Big Education in the Era of Big Data

Irwin King

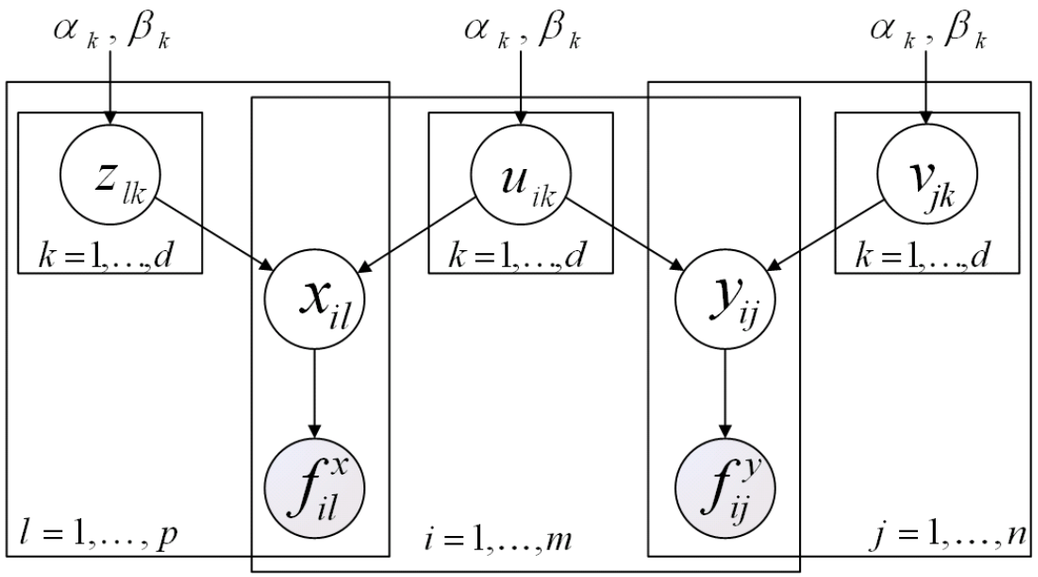
Associate Dean (Education), Faculty of Engineering, CUHK
Director, Shenzhen Rich Media and Big Data Analytics and Application Key Lab, SZRI, CUHK
PI, The Knowledge and Education Exchange Platform (KEEP), CUHK

Department of Computer Science and Engineering
The Chinese University of Hong Kong

king@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~king

©2015 Irwin King. All rights reserved.

Collective Probabilistic Factor Model



$$\mathcal{L}(U, V, Z; F^x, F^y)$$
$$= \sum_{i=1}^{m}\sum_{l=1}^{p}(f^x_{il}\ln x_{il} - x_{il}) + \sum_{i=1}^{m}\sum_{j=1}^{n}(f^y_{ij}\ln y_{ij} - y_{ij})$$
$$+ \sum_{i=1}^{m}\sum_{k=1}^{d}((\alpha_k - 1)\ln(u_{ik}/\beta_k) - u_{ik}/\beta_k)$$
$$+ \sum_{j=1}^{n}\sum_{k=1}^{d}((\alpha_k - 1)\ln(v_{jk}/\beta_k) - v_{jk}/\beta_k)$$
$$+ \sum_{l=1}^{p}\sum_{k=1}^{d}((\alpha_k - 1)\ln(z_{lk}/\beta_k) - z_{lk}/\beta_k) + \text{const.}$$

$$u_{ik} \leftarrow u_{ik}\frac{\sum_{j=1}^{n}(f^y_{ij}v_{jk}/y_{ij}) + \sum_{l=1}^{p}(f^x_{il}z_{lk}/x_{il}) + (\alpha_k - 1)/u_{ik}}{\sum_{j=1}^{n}v_{jk} + \sum_{l=1}^{p}z_{lk} + 1/\beta_k}$$

$$v_{jk} \leftarrow v_{jk}\frac{\sum_{i=1}^{m}(f^y_{ij}u_{ik}/y_{ij}) + (\alpha_k - 1)/v_{jk}}{\sum_{i=1}^{m}u_{ik} + 1/\beta_k},$$

$$z_{lk} \leftarrow z_{lk}\frac{\sum_{i=1}^{m}(f^x_{il}u_{ik}/x_{il}) + (\alpha_k - 1)/z_{lk}}{\sum_{i=1}^{m}u_{ik} + 1/\beta_k}.$$

$$u_{ik} \leftarrow u_{ik}\frac{\theta\sum_{j=1}^{n}(f^y_{ij}v_{jk}/y_{ij}) + (1-\theta)\sum_{l=1}^{p}(f^x_{il}z_{lk}/x_{il}) + (\alpha_k - 1)/u_{ik}}{\theta\sum_{j=1}^{n}v_{jk} + (1-\theta)\sum_{l=1}^{p}z_{lk} + 1/\beta_k}$$

On-Going Research

Machine Learning

- Combinatorial Pure Exploration of Multi-Armed Bandits. (NIPS'14)
- Exact and Stable Recovery of Pairwise Interaction Tensors. (NIPS'13)
- Simple and Efficient Multiple Kernel Learning By Group Lasso (ICML'10)
- Online Learning for Group Lasso (ICML'10)
- Heavy-Tailed Symmetric Stochastic Neighbor Embedding (NIPS'09)
- Adaptive Regularization for Transductive Support Vector Machine (NIPS'09)
- Learning with Consistency between Inductive Functions and Kernels (NIPS'08)
- An Extended Level Method for Efficient Multiple Kernel Learning (NIPS'08)
- Transductive Support Vector Machine (NIPS'07)

The grass is greener on the other side...

Be inspired!

Stories and more stories...

Be informed!

The devil is in the details...

Be challenged!

Words of Wisdom

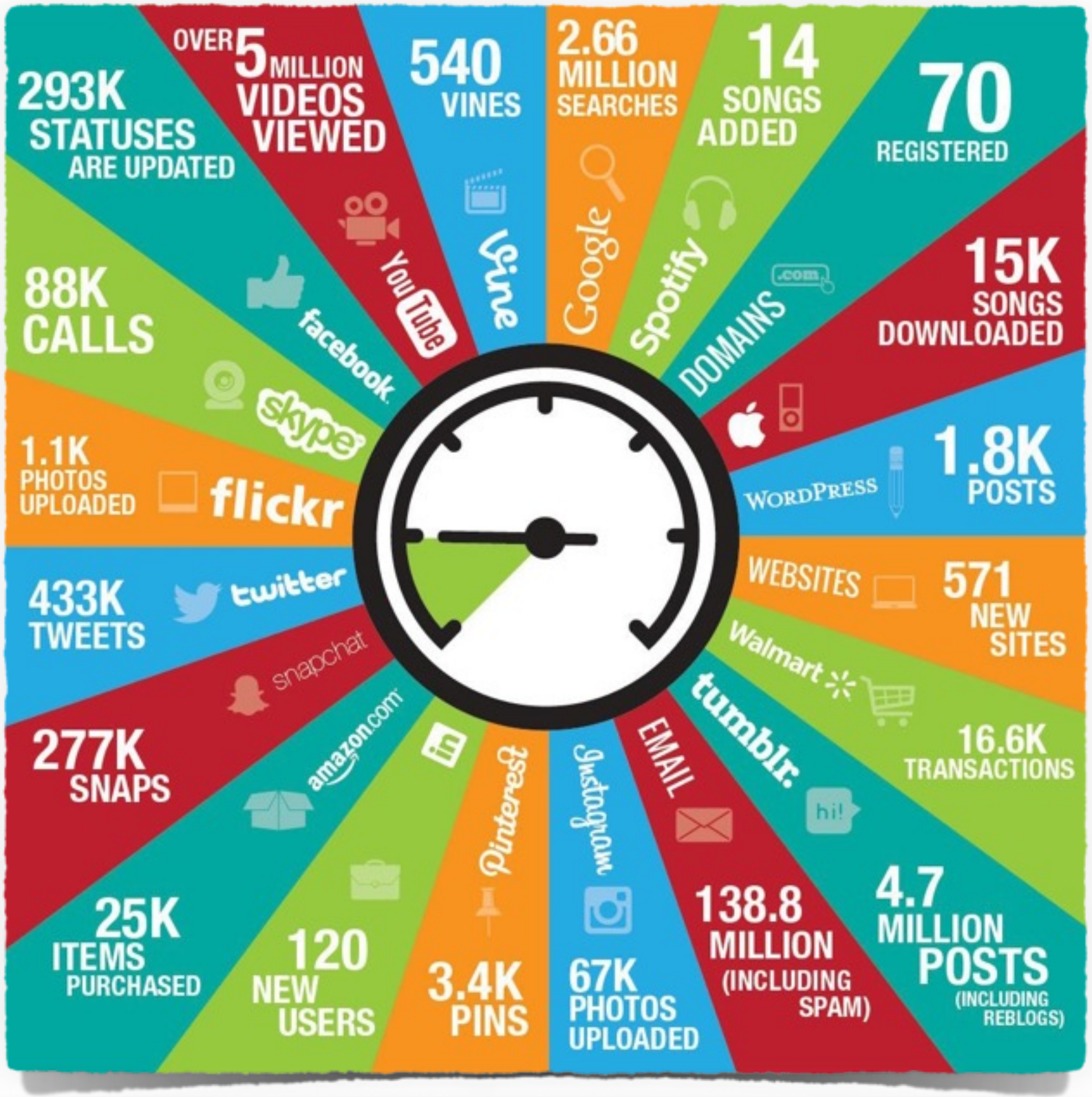
The **BEST** universities focus on EDUCATION!

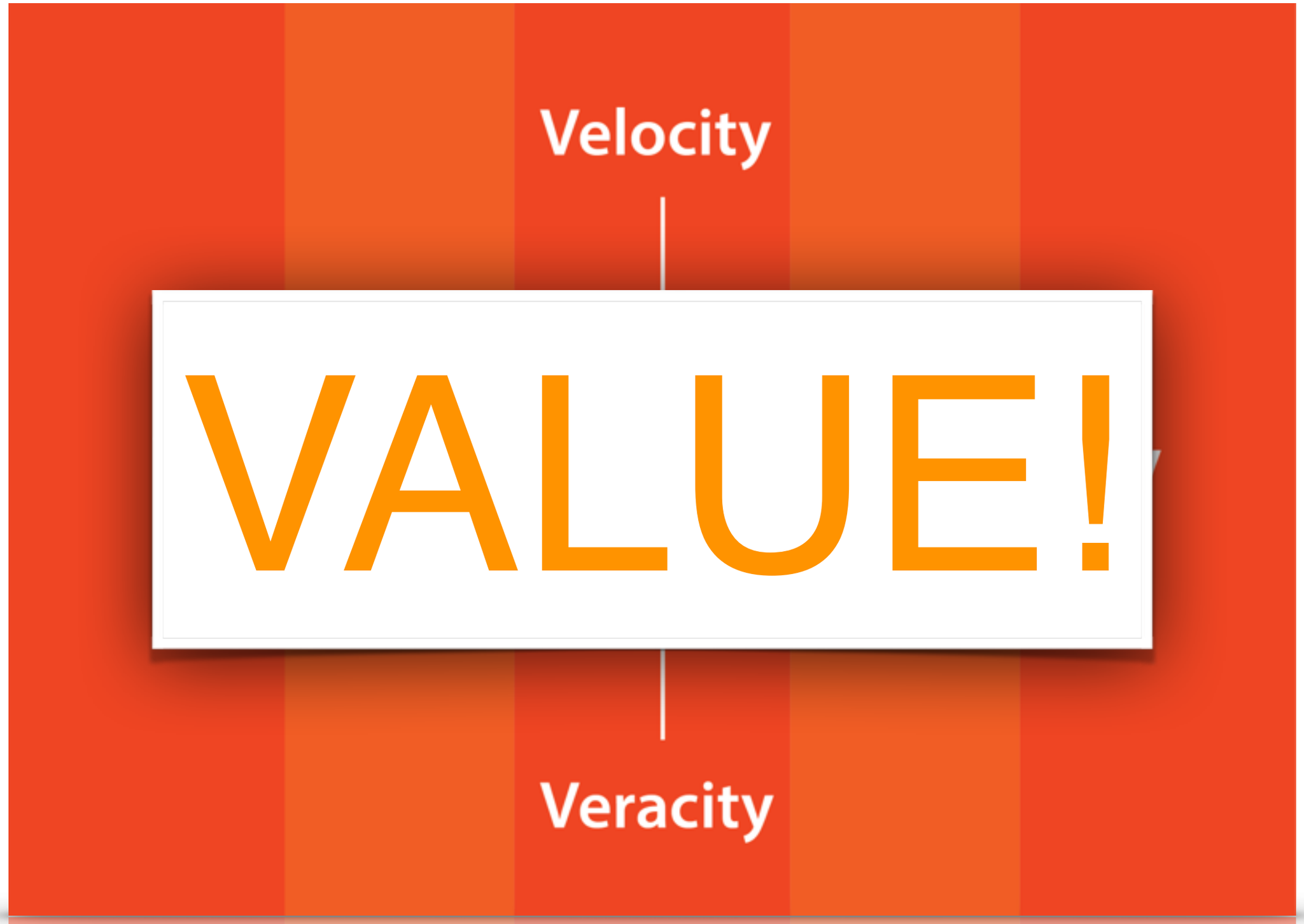
The **BETTER** universities focus on citation numbers and impact factors...

The **GOOD** universities focus on counting the number of publications...

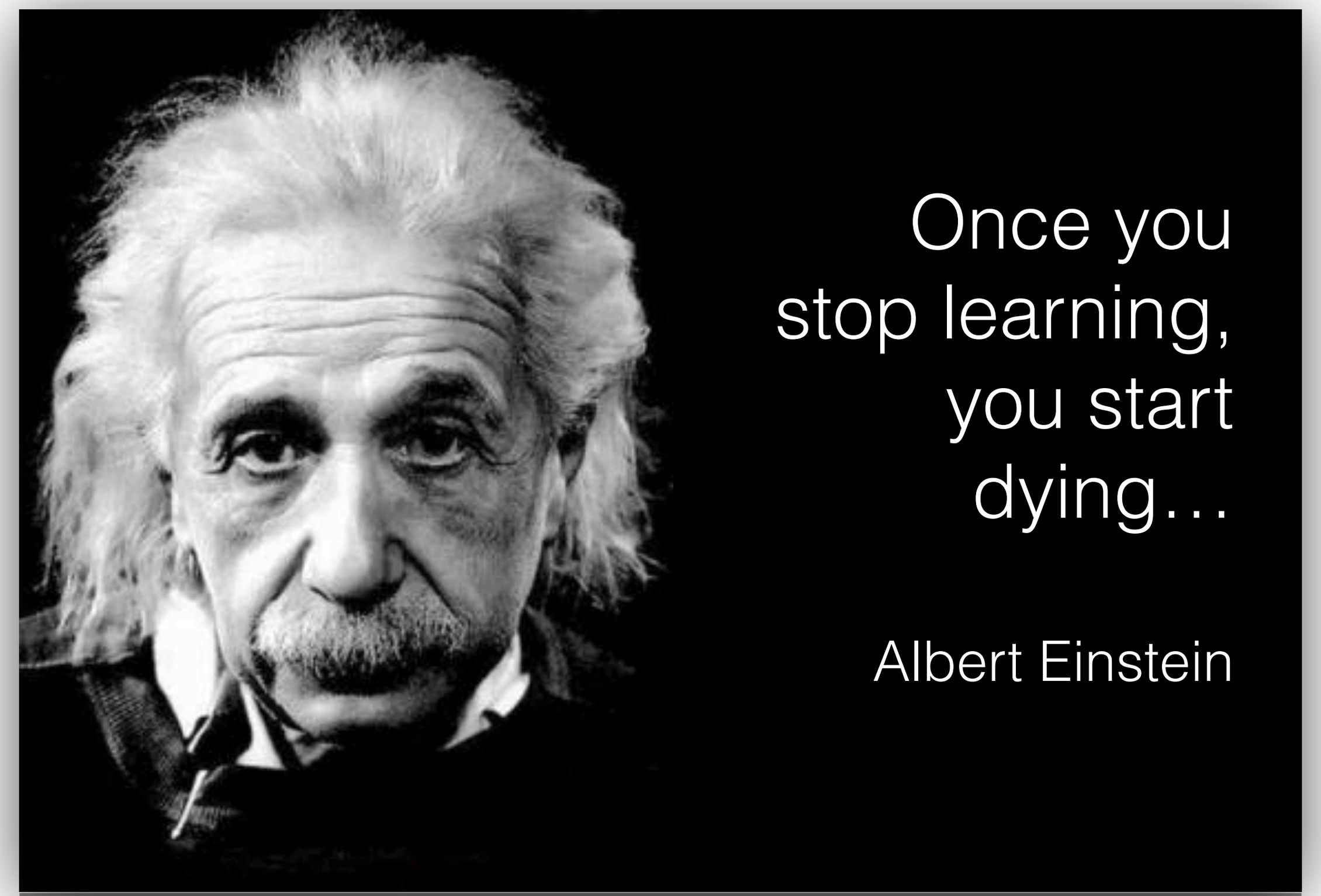
Our Education System

Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid.


293K
STATUSES
ARE UPDATED
OVER 5 MILLION
VIDEOS
VIEWED
540
VINES
2.66
MILLION
SEARCHES
14
SONGS
ADDED
70
REGISTERED
88K
CALLS
facebook
YouTube
Vine
Google
Spotify
DOMAINS
.com
15K
SONGS
DOWNLOADED
skype
1.1K
PHOTOS
UPLOADED
flickr
WordPress
1.8K
POSTS
433K
TWEETS
twitter
WEBSITES
571
NEW
SITES
snapchat
Walmart
277K
SNAPS
amazon.com
16.6K
TRANSACTIONS
tumblr.
hi!
25K
ITEMS
PURCHASED
120
NEW
USERS
Pinterest
3.4K
PINS
Instagram
67K
PHOTOS
UPLOADED
EMAIL
138.8
MILLION
(INCLUDING
SPAM)
4.7
MILLION
POSTS
(INCLUDING
REBLOGS)


Velocity
VALUE!
Veracity

Big Education on Lifelong Learning


Once you
stop learning,
you start
dying…
Albert Einstein

E-LEARNING
MANAGEMENT
CONTENT
COMPUTER
LEARNING
STUDENT
WEB
TEACHING
CBTS
EXAMPLE
ELEARNING
DELIVERED
WELL
OBJECTS
ACTIVITIES
KNOWLEDGE
USED
TERM
MORE
SOFTWARE
SOCIAL
SCHOOLS
HIGHER
SUPPORT
COURSE
VIA
FORM
COST
CLASSES
SECONDARY
ALLOW
EARLY
CLASS
COLLABORATIVE
INCLUDE
COURSES
PARTICIPANTS
DISCUSSION
ONE
EVEN
TEACHING
TIME
EDUCATION
FACE
COMPUTERS
VIRTUAL
WORK
ONLINE
TECHNOLOGIES
TRAINING
REFER
ASSESSMENT
STANDARDS
TECHNOLOGY
USING
BASED
THROUGH
INTERNET
SYSTEMS
CLASSROOM
ENVIRONMENTS
INFORMATION
PROVIDE
ELECTRONIC
TRADITIONAL
SYSTEM
LMS
GENERALLY
MUCH
EDUCATIONAL
PROCESS
ENHANCED

MILESTONES IN E-LEARNING

1971

The Open University opens in England with an open admissions policy, and begins broadcasting lectures on television. **25,000 students enroll.**

1989

University of Phoenix launches its private, for-profit online school **12 students enroll.**

1993

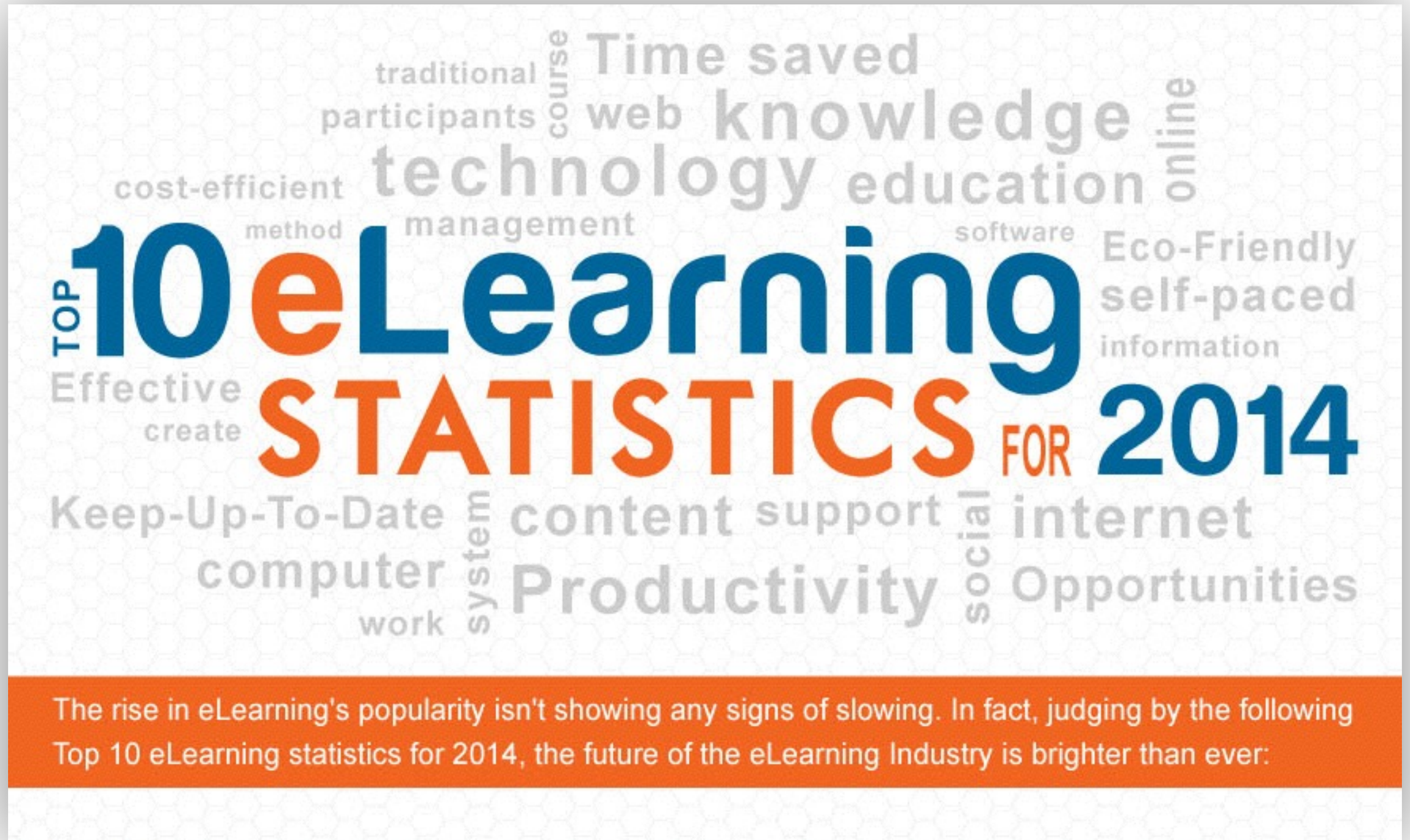
Criteria is created by pioneer William Graziadei III, Ph.D: e-learning systems must be easy to use, portable, replicable, scalable, and affordable.

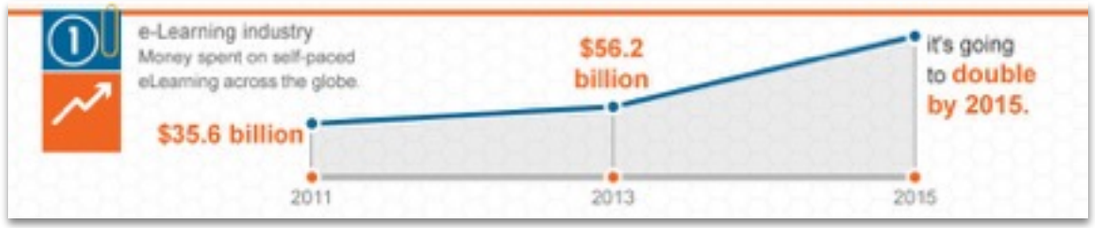
1999

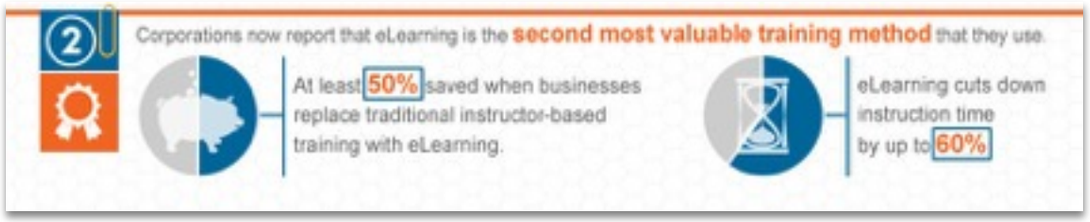
The term **'e-Learning' is coined** at an educational seminar.

2004

Salman Khan records instructional YouTube videos to help his cousins with math. **The rising popularity of these videos leads him to found the Khan Academy,** a not-for-profit, free, online educational organization.


traditional
course
Time saved
participants
web
knowledge
online
technology
education
cost-efficient
method
management
software
Eco-Friendly
TOP
10 eLearning
self-paced
information
Effective
create
STATISTICS FOR 2014
Keep-Up-To-Date
system
content
support
social
internet
computer
Productivity
Opportunities
work
The rise in eLearning's popularity isn't showing any signs of slowing. In fact, judging by the following Top 10 eLearning statistics for 2014, the future of the eLearning Industry is brighter than ever:


1
e-Learning industry
Money spent on self-paced eLearning across the globe.
$35.6 billion
$56.2 billion
it's going to double by 2015.
2011
2013
2015

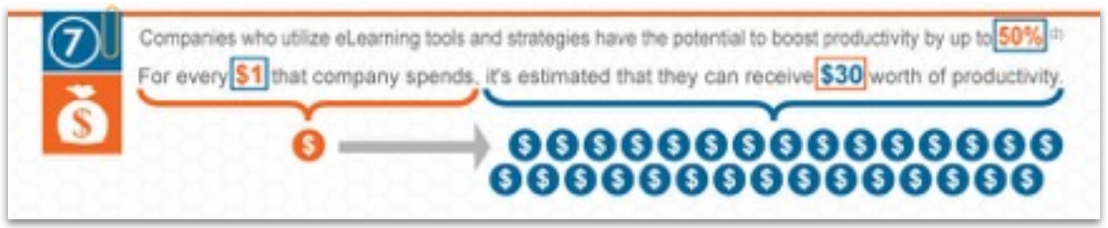

2
Corporations now report that eLearning is the second most valuable training method that they use.
At least 50% saved when businesses replace traditional instructor-based training with eLearning.
eLearning cuts down instruction time by up to 60%

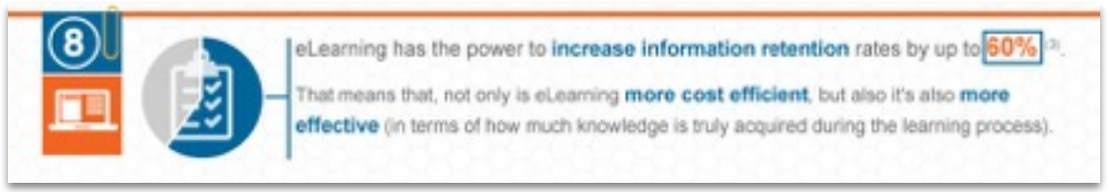
3
2013: About 4.6 out of ten college students* are taking at least one course online.
2019: Roughly half of all college classes* will be eLearning-based.

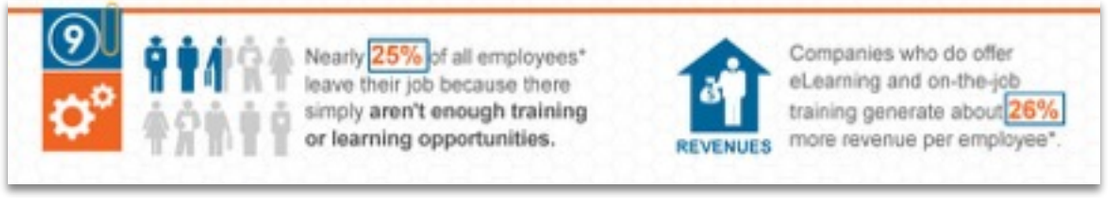
4
eLearning is Eco-Friendly [1]
ENERGY
Consumes 90% less energy than traditional education courses.
The amount of CO2 emissions (per student) is also reduced by up to 85%.
EMISSIONS
CO2

5
Over 41.7% of global Fortune 500 companies now use some form of technology to instruct employees during formal learning hours, and this is going to steadily increase in future years.
USAGE

6
Self-paced eLearning's estimated 5 year annual growth rate*
Western Europe: 5.8%
Revenues by 2015:
$8.1 billion
The world's most rapidly growing eLearning markets are:
Malaysia and Vietnam.
Asia: 17.3%
Middle East: 8.2%
Revenues by 2016:
$560.7 million
Africa: 15.4%
Revenues by 2016:
$512.8 million


7
Companies who utilize eLearning tools and strategies have the potential to boost productivity by up to 50% [2]
For every $1 that company spends, it's estimated that they can receive $30 worth of productivity.


8
eLearning has the power to increase information retention rates by up to 60% [3].
That means that, not only is eLearning more cost efficient, but also it's also more effective (in terms of how much knowledge is truly acquired during the learning process).


9
Nearly 25% of all employees* leave their job because there simply aren't enough training or learning opportunities.
REVENUES
Companies who do offer eLearning and on-the-job training generate about 26% more revenue per employee*.

10
72% of companies who were included in a recent survey stated that eLearning helps them to keep up-to-date with changes in their industry, which helps them to remain competitive within their niche.[4]
Companies and organizations that did have a strong learning culture did better in their market than those who do not. For example:
These organizations are 46% more like to be the leader in their industry
34% increase in their ability to respond to the needs of the customer
17% more likely to become the market share leader.

BIG MONEY

BIG USERS

BIG IMPROVEMENT

TURNIT IN
CLASS BLOGS
e-books
VIRTUAL PROFESSOR
SOCIAL NETWORK
EMAILS
BLACKBOARD DISCUSSIONS
TEAMS INSTANT MESSAGES
CHAT
PODCASTS
WIKI
WI-FI
GPS
iPhone
CD
FM
HD BACK
E-LEARNING OVERLOAD
UPS
© WWW.JOSETOONS.COM
NOV 19, 08

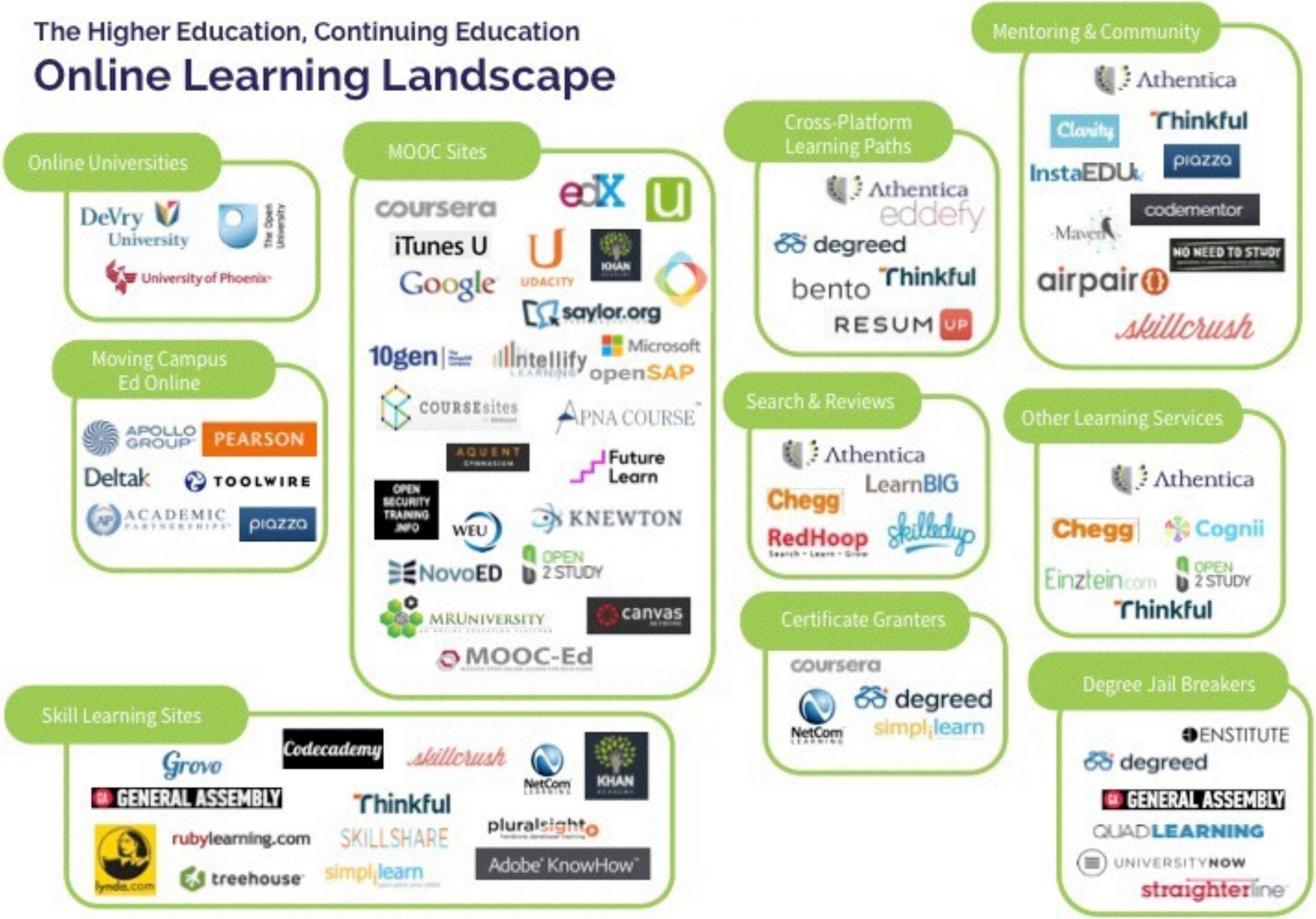
The task of the modern educator is not to cut down jungles, but to irrigate deserts.

C.S. Lewis

Trends in Big Education

The Higher Education, Continuing Education

Online Learning Landscape



http://athentica.com/wp-content/uploads/2013/10/Online-Learning-Landscape-Oct-2013.jpg

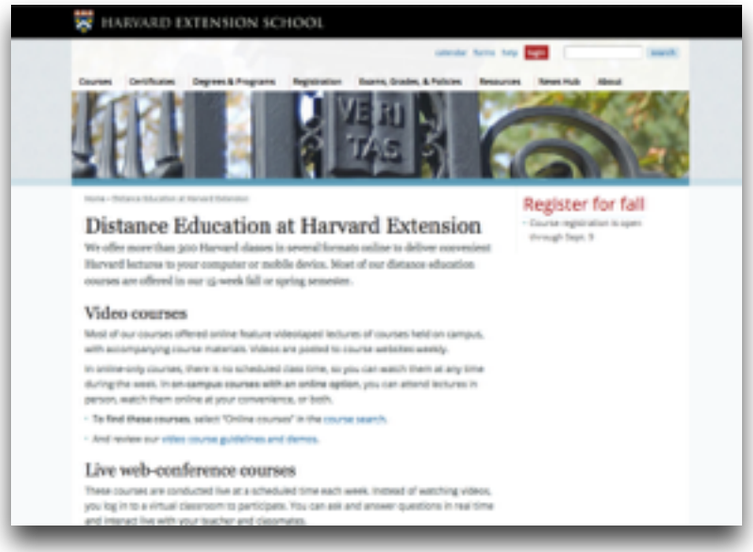
Cost effectiveness

Collaboration

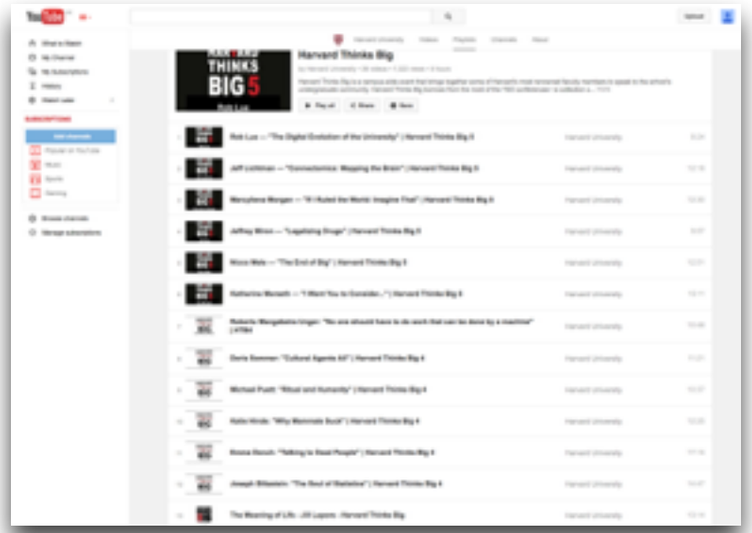
Customization

Multimodal Learning

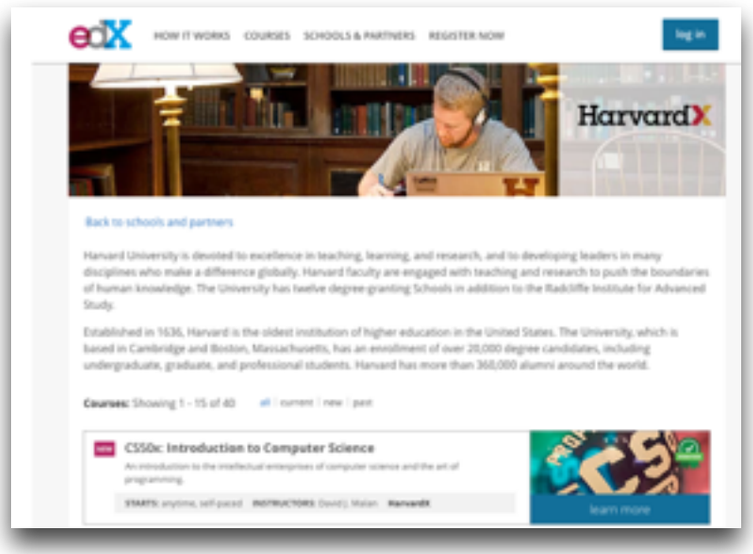
Harvard @ iTunes U



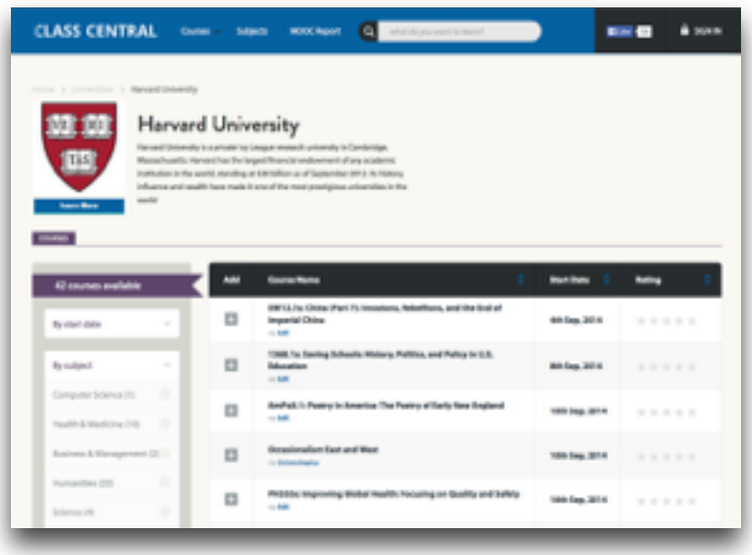
harvard.edu



Harvard @ YouTube



Harvard @ edX



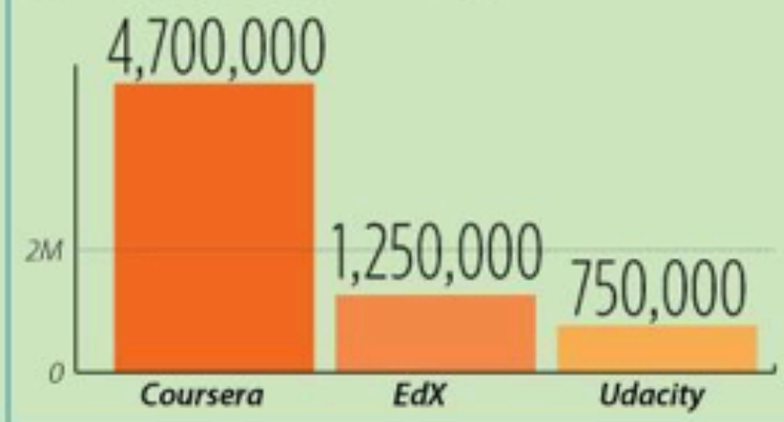
Harvard @ Class Central

MOOC

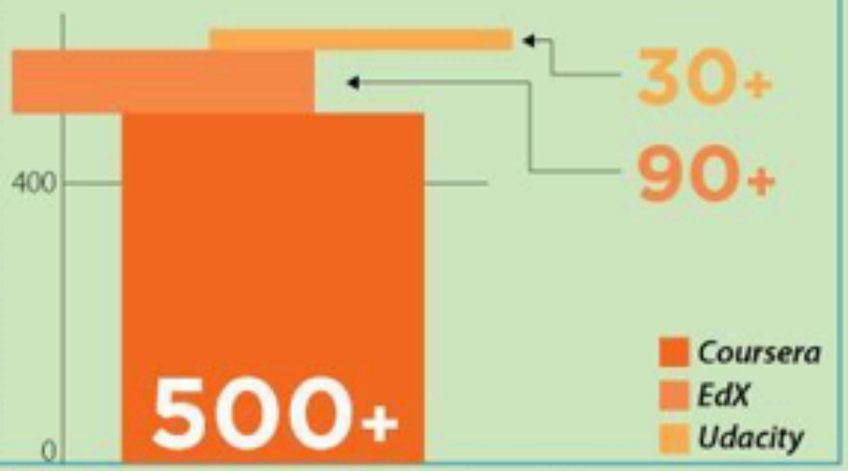
Massive Open Online Course

MASSIVE

Students enrolled in MOOCs



Courses offered by major platforms

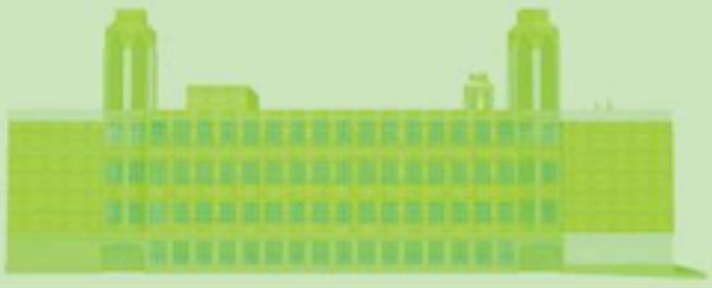


OPEN

These courses are provided by many different universities and open to anyone who wishes to enroll.

87

The number of academic partners that offer courses on Coursera's platform



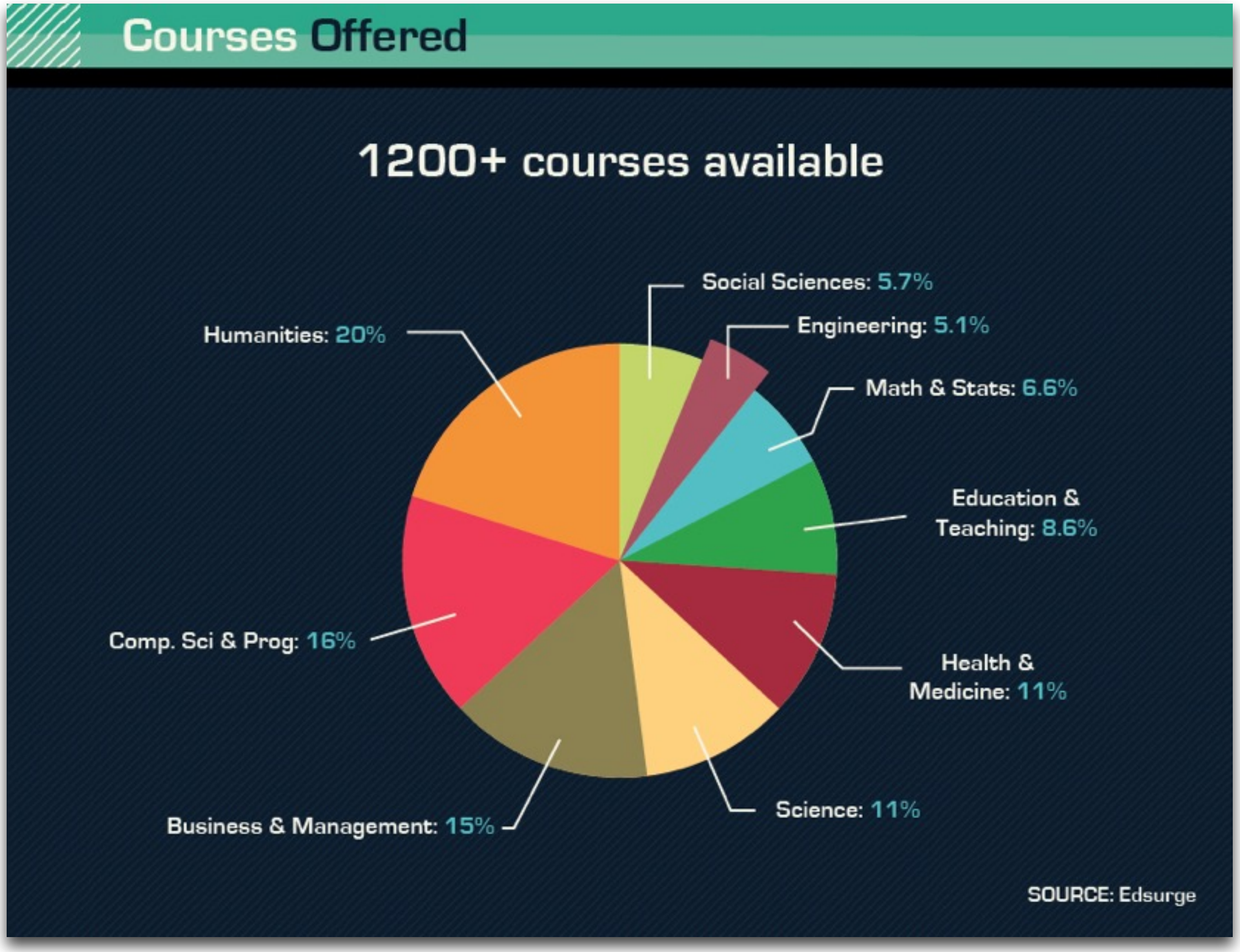
ONLINE

Courses are even reaching as far as developing countries like Mongolia, where high school students are taking courses from:

Major Providers

COURSERA	EDX	UDACITY
Started by Stanford computer science professors Andrew Ng and Daphne Koller in April 2012. For profit	Started by Harvard and MIT in May 2012. Non-profit	Started by Sebastian Thrun and Peter Norvig in Feburary 2012. For profit
5+million students	1.65 million students	1.8 million students
532 courses	125 courses	33 courses
107 partner schools	30 partners	16 partners
Students from 190 countries	225 countries and territories	190 countries

Courses Offered



Small Private Online Course (SPOC) with Degree

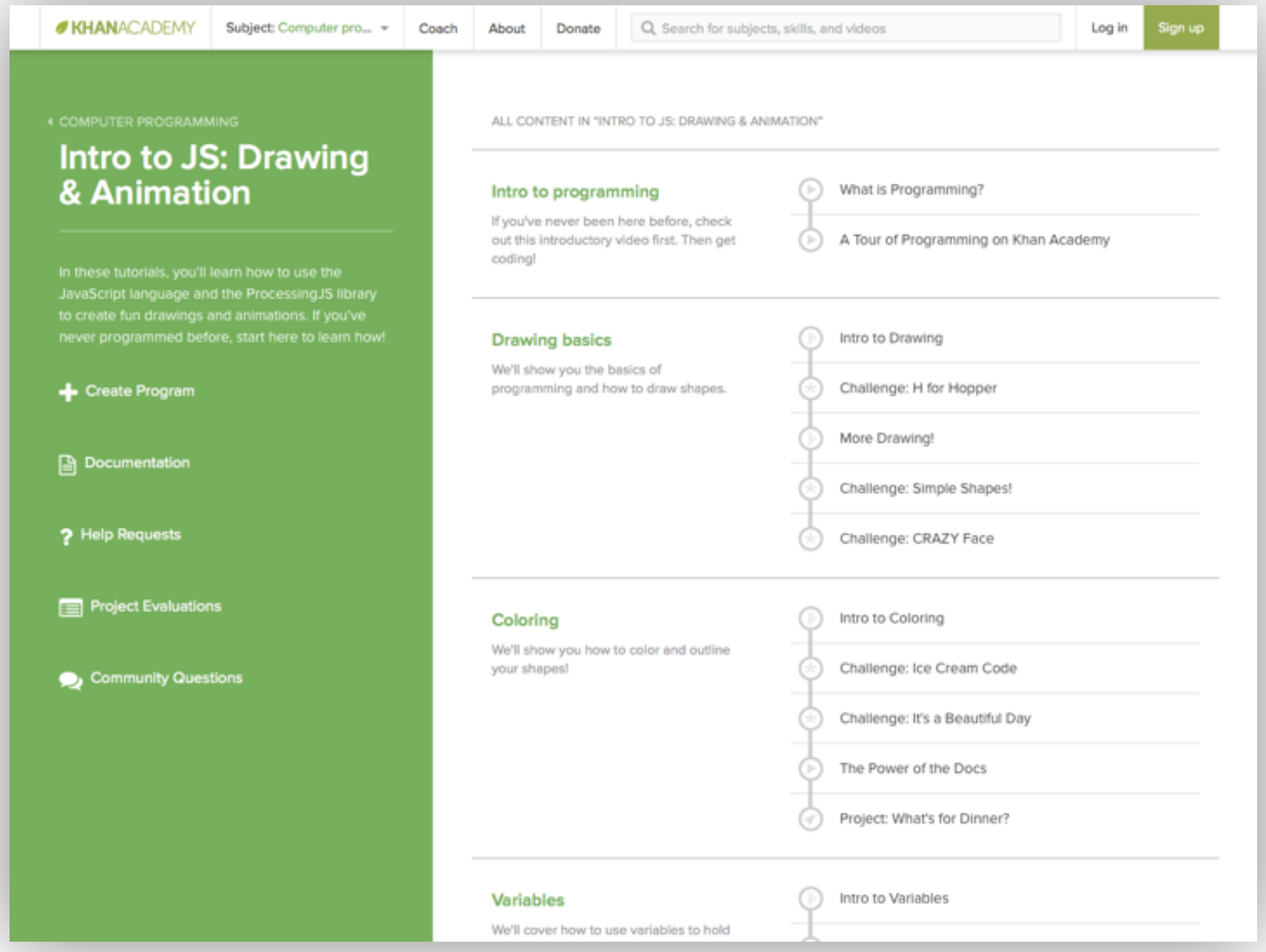
UDACITY

Georgia Tech College of Computing

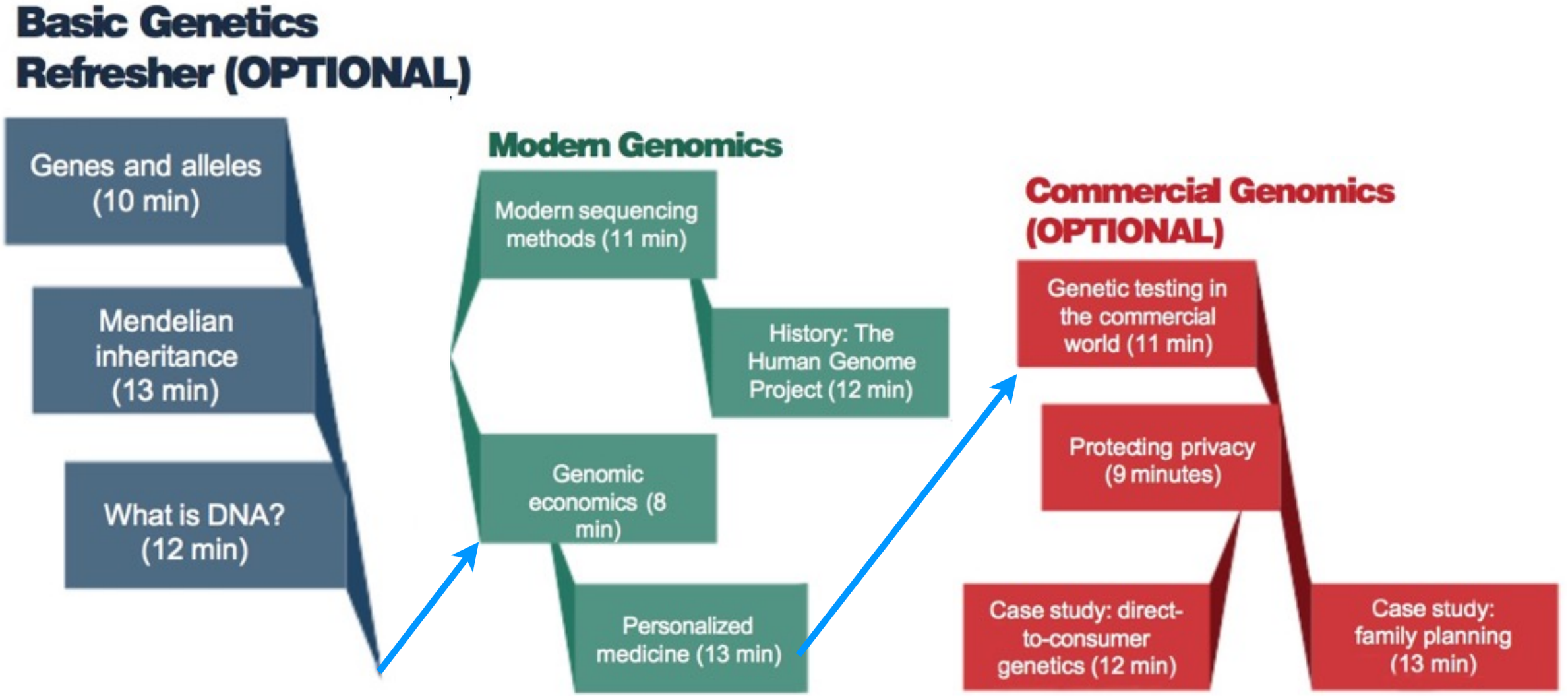
at&t

Flipped Classroom

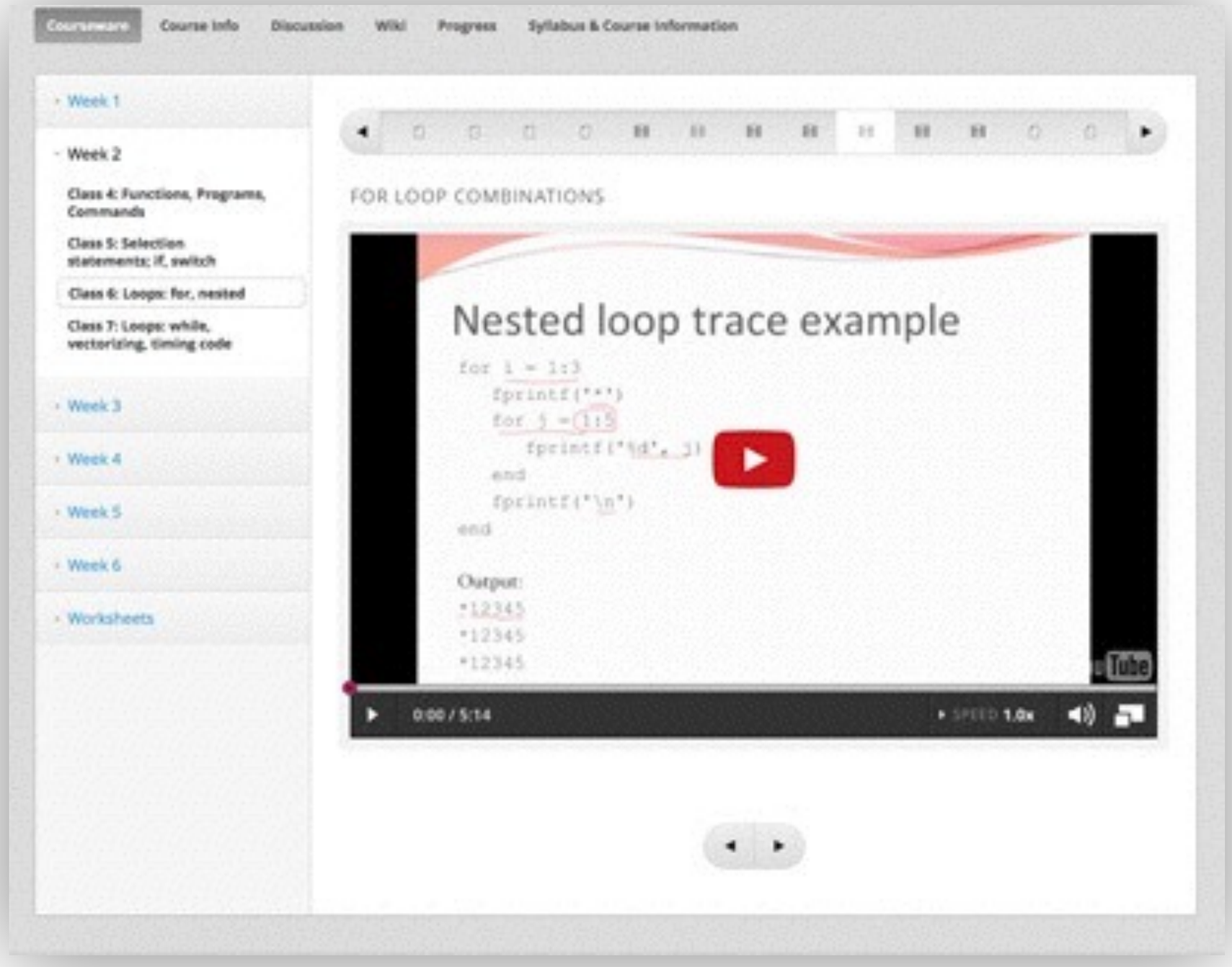
Microlearning



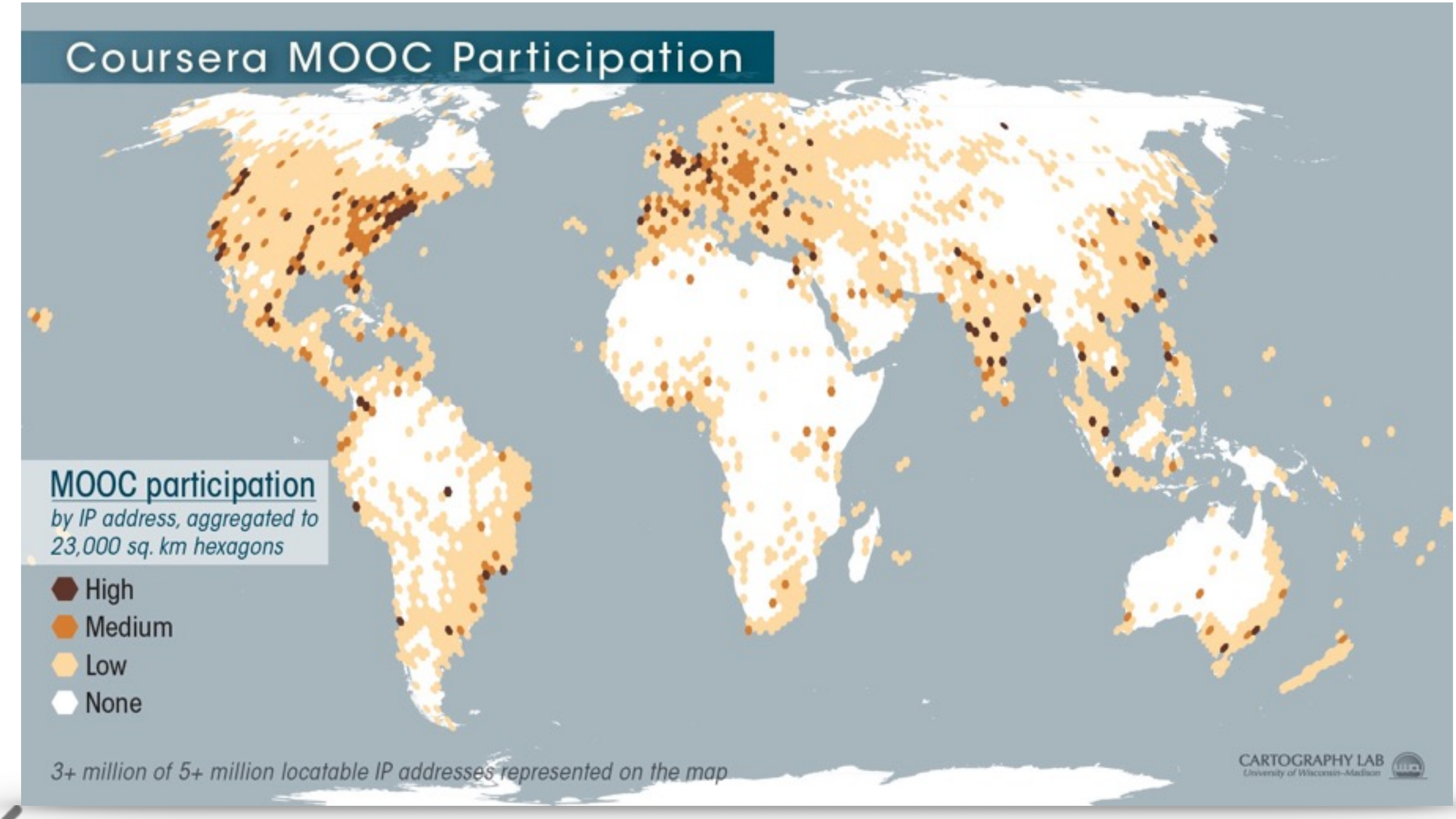
Personalized Learning



Active Learning



Peer Learning



Knowledge and Education Exchange Platform

KEEP

What is KEEP?

- KEEP is a **multi-year** and **cross-institutional** project with strong partners and alliances
- KEEP is the **big data learning analytics** cloud platform
- KEEP serves as a creative **online learning gateway** for educators and learners around the world
- KEEP encourages and promotes **flexible and active learning**
- KEEP is a **knowledge aggregator** and a **technology integrator**

Our Vision

To **empower people** by providing and promoting the **best education resources** in order to facilitate collaboration and innovation for **teaching and learning** through **knowledge aggregation** and **technology integration**

Our Mission

- Organize and search accessible and useful materials
- Gain fresh insights through analytics
- Collect the best online courses
- Promote innovative education applications

KEEP's Partners & Alliances

UGC 大學教育資助委員會 University Grants Committee

coursera edX

course-builder Course Builder

eWant education you want! 育網開放教育平台

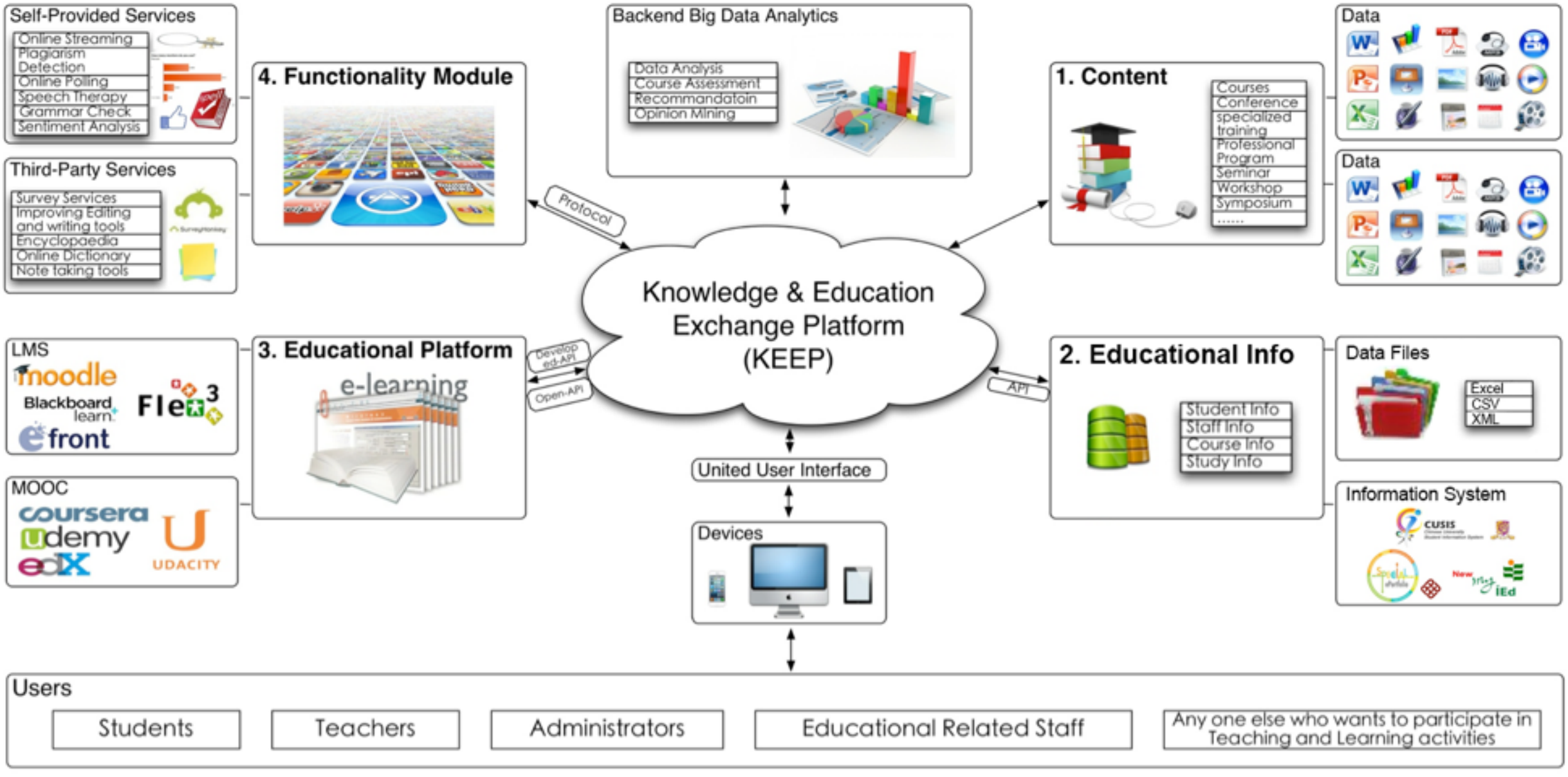
OPENedX

xuetangX.com 学堂在线

Lingnan 嶺南大學 University

Google amazon web services™ android HUAWEI Microsoft

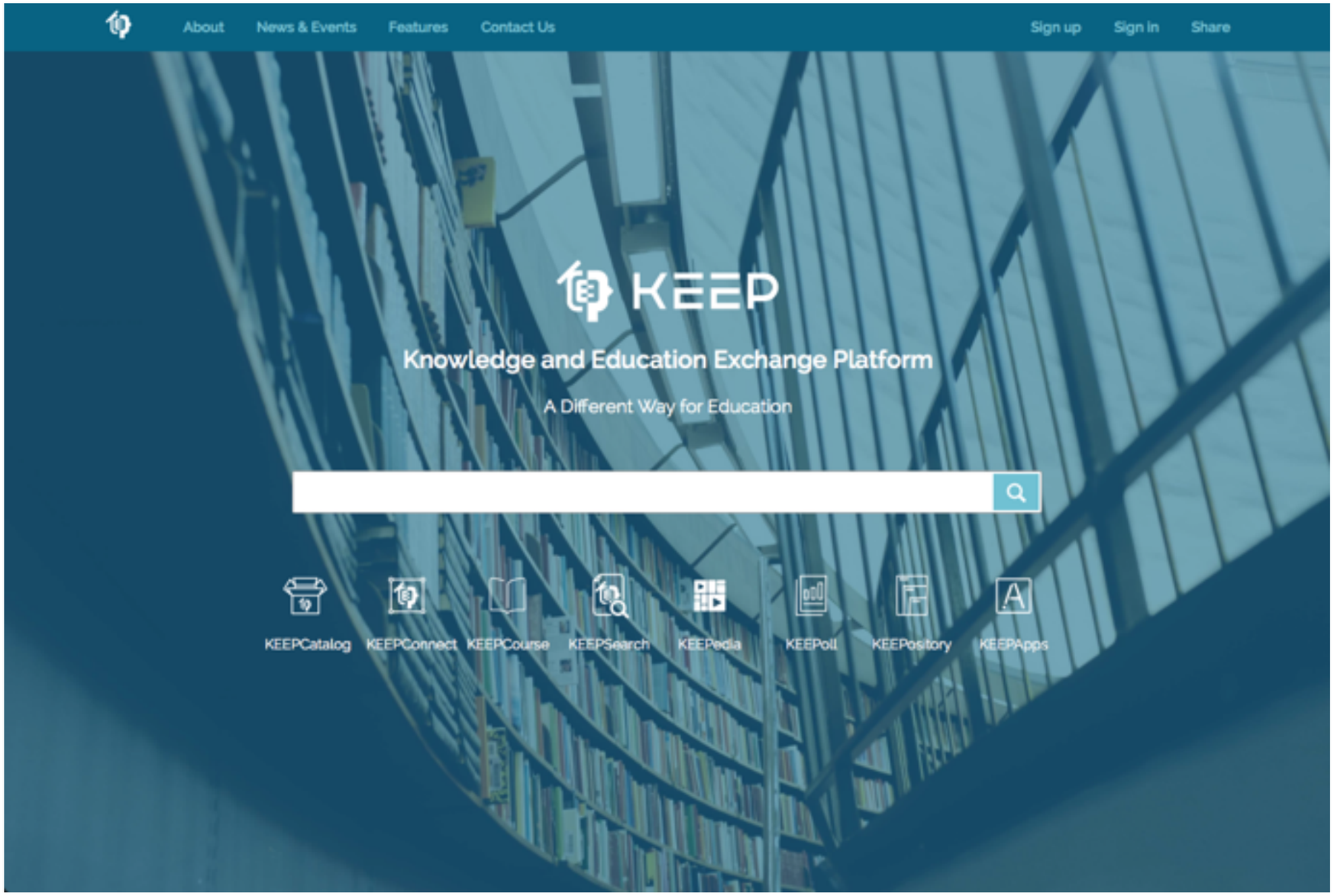
KEEP Education Cloud



KEEP education cloud

Our Webpage

www.keep.edu.hk



KEEPSearch

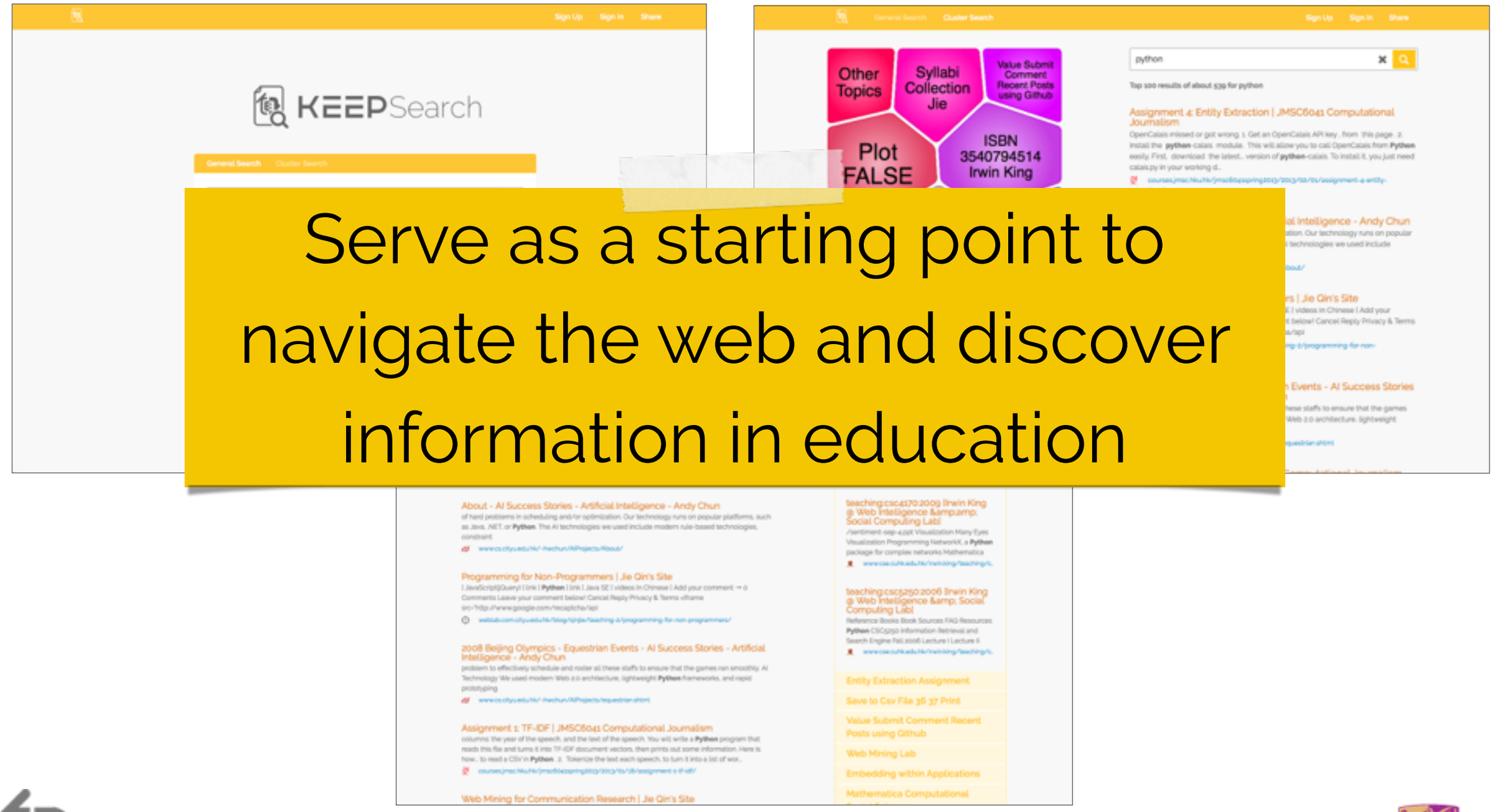
Problem we are facing:

- Various resources and information online
- People are struggling for these kind of resources
- Google and other search engine do not have this kind of ability that focus on education subject information searching

Benefit:

- Organize these kind of resources and information
- Bring the best and efficient way for people to retrieve useful and accessible resources

KEEPSearch



Serve as a starting point to navigate the web and discover information in education

KEEPCourse

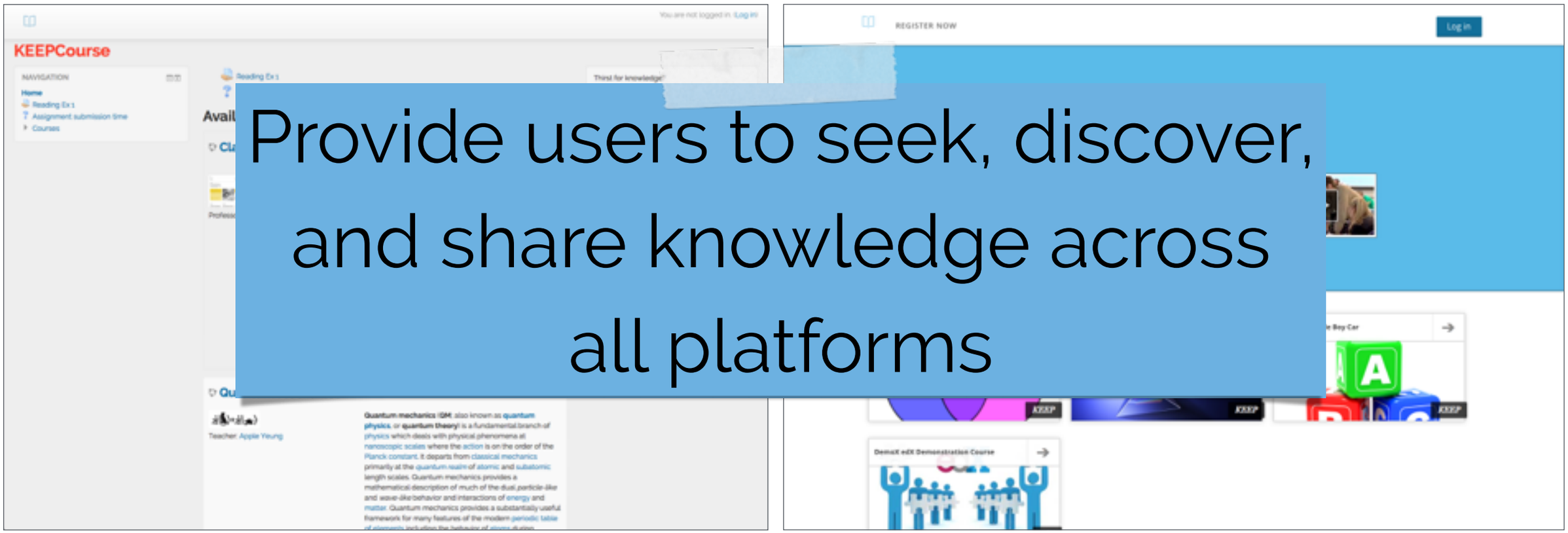
Problem we are facing:

- The idea of MOOC/LMS is popular around the world for a period of time
- Different kind of education platforms are built, and run independently e.g. Coursera, edX, Course Builder, Moodle, Sakai
- Learners have to visit them separately

Benefit:

- Provide one stop solution to search all kinds of courses from different platform
- Provide content hosting that helps educator to promote their courses

KEEPCourse



Provide users to seek, discover, and share knowledge across all platforms

KEEPedia

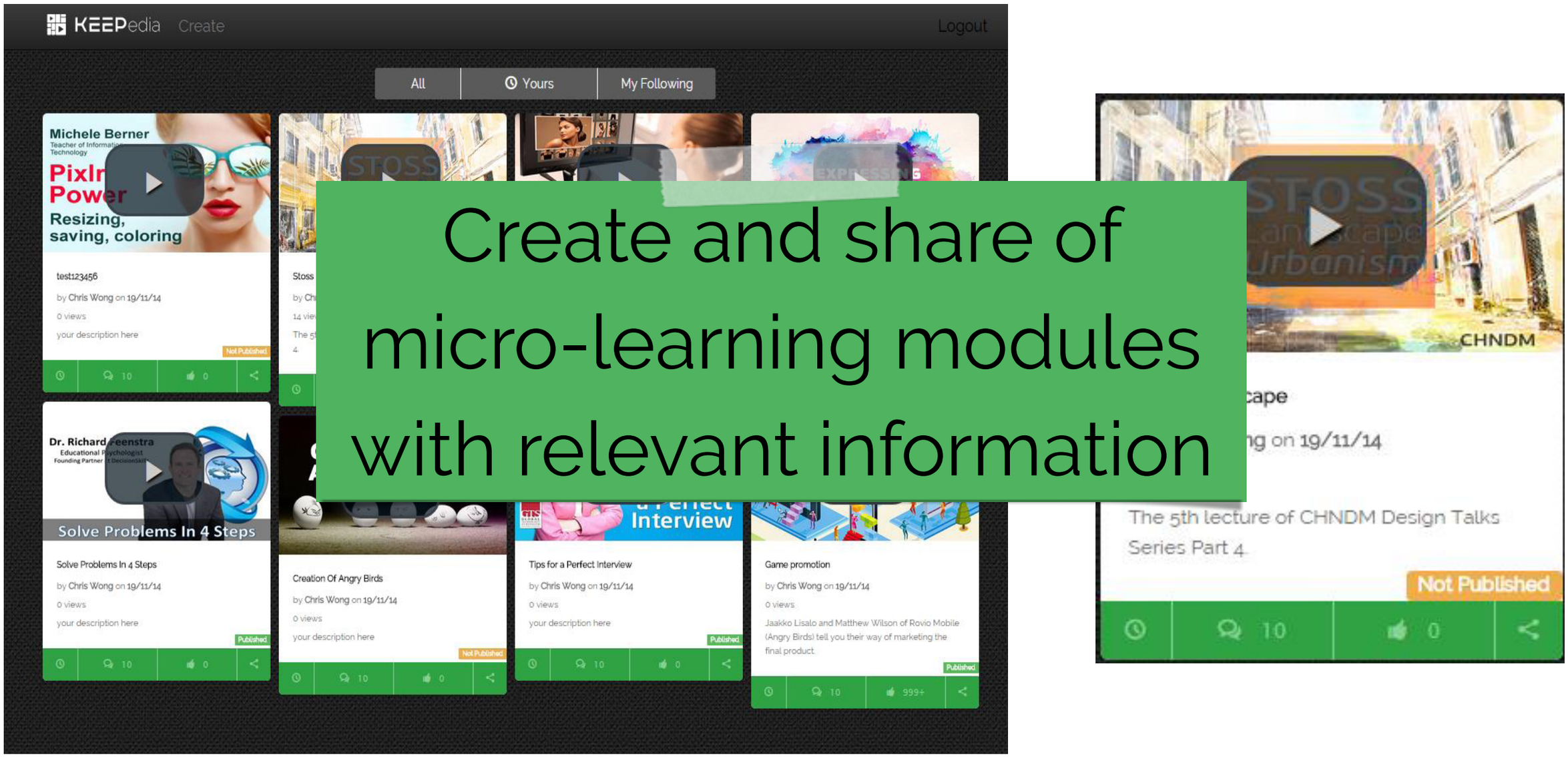
Problem we are facing:

- Few "integrated" / "one-stop" / "comprehensive" video delivery services designed for education and provide materials closely related to education purpose
- Few services have tracking ability to collect users video behavior data for learning analytics

Benefit:

- Each micro-module will focus on a single and simple topic
- Study a topic through comprehensive learning materials, e.g. video, text, exercise, quiz, etc.
- Able to track user behavior for learning analytics

KEEPedia



Create and share of micro-learning modules with relevant information

KEEPCatalog

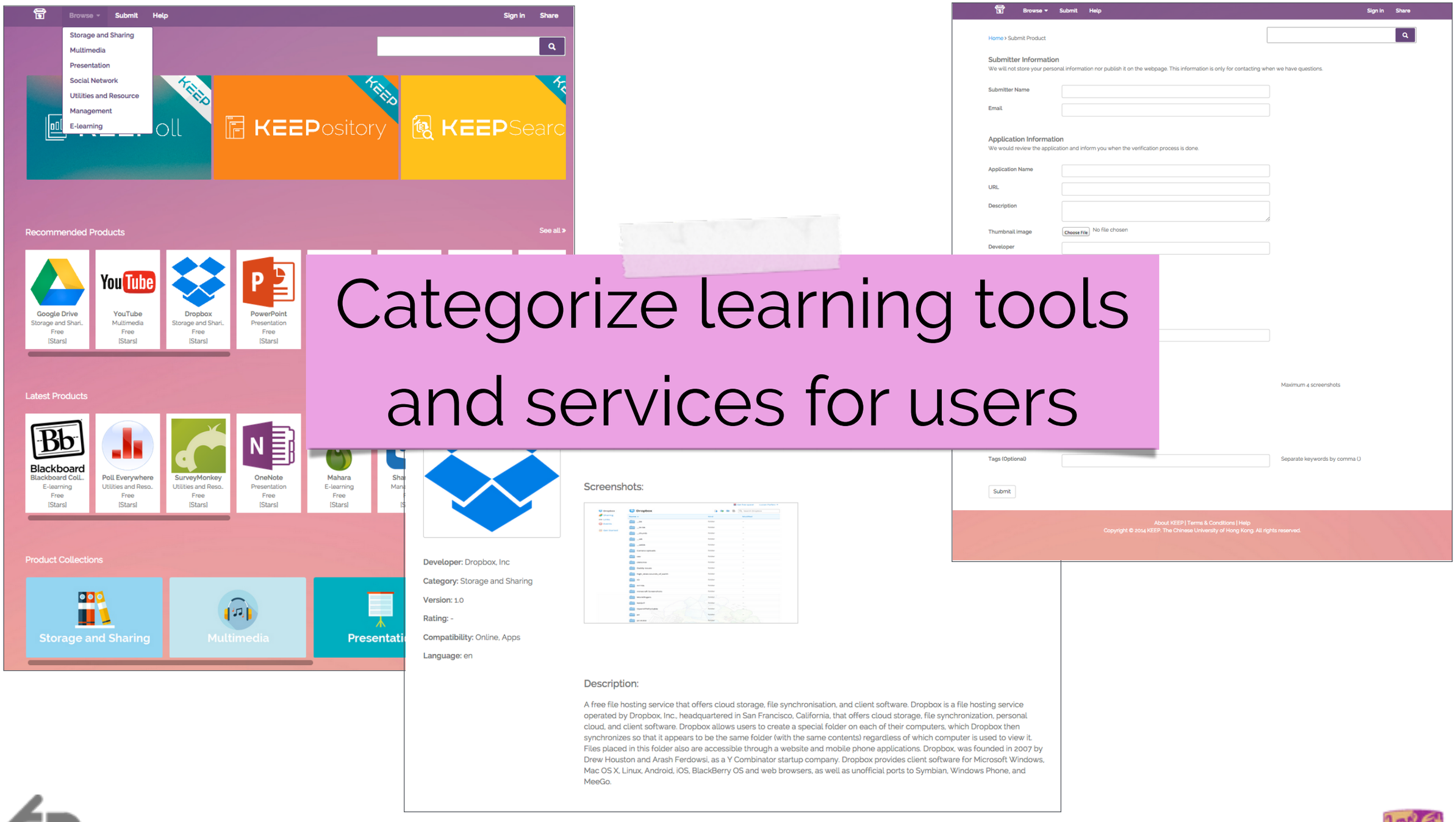
Problem we are facing:

- Too many software products on the Internet that facilitates knowledge spreading and self-learning in an electronic way
- A well-organized list is needed for people to find suitable education products

Benefit:

- Collect, categorize and show educational applications available online
- Allow users to submit new products to enhance the completeness of the collection

KEEPCatalog



Categorize learning tools and services for users

KEEP Apps

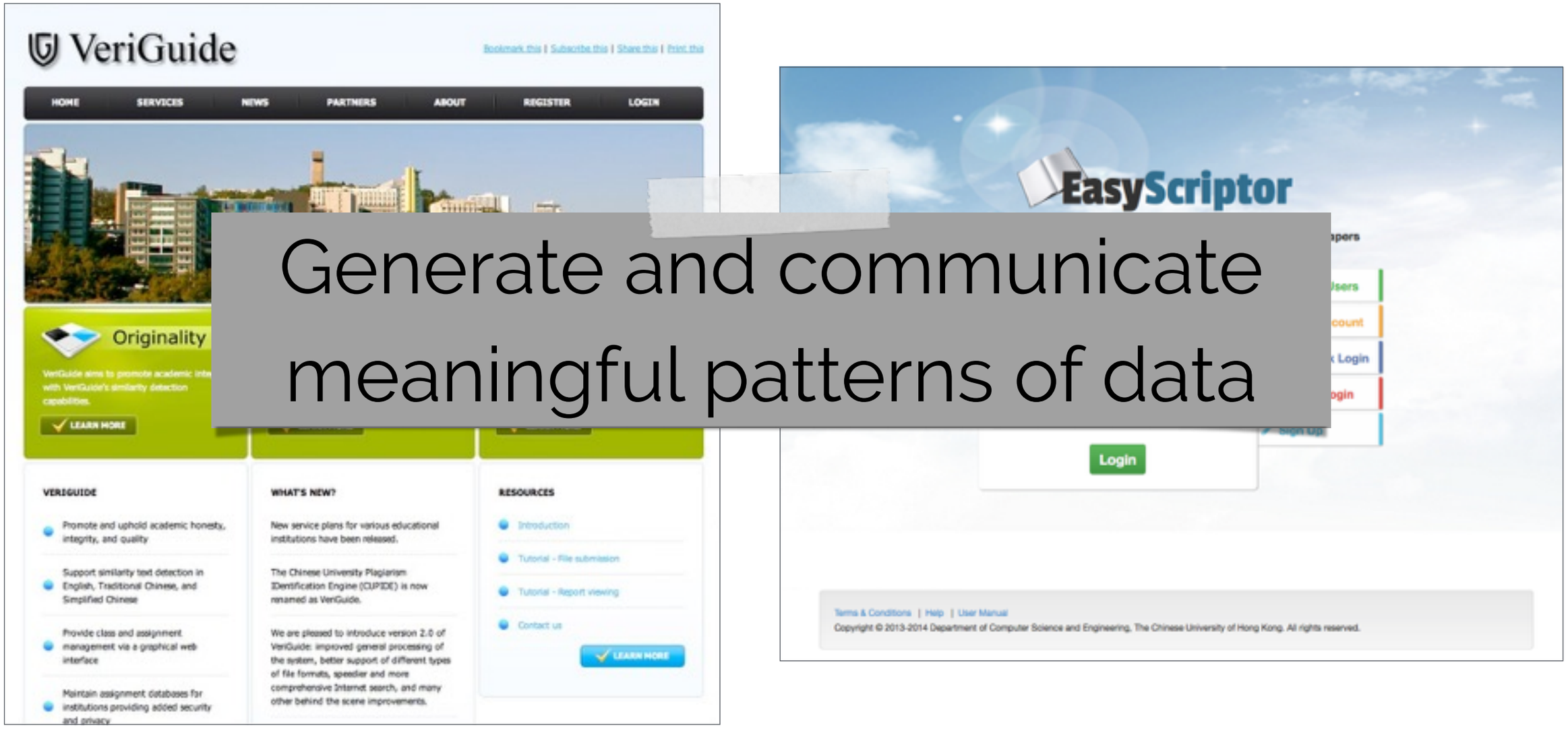
Problem we are facing:

- Applications are too scattered
- Lack of available application data analytics online

Benefit:

- Facilitate the development and exhibition of innovative education applications
- Facilitate application data collection and behavior tracking

KEEP Apps



Generate and communicate meaningful patterns of data

KEEPoll

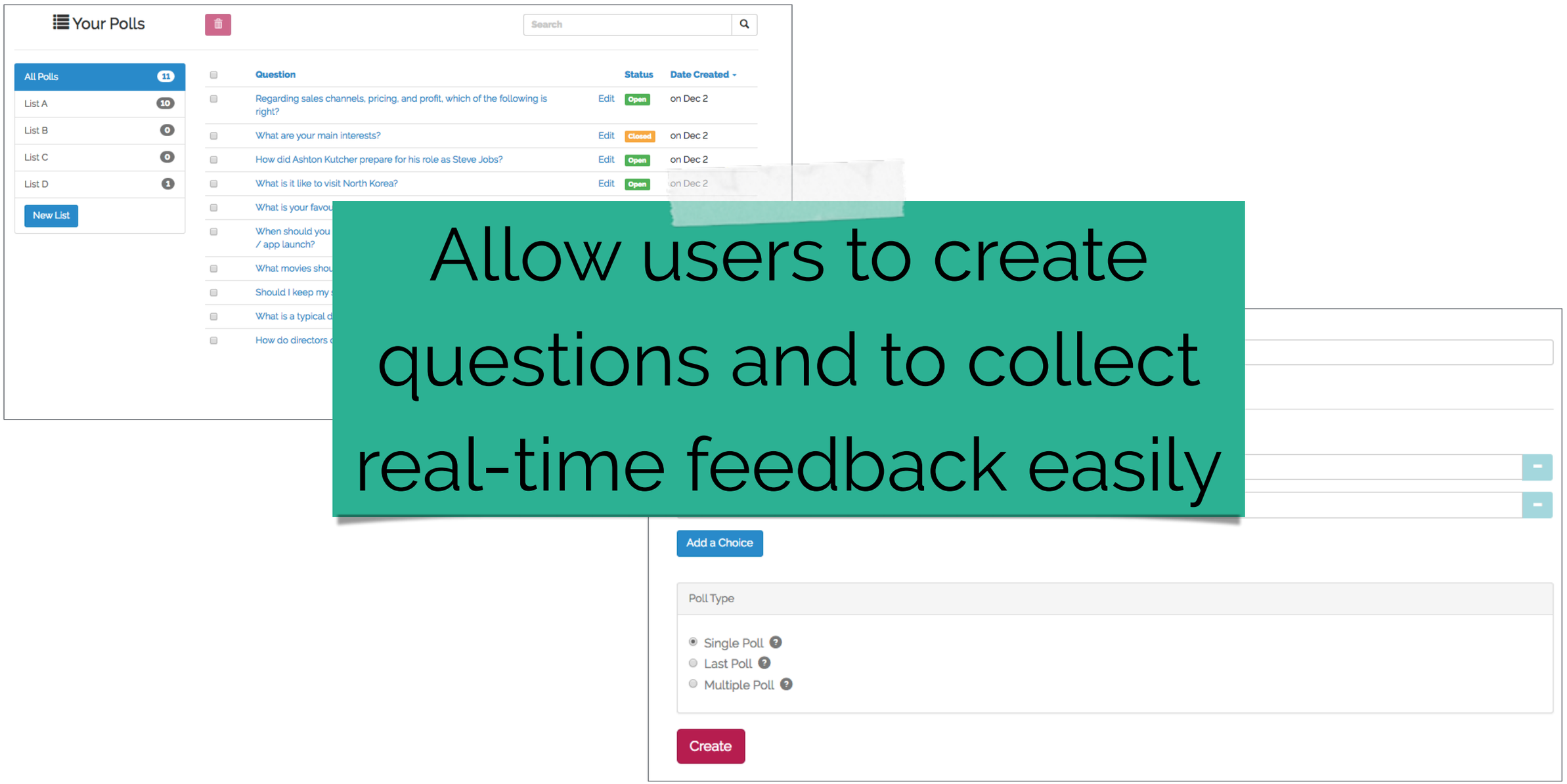
Problem we are facing:

- Hard to get real-time feedback
- Old voting methods are inaccurate and take far too long (rise hand…)
- Voting machines are expensive and difficult to setup

Benefit:

- Allow users to create questions to collect real-time feedback easily
- Able to respond with laptops, tablets and mobile phones

KEEPoll



Allow users to create questions and to collect real-time feedback easily

KEEPository

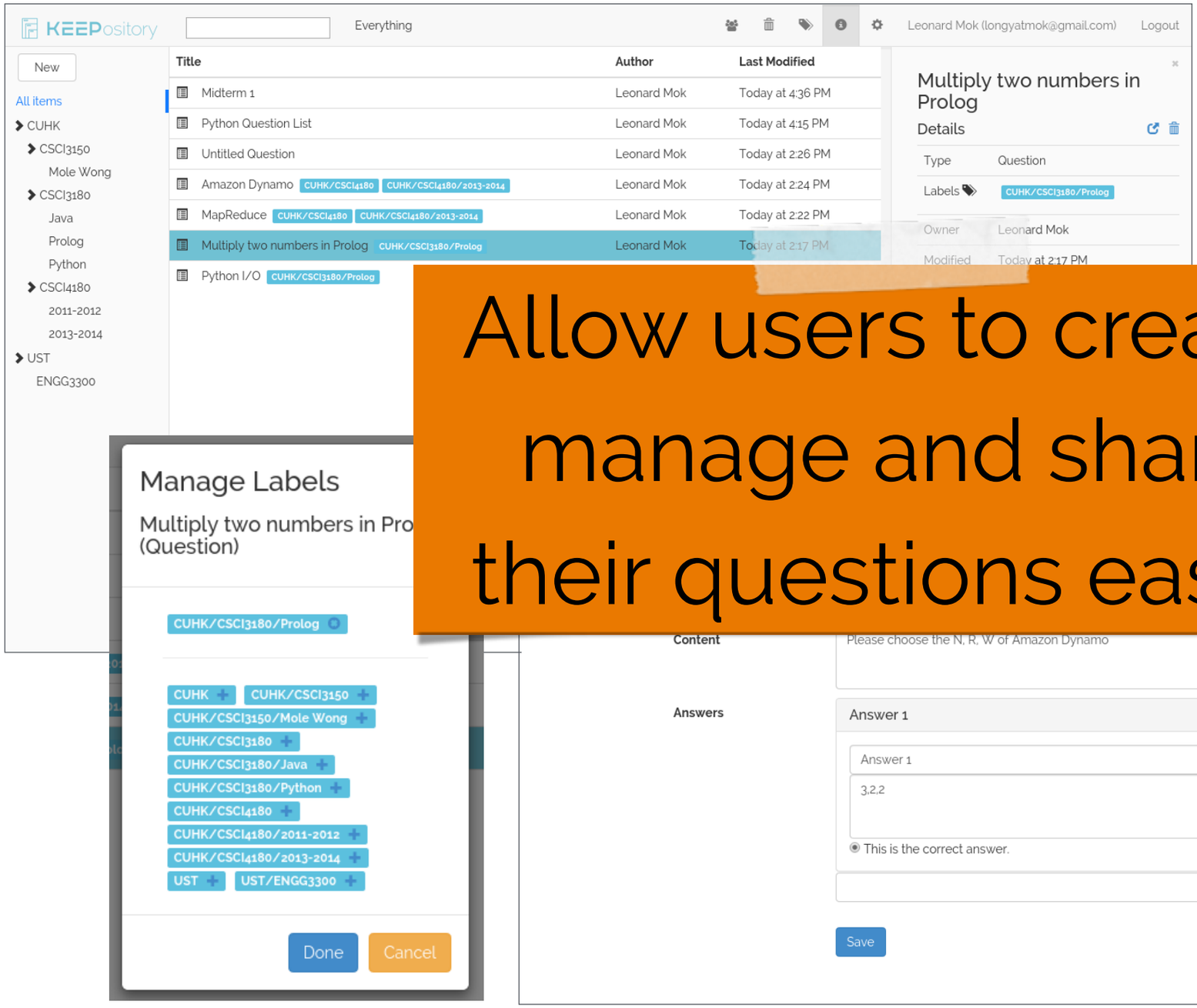
Problem we are facing:

- Learning information and statistics hard to store and analyze
- Questions, quizzes and learning materials cannot be reuse or share
- Student cannot immediately get result for self-review

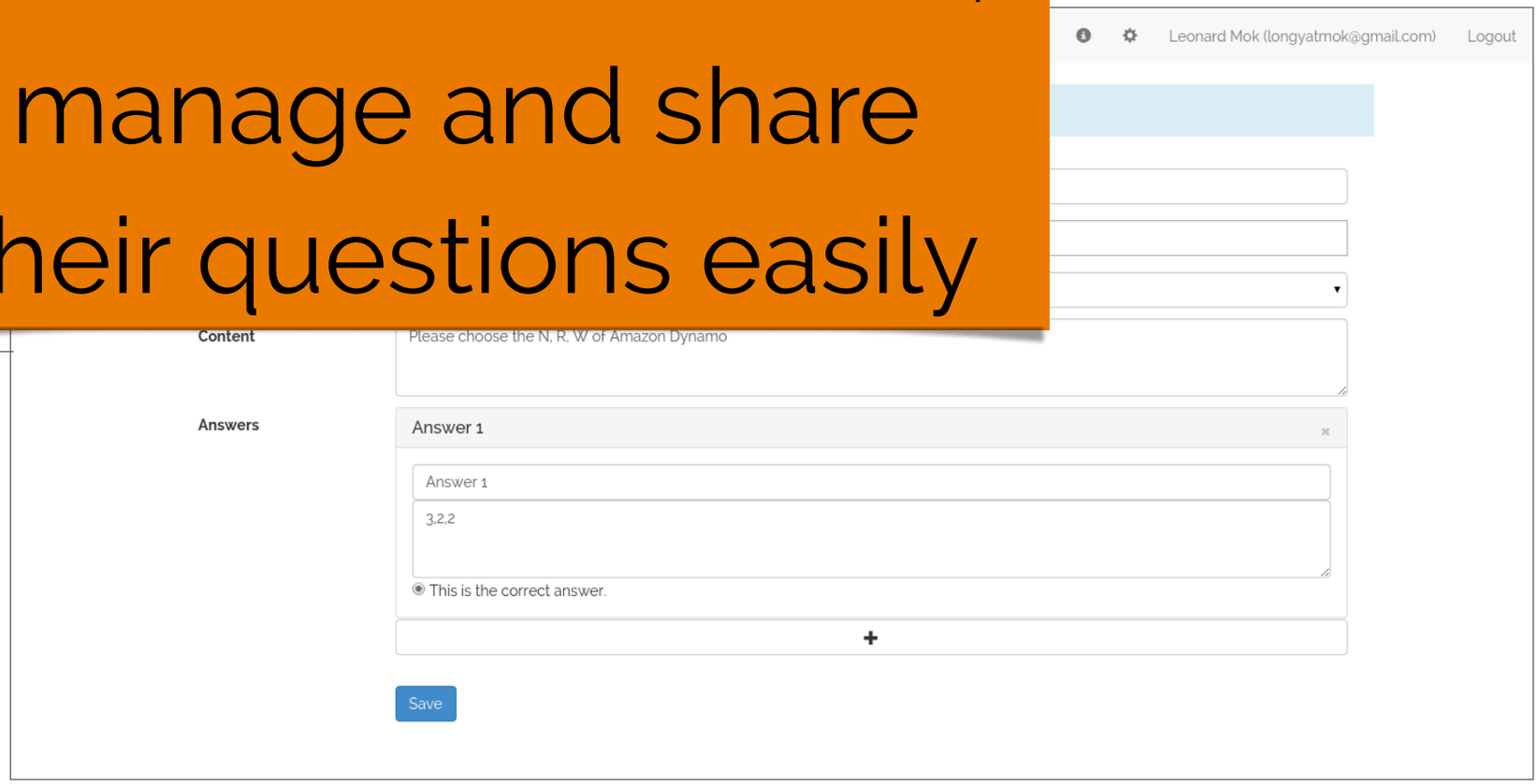
Benefit:

- View and examine questions created by other teachers
- Create online quizzes to evaluate students' learning progress
- Get quizzes statistics for improving teaching strategies
- Provide high quality online assessment for students

KEEPository



Allow users to create, manage and share their questions easily



Work in progress
&
Future works

Natural language processing

- Text and semantic analysis
- Summarization
- Sentiment analysis
- Automated grading
- Q&A systems

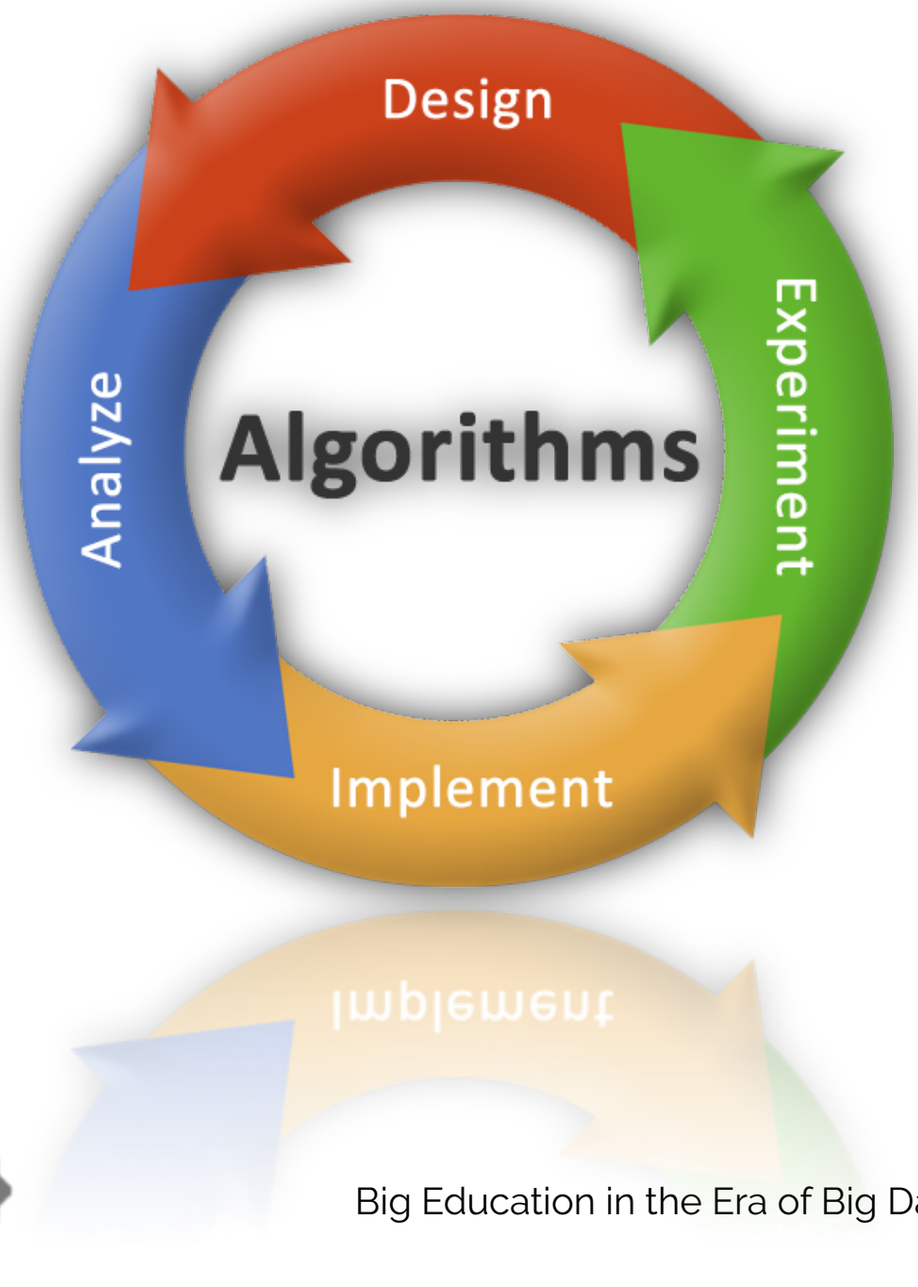
Recommendations

- Personalized learning
- Courses, tutors, peering learning partners, etc.
- Learning resources, time allocation, etc.
- Career planning

Knowledge map

- Explore topics
- Track topic changes
- Make topic comparisons and inferences
- Better search on concepts

Algorithms & Techniques



- Machine learning
- Data analytics
- Social computing
- Web intelligence Multimedia information processing
- Gamification

Are You Ready?

Get involved

"I think you'll find that mine is bigger..."

Concluding Remarks

- *Be Inspired*
  - Big Education is the focus!
- *Be Informed*
  - Big Data in Education is the ***VALUE*** proposition!
- *Be Challenged*
  - Use technologies to transform education in the Big Data Era!

Acknowledgements

- Byron Lai
- Daisy Lau
- Jamie Yeung
- Jean Yao
- Junfeng Yang
- Lin Tsang
- Patrick Lau
- Raymond Yuen
- Roger Cheung
- Sophia Man

Looking for more engineers, programmers, system analysts, etc. to work on KEEP...

Acknowledgements

- Shouyuan Chan (Microsoft, USA)
- Xixian Chen (Ph.D.)
- Chen Cheng (Ph.D.)
- Junjie Hu (Ph.D.)
- Baichuan Li (Baidu, China)
- Guang Ling (Ph.D.)
- Haiqin Yang (Postdoc)
- Connie Yuen (Ph.D.)
- Hongyi Zhang (Ph.D.)
- Shenglin Zhao (Ph.D.)
- Tong Zhao (Ph.D.)

- Looking for more PhD students working on machine learning, Big Data, social computing,...

VeriGuide™

- Similarity text detection system created in 2005
- 145,000 users and 280,000 submissions for year 2005-2013
- Track students' progress in writing

ACML 2015

7th Asian Conference on Machine Learning
November 20-22, 2015, Hong Kong

Home
Committees
Call for Papers
Paper Submission Site
Invited Talks
Accepted Papers
Keynote Speakers
Workshops
Program
Registration
Venue Information
Accommodation
Contact Information
Sponsors
Previous Conferences
About Hong Kong

Organizers

Seventh Asian Conference on Machine Learning, Hong Kong

The Chinese University of Hong Kong and Hong Kong Science Park
November 20-22, 2015

ANNOUNCEMENTS

- We are calling for paper submissions, workshop proposals, and tutorial proposals.

ACML 2015

The 7th Asian Conference on Machine Learning (ACML2015) will be held in Hong Kong on November 20-22, 2015. The conference aims to provide a leading international forum for researchers in machine learning and related fields to share their new ideas, progresses and achievements. Submissions from regions other than the Asia-Pacific are also highly encouraged.

The conference calls for high-quality, original research papers in the theory and practice of machine learning. The conference also solicits proposals focusing on frontier research, new ideas and paradigms in machine learning. The conference proceedings will be published in The Journal of Machine Learning Research (JMLR): Workshop and Conference Proceedings series.

See the Call for Papers for details and submit through the paper submission site.

IMPORTANT DATES

BigScholar 2015

Welcome to
BigScholar 2015

The Second WWW Workshop on
Big Scholarly Data: Towards the Web of Scholars
http://msclab.org/bigscholar/

A workshop of WWW 2015 (The 24th International World Wide Web Conference)
Florence, Italy, May 19, 2015

The BigScholar 2015 workshop aims at bringing together researchers and practitioners working on Big Scholarly Data to discuss what are emerging research issues and how to explore the Web of Scholars. Several core challenges, such as the tools and methods for analyzing and mining scholarly data will be the main center of discussions at the workshop. The goal is to contribute to the birth of a community having a shared interest around the Web of Scholars and exploring it using data mining, recommender systems, social network analysis and other appropriate technologies.

The workshop will be a half-day workshop. The format of the workshop will include one invited talk (keynote), research and position paper presentations and one discussion panel. The workshop will be held in the afternoon on May 19, 2015 in Florence, Italy, in conjunction with the 24th International World Wide Web Conference (WWW 2015).

Important Dates

Paper submissions due:
Jan 24, 2015
Notification of acceptance:
Feb 22, 2015
Camera ready version due:
Mar 8, 2015
Workshop date:
May 19, 2015

History

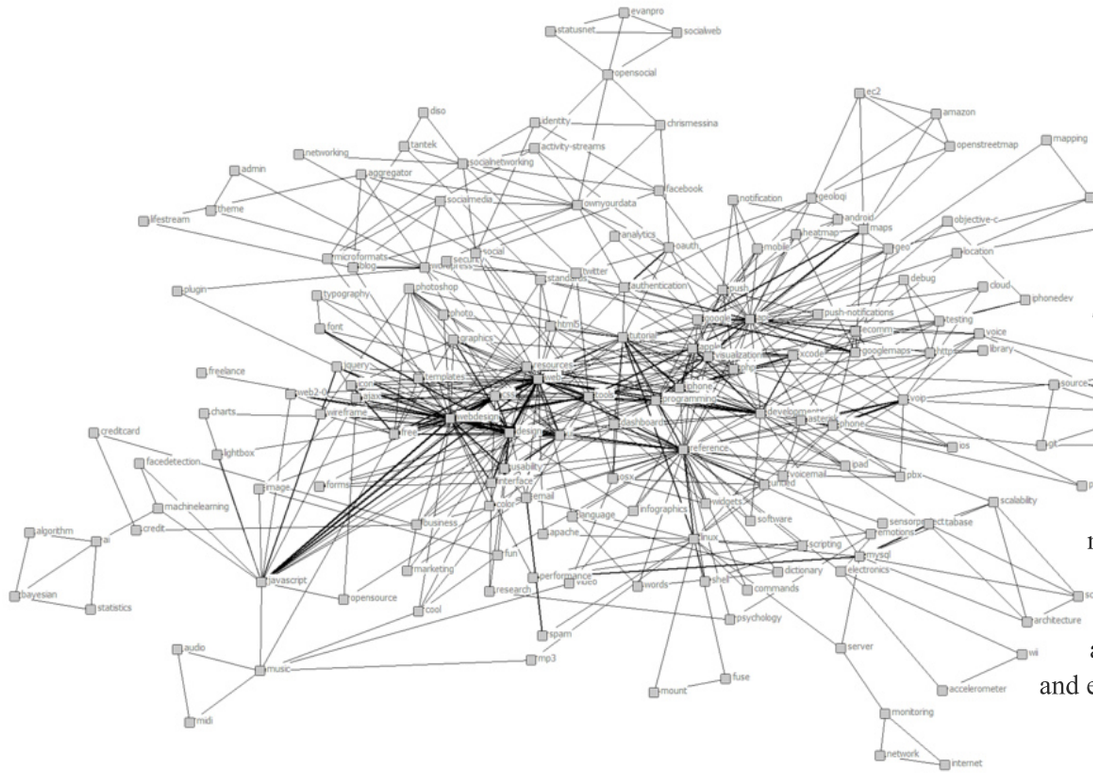
- BigScholar 2014
  Seoul, Korea, April 2014

WWW2015 Workshop on Web-based Education Technologies (WebET 2015) May 19, 2015, Florence, Italy

New Book Series

SOCIAL MEDIA & SOCIAL COMPUTING

CALL FOR BOOKS!



The Social Media and Social Computing Series focuses on publishing high quality references in the rapidly emerging area of social media and social computing. Both experimental/practical as well as theoretical investigations are welcome. The series targets both scholars and practitioners in social media and social computing for work in the intersection of computer science, information technology, psychology, economics, education and other social sciences.

The advent of the Internet and the Web has resulted in social interactions and behaviors through the use of technologies and web services, e.g., hardware devices such as smart phones, tablets, RFID, etc., software services such as wikis, blogs, micro-blogs, social network sites, recommender systems, social bookmarking, social news, multimedia sharing sites, etc. Analyzing these technologically-enabled interactions in their social context will benefit information providers and information consumers. However, the large volume and scale of user-generated contents require effective modeling methods and efficient algorithms to handle these chalenging problems.

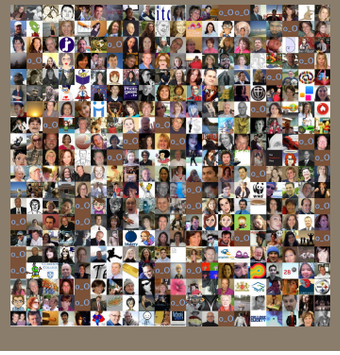
Series Editor:

Irwin King

Prof. King is Associate Editor of the IEEE Transactions on Neural Networks (TNN) and IEEE Computational Intelligence Magazine (CIM). He is a senior member of IEEE and a member of ACM, International Neural Network Society (INNS), and VP & Governing Board Member of the Asian Pacific Neural Network Assembly (APNNA) . He serves the Neural Network Technical Committee (NNTC) and the Data Mining Technical Committee under the IEEE Computational Intelligence Society.

https://www.cse.cuhk.edu.hk/irwin.king/home

kin@cse.cuhk.edu.hk



Pleas send Proposals to either the Series Editor or Directly to :

LEONG Li-Ming
Editor, CRC Press
240 Macpherson Road, #08-01,
Pines Industrial Building,
S'pore 348574
li.ming@tandf.com.sg
Tel: (65) 67415166 x 115

The Chinese University of Hong Kong

Q & A

Q & A