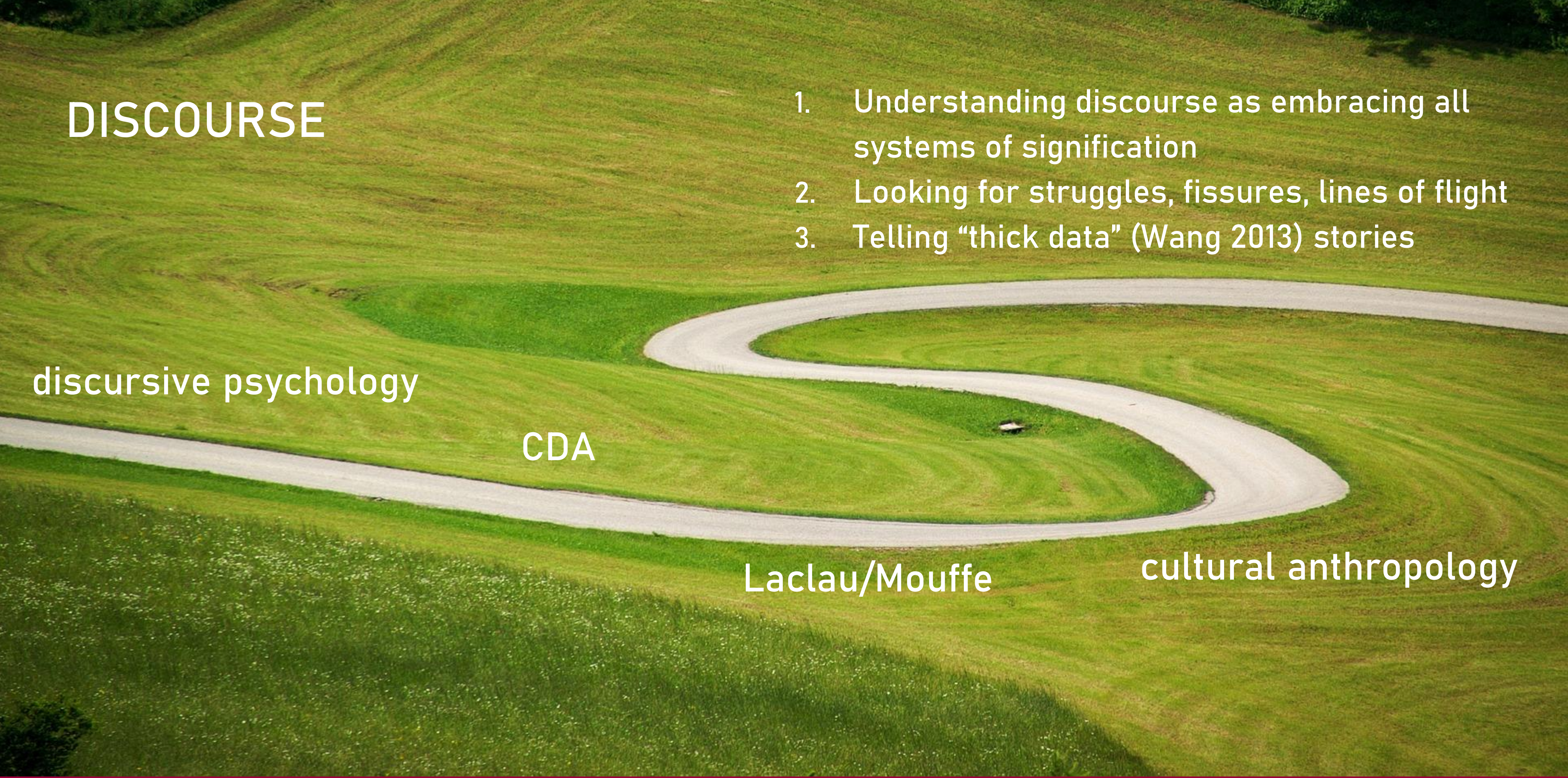
1) Discourse about education
2) Discourse in education
3) Discourse encoded into education

How do discourse studies rise to the challenges posed by these new forms of data?





Augmented Nonsense CC BY 2.0 Surian Susay



# DISCOURSE

1. Understanding discourse as embracing all systems of signification
2. Looking for struggles, fissures, lines of flight
3. Telling “thick data” (Wang 2013) stories

discursive psychology

CDA

Laclau/Mouffe

cultural anthropology



GEORG ECKERT INSTITUT
Leibniz-Institut für internationale Schulbuchforschung

# [1] DISCOURSE “ABOUT” EDUCATION

- Datafication of education through
  - (a) large-scale assessments (PISA, ICILS, TIMMS)
  - (b) data analytics
- New forms of marketization and psychologization of education: Through the promises of new actors (Pearson, IBM, etc.) and performed through software in educational settings
- New modes of analysis: Global networks of “policy experts”

(Ball/Junemann/Santori 2017; Roberts-Holmes/Bradbury 2016; Roberts-Mahoney/Means/Garrison 2016; Rowe 2019; Williamson 2017, 2019 )

# [2] DISCOURSE “IN” EDUCATION

- National testing goals
- School grade targets
- “Data conversations” with teachers and senior staff about specific students
- Tensions: Teachers found these data practices beneficial for their practice with students. But this broader data-centric discourse “came to constitute what was valued about their work and learning, and that of their students”.

  (Hardy 2018)
- New methods?
- New sites? Research *on* schools can not be only research *in* schools



GEORG ECKERT INSTITUT
Leibniz-Institut für internationale Schulbuchforschung

# [3] DISCOURSE “ENCODED INTO” EDUCATION

- Critical analyses of edtech from science and technology studies (STS), critical design research, educational theory, governance studies and ethnographic media studies

(Edwards/Fenwick 2015; Hartong 2016; Perrotta/Selwyn 2019; Richter/Allert 2017; Sims 2017; Thompson/Sellar 2018).

- Rare discourse analysis

(social science fiction critique of edtech discourse in Macgilchrist/Allert/Bruch 2020)

With Juliane Jarke

Predictive analytics in education

Using data to predict students risk of dropping out of school/college. Usually ABCs (attendance, behaviour, course performance). Sometime demographics. Increasingly also use of learning management systems or other software.

DATAFIED
Data for and in Education

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

RAHMENPROGRAMM
EMPIRISCHE BILDUNGS-FORSCHUNG



DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation
HELMUT SCHMIDT UNIVERSITÄT Universität der Bundeswehr Hamburg
ifib Institut für Informationsmanagement Bremen GmbH
GEORG ECKERT INSTITUT Leibniz-Institut für internationale Schulbuchforschung

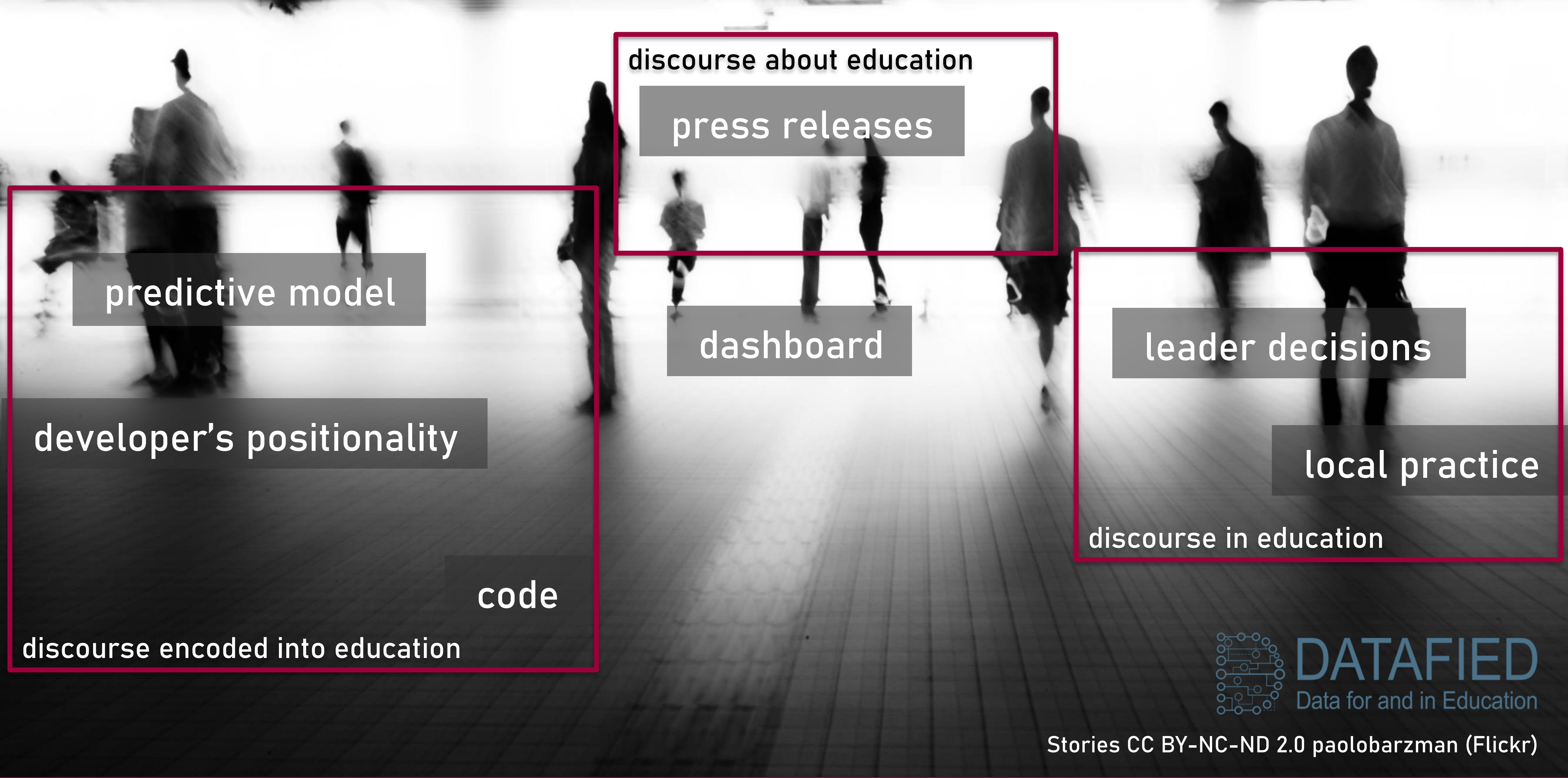
"In simple terms, though it is far from simple, it [entextualization] is the process of rendering discourse extractable, of making a stretch of linguistic production into a unit – a text – that can be lifted out of its interactional setting. [...] Entextualization may well incorporate aspects of context, such that the resultant text carries elements of its history of use within it." (Bauman/Briggs 1990: 73)

DATAFIED
Data for and in Education





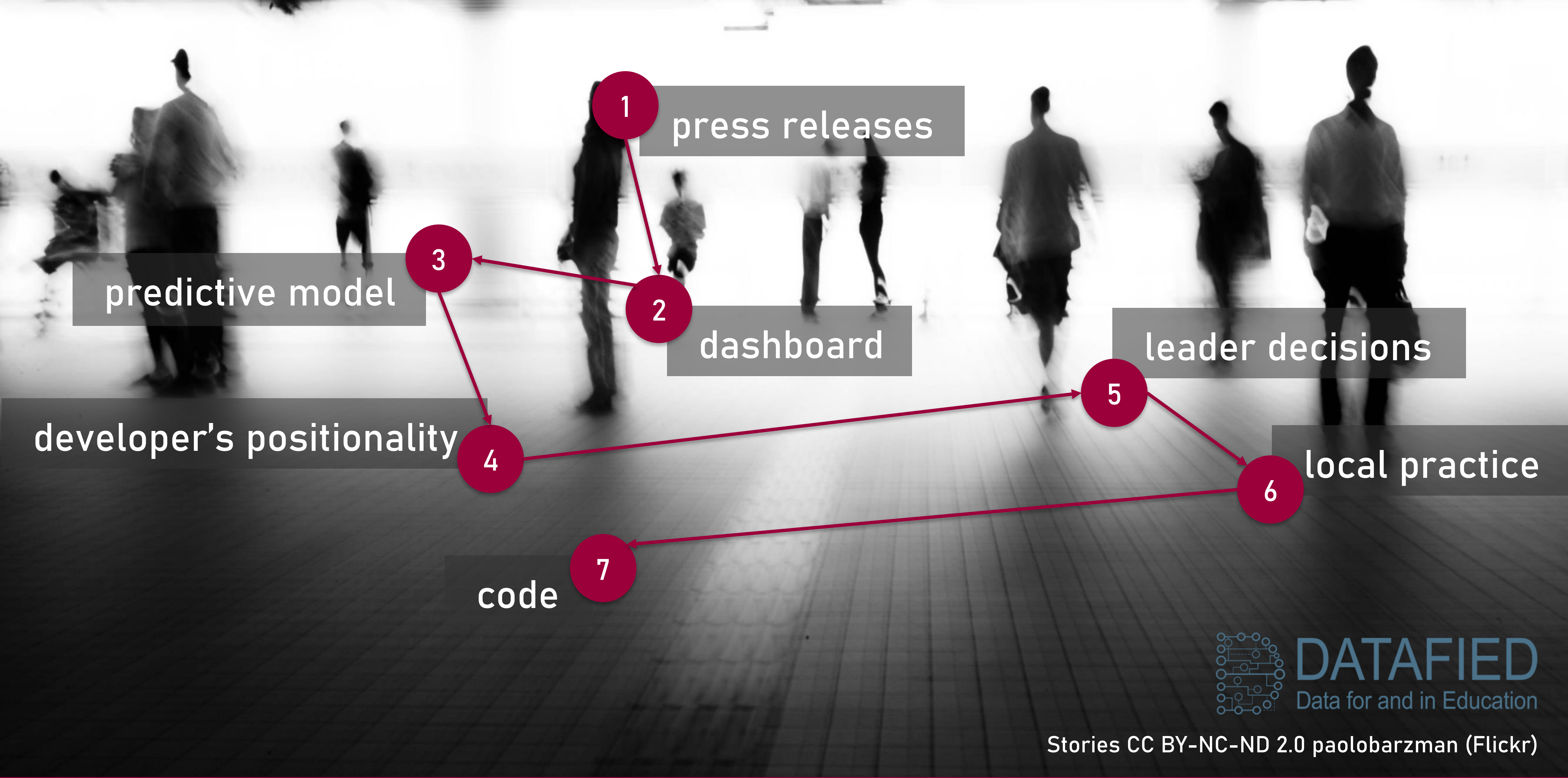
DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation
Helmut Schmidt Universität Universität der Bundeswehr Hamburg
ifib Institut für Informationsmanagement Bremen GmbH
GEORG ECKERT INSTITUT Leibniz-Institut für internationale Schulbuchforschung








DIPF
Leibniz-Institut für Bildungsforschung und Bildungsinformation
Helmut Schmidt Universität
Universität der Bundeswehr Hamburg
ifib
Institut für Informationsmanagement Bremen GmbH
GEORG ECKERT INSTITUT
Leibniz-Institut für internationale Schulbuchforschung




DATAFIED
Data for and in Education





DIPF
Leibniz-Institut für Bildungsforschung und Bildungsinformation

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

ifib
Institut für Informationsmanagement Bremen GmbH

GEORG ECKERT INSTITUT
Leibniz-Institut für internationale Schulbuchforschung

# PRESS RELEASE

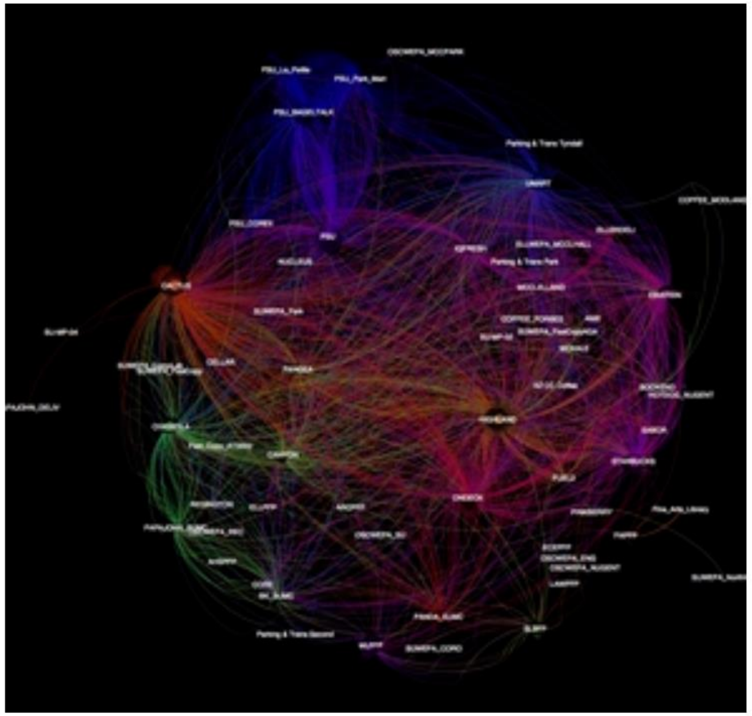
"Sudha Ram, a professor of management information systems, directs the UA's INSITE: Center for Business Intelligence and Analytics in the Eller College of Management. The center focuses on harnessing the power of big data, using machine learning and network science, to help businesses and organizations make better-informed decisions.

The goal of Ram's Smart Campus research is to help educational institutions repurpose the data already being captured from student ID cards to identify those most at risk for not returning after their first year of college.

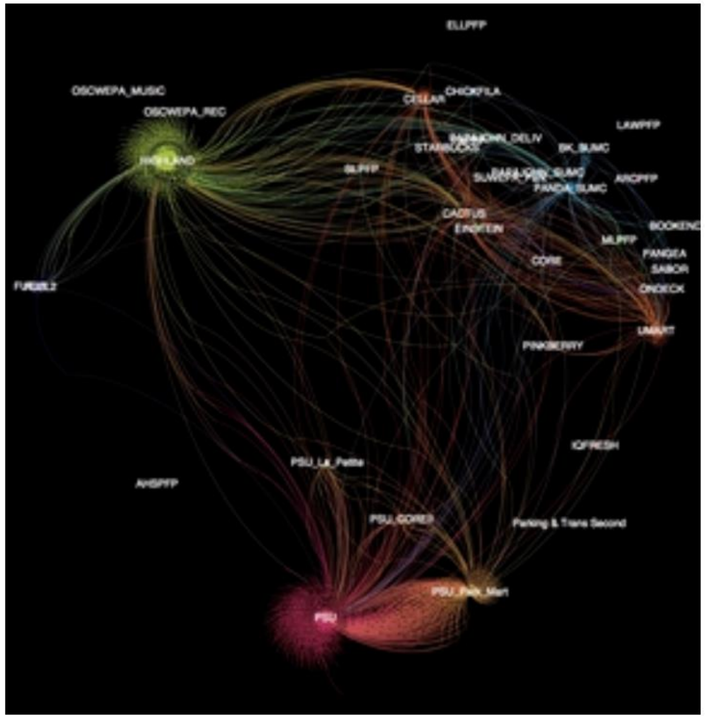
"By getting their digital traces, you can explore their patterns of movement, behavior and interactions, and that tells you a great deal about them," Ram said.

Freshman retention is an ongoing challenge for public universities nationwide. It's important not only for the obvious reason — that a university's goal is to educate students — but also because retention and graduation rates influence a university's reputation and national rankings."

# PRESS RELEASE



This data visualization shows the traffic patterns of UA freshmen, as told by their CatCard usage, in a select area of campus between 10 a.m. and 2 p.m. on a weekday. (Image courtesy of Sudha Ram)



The traffic patterns of UA freshmen, as told by their CatCard usage, in a select area of campus between 6 a.m. and 10 p.m. on a weekday (Image courtesy of Sudha Ram)

“Of all the students who drop out at the end of the first year, with our social integration measures, we're able to do a prediction at the end of the first 12 weeks of the semester with 85 to 90 percent recall," Ram said. "That means out of the 2,000 students who drop out, we're able to identify 1,800 of them.”

# PRESS RELEASE

“Sudha Ram, a professor of management information systems, directs the UA's INSITE: Center for Business Intelligence and Analytics in the Eller College of Management. The center focuses on harnessing the power of big data, using machine learning and network science, to help businesses and organizations make better-informed decisions.

The goal of Ram's Smart Campus research is to help educational institutions repurpose the data already being captured from student ID cards to identify those most at risk for not returning after their first year of college.

“By getting their digital traces, you can explore their patterns of movement, behavior and interactions, and that tells you a great deal about them," Ram said.

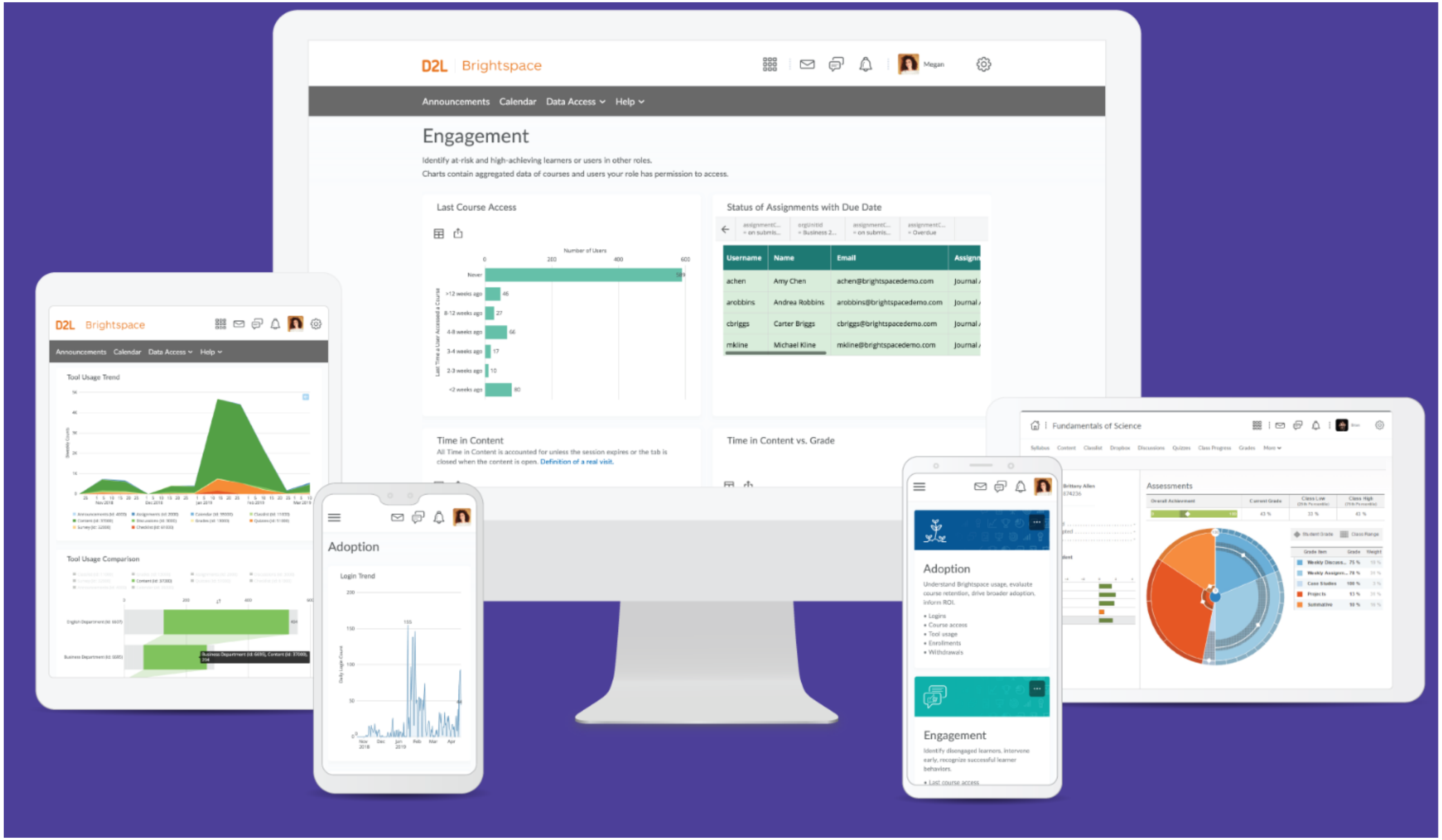
Freshman retention is an ongoing challenge for public universities nationwide. It's important not only for the obvious reason — that a university's goal is to educate students — but also because retention and graduation rates influence a university's reputation and national rankings.”




DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation
Helmut Schmidt Universität Universität der Bundeswehr Hamburg
ifib Institut für Informationsmanagement Bremen GmbH
Georg Eckert Institut Leibniz-Institut für internationale Schulbuchforschung

# DASHBOARDS

D2L Brightspace




## Predict at-risk students.

Issues can be identified as early as two weeks into a course: an embedded widget displays at-risk students directly in an instructor's workflow.

- Simplified, interactive reports make risk patterns easier to see
- "Success Index" shows a student's predicted final grade
- Class list can be filtered to show at-risk or potential risk students

## Identify the risk factors.

A struggling student's content usage, grades, and social engagement can be compared at-a-glance to find the cause faster.

- Interactive win-loss charts show a student's position relative to course expectations
- Individual performance can be viewed across all assessments relative to peers
- Social network within a class is visualized based on discussion forums

## Get students back on track.

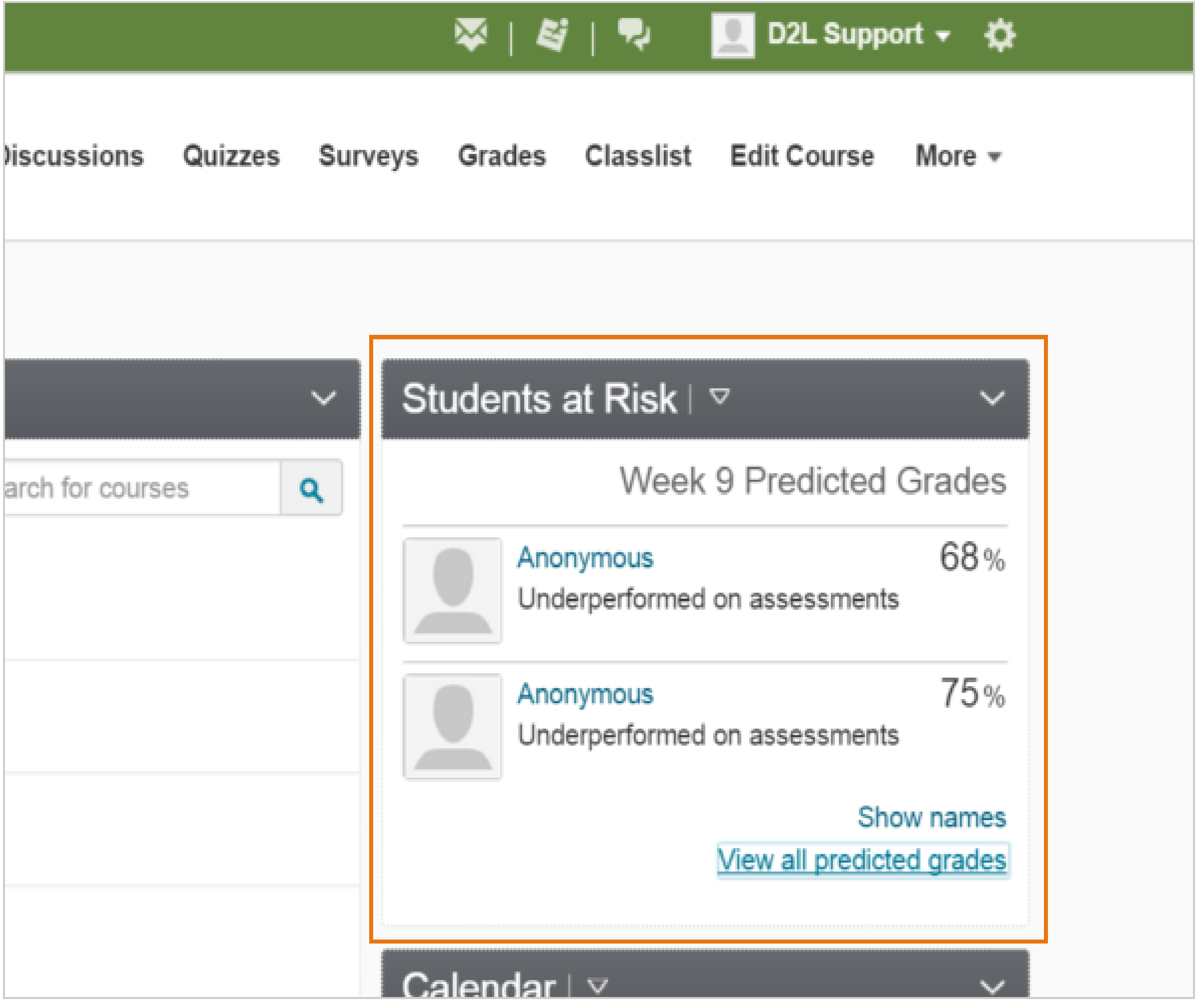
Knowing which content and delivery models are working helps instructors take action quickly with the right mentoring tools.

- Specific risk types are identified: academic, engagement, withdrawal, dropout
- The entire class, subsets, or individuals can be emailed directly from the tool
- Week-over-week trend indicators demonstrate the impact of interventions

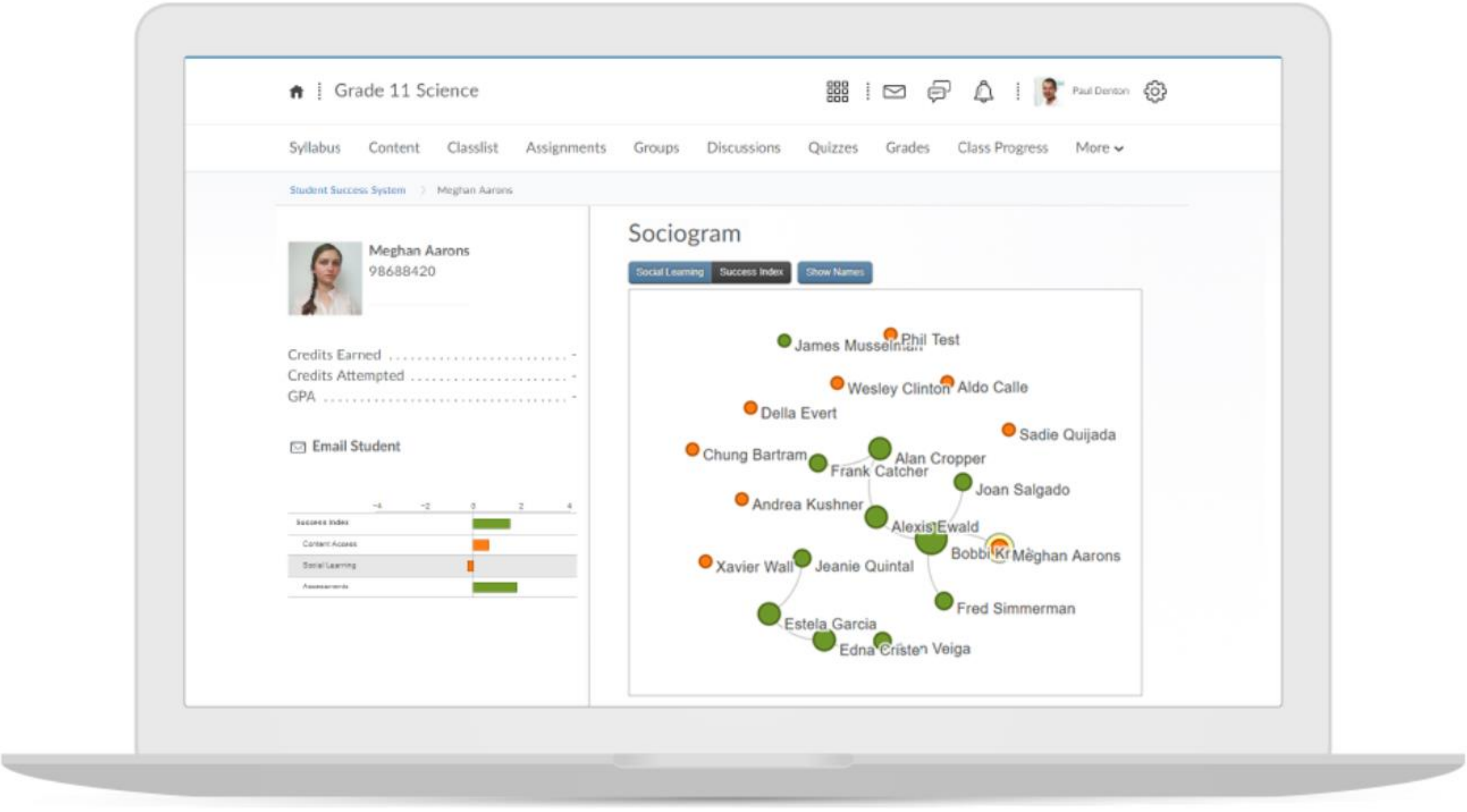
# Using the Students at Risk widget

After you add the Students at Risk widget to a course home page, you can begin using the widget to monitor learner success.

By default, learner names display as **Anonymous** in the widget. This allows instructors to protect learner names in situations when the instructor displays the course home page to other learners, for example, in a classroom or during a web cast. Next to each learner name, a predicted grade for the week displays.




D2L | Brightspace




D2L | Brightspace






DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation

Helmut Schmidt Universität Universität der Bundeswehr Hamburg

ifib Institut für Informationsmanagement Bremen GmbH

GEORG ECKERT INSTITUT Leibniz-Institut für internationale Schulbuchforschung

Engagement dashboard

“These domains are included by default in predictive models, however administrators can exclude domains during configuration. The domains for a course's predictive model display in the win-loss chart on each learner's personal dashboard. These individual domains combine to provide the overall value of the success index for each learner in the course.”

D2L | Brightspace

| Domain | Description |
|---|---|
| Course Access | This domain describes engagement in terms of how often a learner logs in to Brightspace Learning Environment and accesses the course's home page. |
| Content Access | This domain describes engagement in terms of how often a learner accesses Content topics and modules in the course. |
| Social Learning | This domain describes learner engagement in discussion forums within the course. You can click this domain in the win-loss chart to see a detailed visualization of social learning for the learner. |
| Assessments | This domain describes learner performance on assessments. You can click this domain in the win-loss chart to see a detailed visualization of learner performance. |
| Preparedness | This domain consists of a set of data elements from the Student Information System (SIS), including admission scores, overall organization performance, and demographics. This option may not be available for your organization. If you would like more information about getting SIS, contact your D2L Account Manager. |

Data are never raw;
always cooked.
Bias is baked in.
(Gitelman 2013; Noble 2018)

That data are messy is so commonsensical for developers that it gets forgotten in public discourse.

Discourse analysis reflects on the forgotten.

“[M]any factors appear to lead to student dropout, including lack of social support from parents, poor motivation, low self-esteem, parental educational achievement and value, and economic factors, making it difficult to create a single intervention that works for all students. While demographic factors correlate with eventual dropout, these indicators are not considered actionable. A school district generally does not have the capacity to improve a student’s economic condition, nor is it possible to alter a student's racial identity or gender. As such, the educational research community has focused on more actionable factors such as behavior, attendance, engagement, and social-emotional learning.”

(Coleman/Baker/Stephenson 2019)

# …entextualized into

- General Coursework: student academic performance such as total credits earned or student grade point averages.
- Student Assessments: interim or summative assessments related to math, science, reading and social studies performance.
- Student Attendance: recorded as absences or tardies.
- Student Behavior: disciplinary incidents the student has on file.

(Coleman/Baker/Stephenson 2019)

# …entextualized into

- “For each of these models, we took the data from a single district. We filtered down to only the students who were flagged as ‘dropped’ and ‘graduated’. Even if students took extra time to graduate, they were still counted as graduating. Only these students were used for building the model; all other outcomes (such as transferring to another school district) were removed from the filtered dataset.”
- “The resultant datasets were generally highly imbalanced, with substantially more students graduating than dropping out. To account for this imbalance, the training data was manually rebalanced by adding duplicate copies of students who dropped out to the data set. Specifically, duplicates were created such that every grade level (10th, 11th, 12th, etc.) of students in the training datasets had an equal number of students who dropped out as students who remained.”
- “Since the algorithm was tree based, we utilized arbitrary value substitution to replace missing values with a high integer”. (Coleman/Baker/Stephenson 2019)

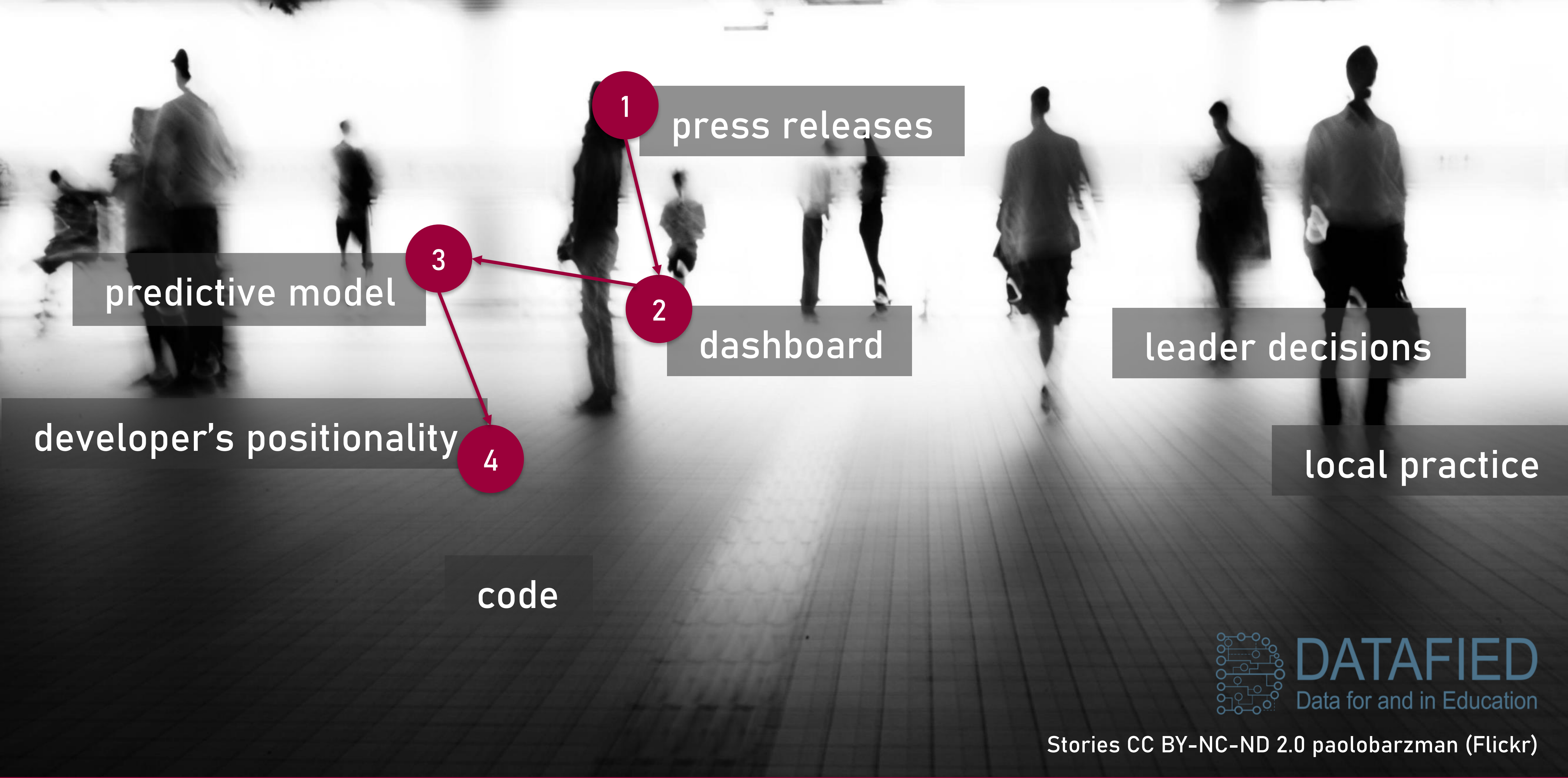
# …and

- Districts were not included as Pillar districts if they had high amounts of missing data, over 40% of values missing" […]

# …was entextualized as

- "Having developed models for Pillar districts, where data are abundant, data quality is high, and where it is possible to develop a high-quality model, we next applied each Pillar model to each Target district."

(Coleman/Baker/Stephenson 2019)








DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation

Helmut Schmidt Universität Universität der Bundeswehr Hamburg

ifib Institut für Informationsmanagement Bremen GmbH

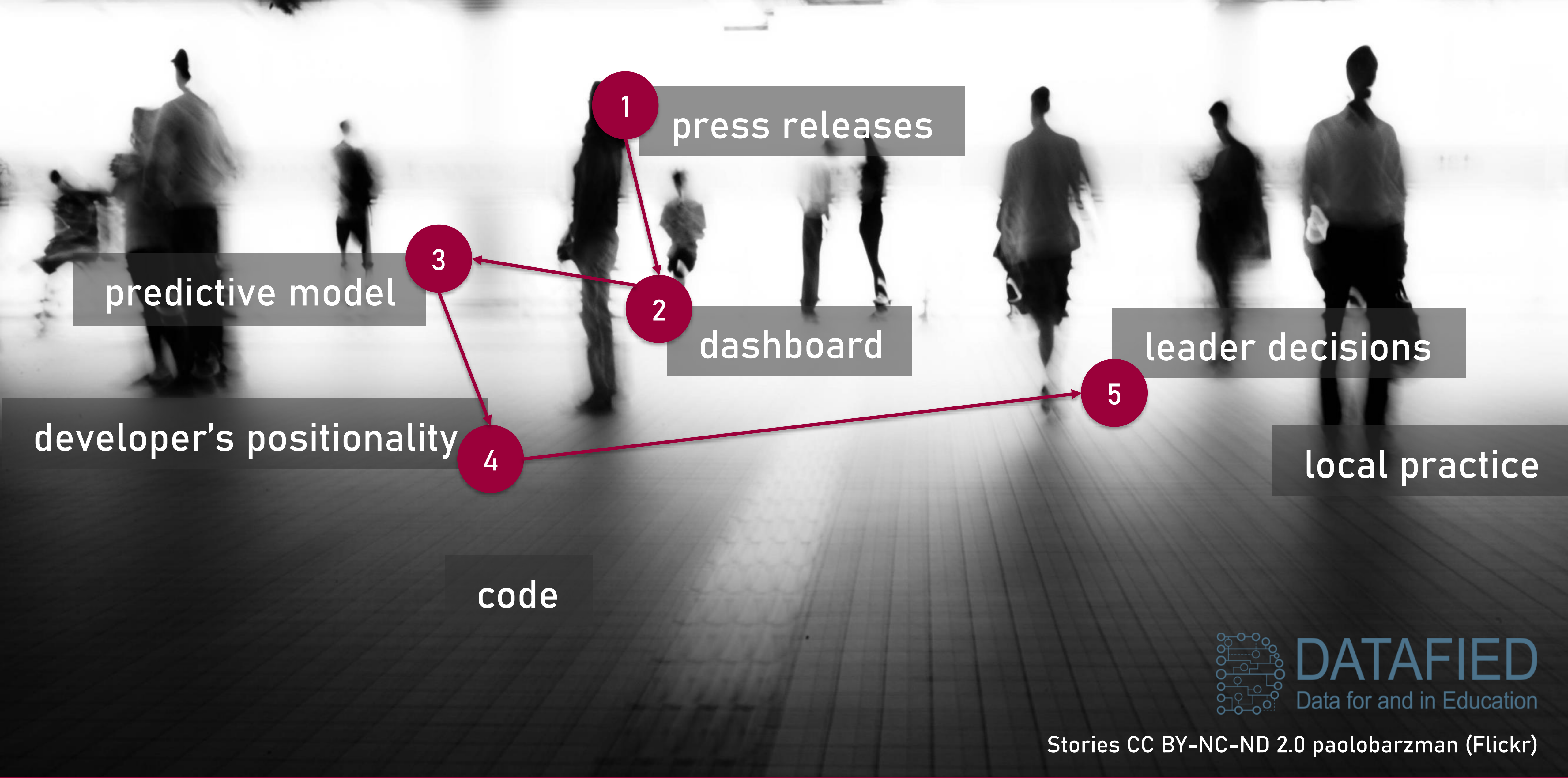
GEORG ECKERT INSTITUT Leibniz-Institut für internationale Schulbuchforschung

"I think, more recently, there has been much more of an understanding of how data sets can encode existing biases, and that if you just take a data set as a neutral thing, then you're making a big mistake. You're not understanding the nature of how data are collected, and what data are collected, and why data are collected."

(Interview 10, 2019; cited in Johanes/Thille 2019)

“if we know the student’s gonna fail, is it ethical that we accept him in the first place?”

(Interview 3, 2019; cited in Johanes/Thille 2019)





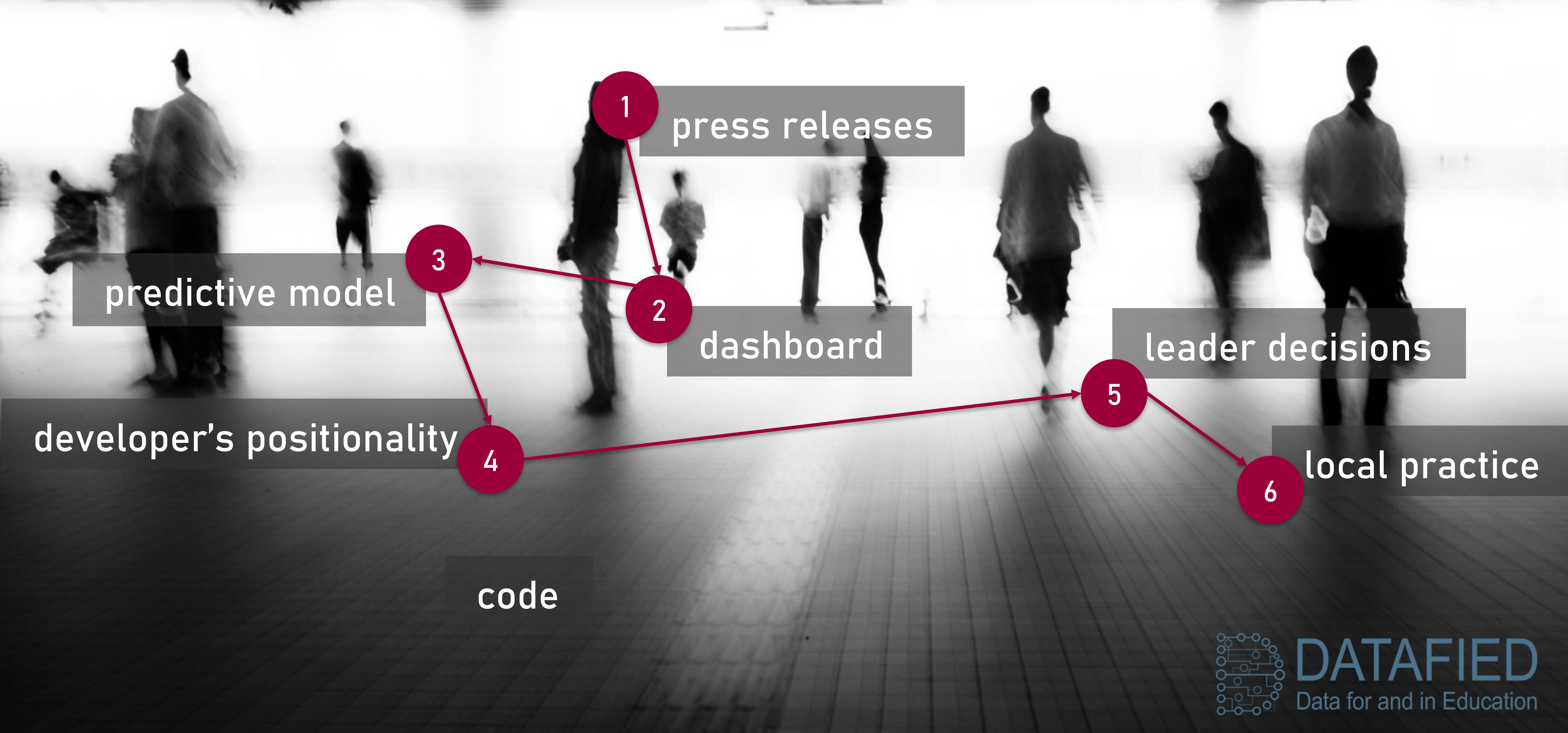



DIPF
Leibniz-Institut für Bildungsforschung und Bildungsinformation
Helmut Schmidt Universität
Universität der Bundeswehr Hamburg
ifib
Institut für Informationsmanagement Bremen GmbH
GEORG ECKERT INSTITUT
Leibniz-Institut für internationale Schulbuchforschung

# LEADERSHIP DECISIONS

Mount St. Mary University, USA. President wanted to use predictive analytics to "weed out students unlikely to be retained". Predictions embedded into rankings and economic logics.

(Jones & McCoy 2019)




DATAFIED
Data for and in Education





DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation
Helmut Schmidt Universität Universität der Bundeswehr Hamburg
ifib Institut für Informationsmanagement Bremen GmbH
GEORG ECKERT INSTITUT Leibniz-Institut für internationale Schulbuchforschung

# LOCAL PRACTICE

School's weekly data meeting. Student getting more Ds and Fs than in the first half of the year; hunger, homeless services. Predictions embedded into relations of care and community.

(NPR 2019)

Child protection service. Professionals make judgement about whether child should be removed from family. System makes different judgement. Professionals defer to system. Predictions embedded into relations of deference.

(Eubanks 2018)

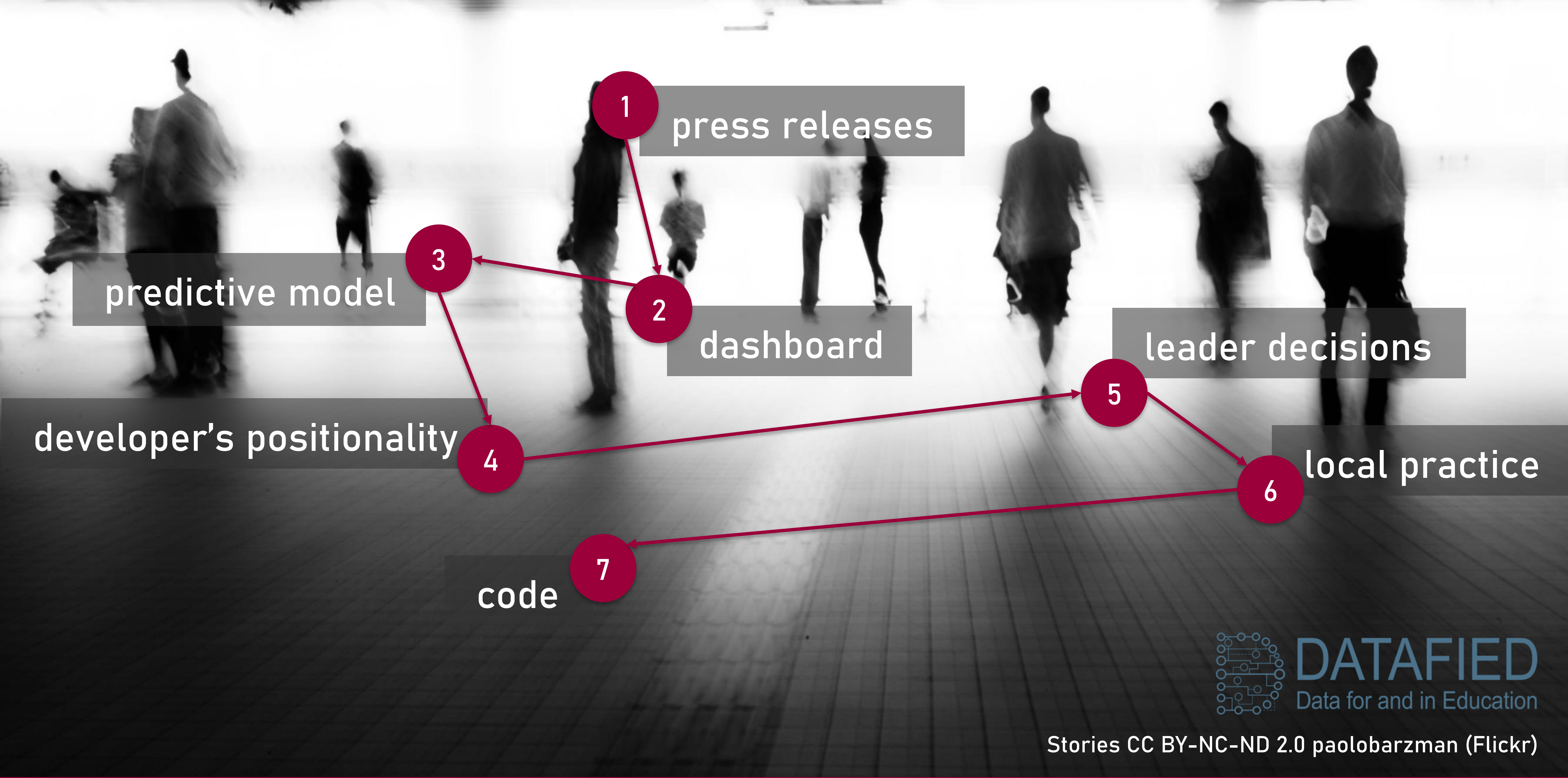

1 press releases
2 dashboard
3 predictive model
4 developer's positionality
5 leader decisions
6 local practice
7 code
DATAFIED
Data for and in Education

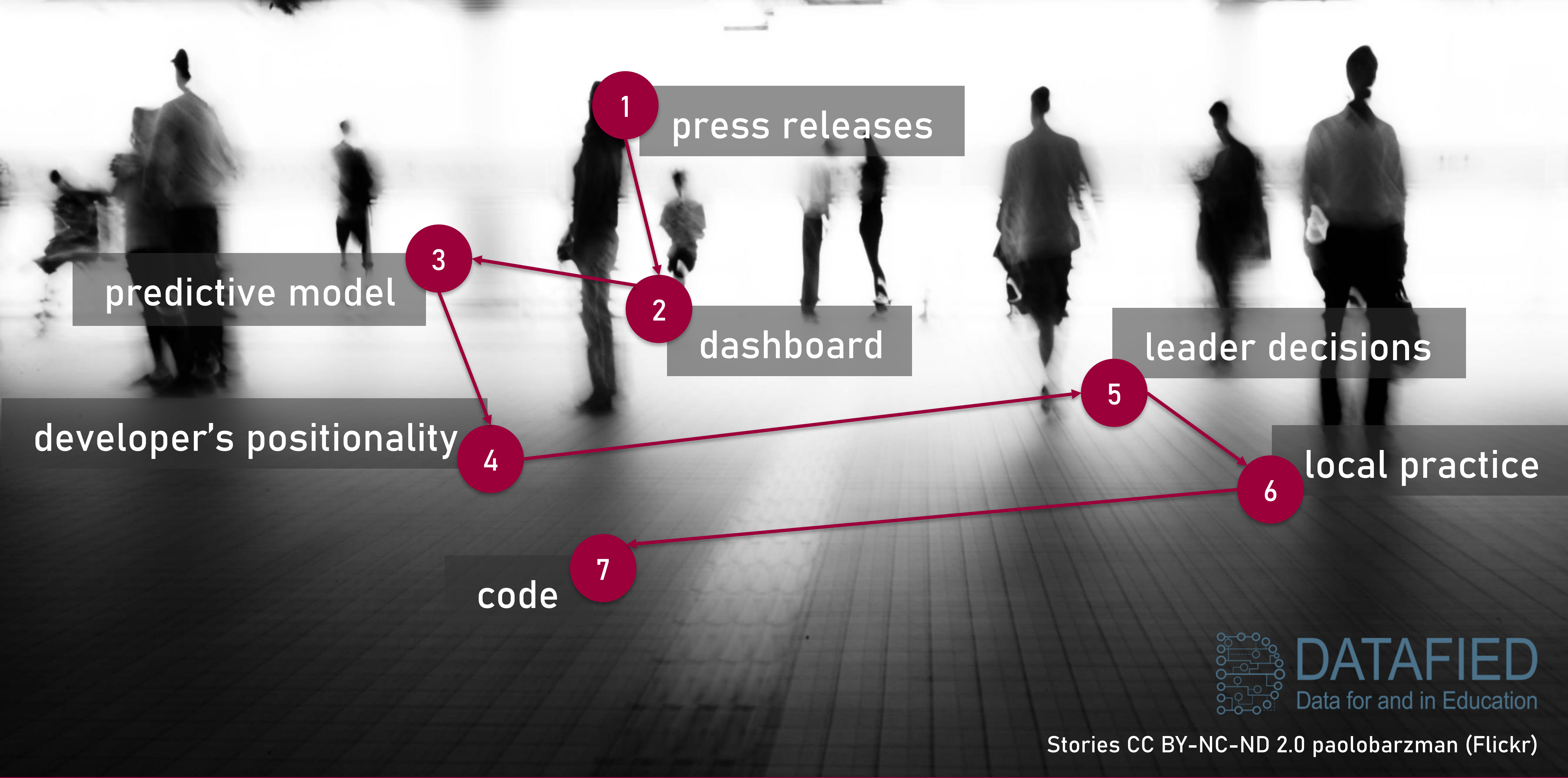

Image found on Mark C. Marino's Twitter @markcmarino

```
function anagram(text) {
  var a = text.split("");
  for (var i = 0; i < a.length; i +=
1) {
    var letter = a[i];
    var j = Math.floor(Math.random() *
a.length);
    a[i] = a[j];
    a[j] = letter;
  }
  return a.join("");
}
```

```
function anagram(text) {
  return   text.split("").sort(function
()     {return     0.5-Math.random()}).
join("");
}
```

"Data in Discourse Analysis"
=> Tau sandal idiosyncrasies

(Marino 2020)







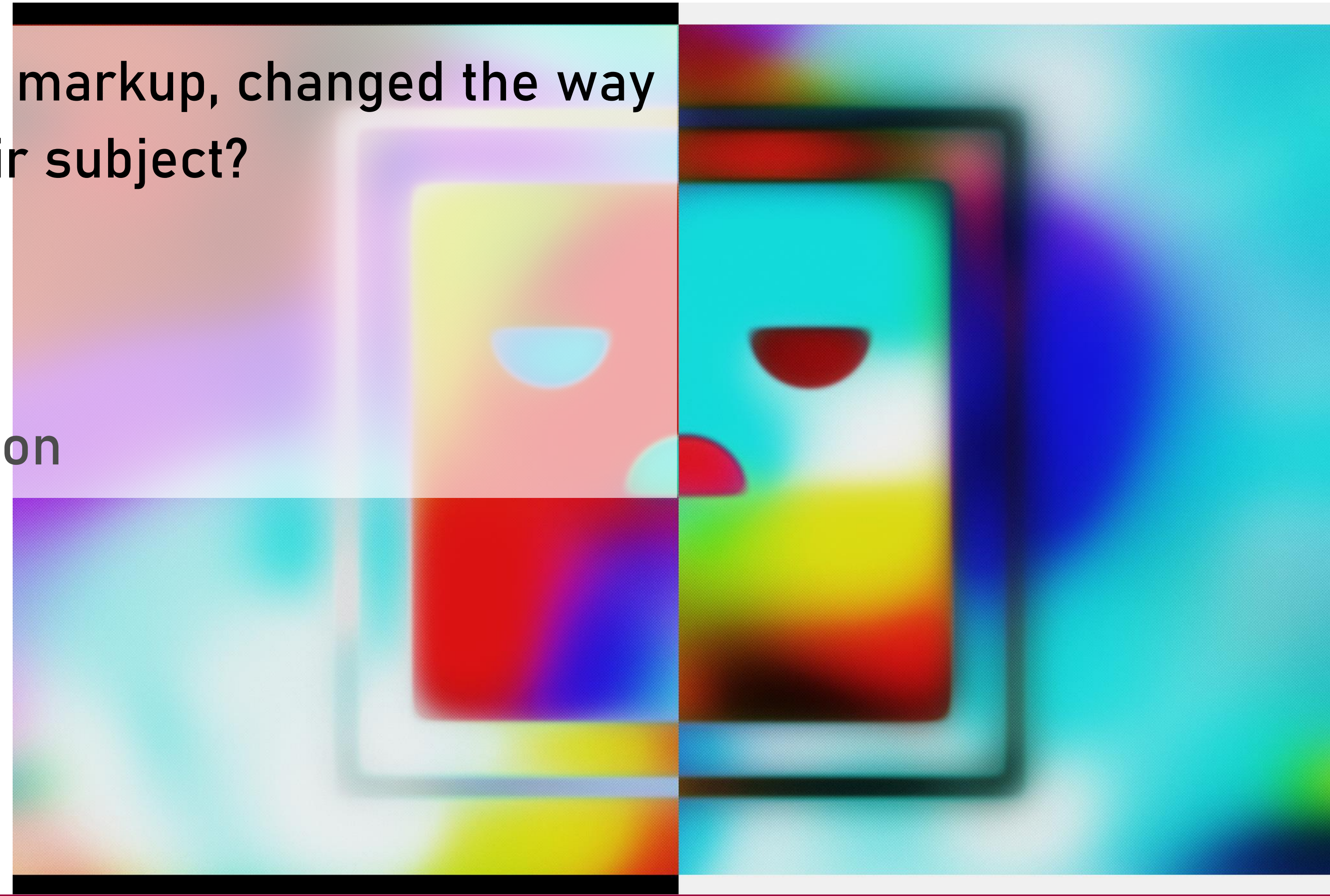

DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation
Helmut Schmidt Universität Universität der Bundeswehr Hamburg
ifib Institut für Informationsmanagement Bremen GmbH
Georg Eckert Institut Leibniz-Institut für internationale Schulbuchforschung

## How has data, its encoding and markup, changed the way discourse studies perceive their subject?

discourse **about** education,

discourse **in** education, and

discourse **encoded into** education

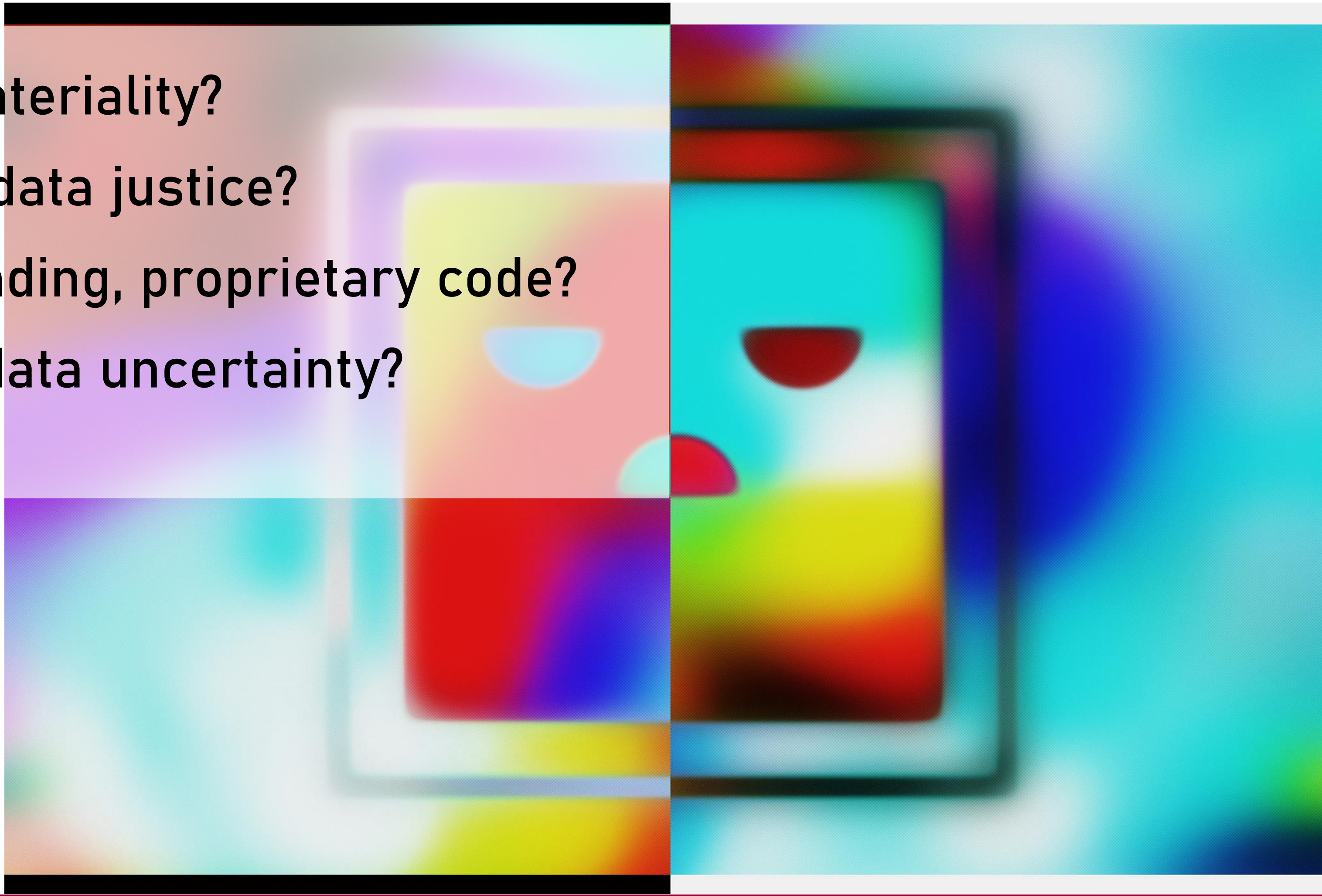


Epistemologically: texts and materiality?

Politically: generative critique, data justice?

Practically: technical understanding, proprietary code?

Pedagogically: coding, design, data uncertainty?



Thank you!
@discoursology
macgilchrist@leibniz-gei.de

Ball, S.J./Junemann, C./Santori, D. (2017): *Edu.net. Globalisation and Education Policy Mobility,* London: Routledge.

Bauman, R./Briggs, C.L. (1990): "Poetics and Performance as Critical Perspectives on Language and Social Life", in: *Annual Review of Anthropology* 19 S. 59-88.

Coleman, C./Baker, R.S./Stephenson, S. (Year) A Better Cold-Start for Early Prediction of Student At-Risk Status in New School Districts. In: International Educational Data Mining Society, ed. Proceedings of the 12th International Conference on Educational Data Mining (EDM 2019), 2019 Montreal, Canada. pp. 732-737.

Edwards, R./Fenwick, T. (2015): "Digital analytics in professional work and learning", in: *Studies in Continuing Education* 38(2), S. 213-227.

Gitelman, L. (Hrsg.) (2013): *"Raw Data" Is an Oxymoron,* Cambridge, MA: MIT Press.

Hardy, I. (2018): "Governing teachers' work and learning through data: Australian insights", in: *Cambridge Journal of Education* 49(4), S. 501-517.

Hartong, S. (2016): "Between assessments, digital technologies and big data: The growing influence of 'hidden' data mediators in education", in: *European Educational Research Journal* 15(5), S. 523-536.

Johanes, P./Thille, C. (2019): "The heart of educational data infrastructures = Conscious humanity and scientific responsibility, not infinite data and limitless experimentation", in: *British Journal of Educational Technology* 50(6), S. 2959-2973.

Macgilchrist, F./Allert, H./Bruch, A. (2020): "Students and society in the 2020s: Three future 'histories' of education and technology", in: *Learning, Media and Technology* 45(1), S. 76-89.

Marino, M.C. (2020): *Critical Code Studies,* Cambridge, MA: MIT Press.

Perrotta, C./Selwyn, N. (2019): "Deep learning goes to school: toward a relational understanding of AI in education", in: *Learning, Media and Technology* 10.1080/17439884.2020.1686017 S. 1-19.

Richter, C./Allert, H. (2017): "Design as critical engagement in and for education", in: *EDeR - Educational Design Research* 1(1), S. 1-20.

Roberts-Holmes, G./Bradbury, A. (2016): "The datafication of early years education and its impact upon pedagogy", in: *Improving Schools* 19(2), S. 119–128.

Roberts-Mahoney, H./Means, A.J./Garrison, M.J. (2016): "Netflixing human capital development: personalized learning technology and the corporatization of K-12 education", in: *Journal of Education Policy* 31(4), S. 405-420.

Rowe, E. (2019): "Capitalism without capital: the intangible economy of education reform", in: *Discourse: Studies in the Cultural Politics of Education* 40(2), S. 271-279.

Sims, C. (2017): *Disruptive Fixation: School Reform and the Pitfalls of Techno-Idealism,* Princeton, NJ: Princeton University Press.

Thompson, G./Sellar, S. (2018): "Datafication, testing events and the outside of thought", in: *Learning, Media and Technology* 10.1080/17439884.2018.1444637 S. 1-13.

Williamson, B. (2017): "Decoding ClassDojo: psycho-policy, social-emotional learning and persuasive educational technologies", in: *Learning, Media and Technology* 42(4), S. 440-453.

DATAFIED: https://datafied.de