



International Journal of Control Theory and Applications

ISSN : 0974-5572

© International Science Press

Volume 9 • Number 45 • 2016

Emerging Role of Crowdsourcing in MOOCs: A Review

Sapan H. Mankad^a

^aCSE Department, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

E-mail: sapanmankad@nirmauni.ac.in

Abstract: With the advent of online and blended learning models, the traditional education is supplemented with many rich features. However, due to simultaneous enormous enrollments in Massively Open Online Courses (MOOCs), it becomes difficult to manage such a large pool of participants. Managing tasks such as assignment grading or feedback response mechanism becomes very challenging for a single or limited instructors. Crowdsourcing has shown significant success to resolve such issues to a great extent. This paper reviews existing approaches on improving current Education scenario with the help of Crowdsourcing and discusses the characteristics of Crowdsourcing-enabled learning strategies, mainly in MOOC settings.

Keywords: Crowdsourcing, Discussion Forums, Peer Assessment, Gamification, MOOC, Human Computer Interaction, Collaborative Learning.

1. BACKGROUND

1.1. Why machines are not enough?

Imagine a scenario in which you want a task to be accomplished in a limited amount of time. It is possible to automate the task using machines, but the constraint is that you want the task to be done by humans; because in contrast to machines, some tasks can be effortlessly completed by humans with the help of rich set of naturally gifted senses. Typical tasks include tagging different persons in an image or filling in critical surveys, or replacing missing values in databases by most appropriate alternatives or translating a text passage¹, or reviewing a product² and so on.

1.2. Why a single domain expert is not enough?

It is reported in literature that rather than a single domain expert, a group of non-experts can yield better performance in such tasks; and the group is called Crowd. The collective intelligence of crowds can be used to solve a wide range of tasks¹.

1.3. What is Crowdsourcing? And how it works?

According to ³, “Crowdsourcing enables programmers to incorporate *human computation* as a building block in algorithms that cannot be fully automated, such as text analysis and image recognition. Similarly, humans can be used as a building block in data-intensive applications.”

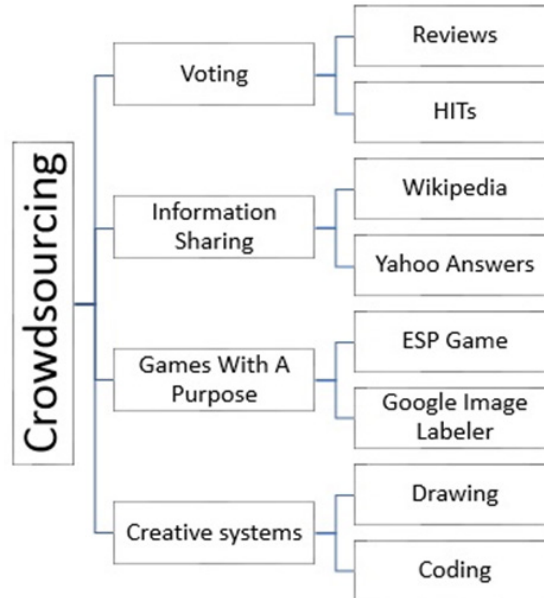


Figure 1: Classification of Crowdsourcing applications

Crowdsourcing is a mechanism through which a requester can hire workforce to execute a task, which is decomposed into small sub-tasks, often referred to as Human Intelligent Tasks (HITs). Each responder in the crowd is assigned one HIT usually and the response is submitted to the requester. Many platforms have emerged which simplify Crowdsourcing operations such as *Amazon Mechanical Turk* (AMT) and *Crowd Flower* to name a few. Figure 1 depicts classification of Crowdsourcing applications, adopted from ⁴.

Jeff Howe in his introductory article on Crowdsourcing ⁵ writes, “Hobbyists, part-timers, and dabblers suddenly have a market for their efforts, as smart companies in industries as disparate as pharmaceuticals and television discover ways to tap the latent talent of the crowd. The labour isn’t always free, but it costs a lot less than paying traditional employees. It’s not outsourcing; it’s crowdsourcing.” The term Crowdsourcing is derived from *Crowd* computing and *Outsourcing*.

1.4. Why would Crowd work for you?

It is straight-forward that very few people would like to work for you without any expectation. Some strategies have been proposed to offer Crowd workers to incentivize for the tasks done by them. The incentives or rewards can be in any form. Crowdsourcing workers can be categorized as money driven or entertainment driven workers⁶. Some examples such as *Google Image Labeler* or *Online ESP Game* work on the principle of offering *fun* as incentive to the users. These are examples of *Games With A Purpose* (GWAP). But as stated earlier, the key motivating factor for Crowd is financial incentive. Most of the companies, nowadays, offer some form of monetary reward to their crowd workers. Typically, there is a fixed amount associated with each HIT assigned to the worker. Again, this kind of offering has a major lacuna as follows. Some of the *malicious* workers may indulge into degrading the quality of the task in an attempt to increase their incentives by submitting more HITs. Hence, quality suffers when quantity of tasks done by a single worker increases. This issue is addressed in later section of this article.

1.5. Applications of Crowdsourcing

Crowdsourcing has been extensively used in different fields of applications ranging from mobile data management⁷ to Information Retrieval Systems⁸, and Music Production⁹ to News Reporting¹⁰. It has also become prominent in e-Governance as reported in¹¹. The author believes that days are not far when the court judgments will not be given by a single judge or a panel of judges, but by a crowd of general public who are part of the society. One of the promising application area where Crowdsourcing is prevalent nowadays is Education. The rest of the paper describes how Education can be enhanced using Crowdsourcing.

2. CURRENT EDUCATION SCENARIO

2.1. Issues in current offline learning

In typical offline classroom scenario, one instructor manages around 50 – 70 students at a time and the communication is unidirectional primarily. The issue is that the instructor may not be able to focus on each and every student individually; and hence cannot modify his teaching approach accordingly. Current e-Learning tools support typical classroom strength not exceeding a hundred students. They also lack support for informal learning (24 X 7 learning) in which users are motivated to use community blogs, social networks, discussion forums and so on¹².

2.2. MOOCs: Good and Bad

Education is an evolutionary process which always keeps improving with introduction of new tools and technologies¹³. “From an educational perspective, MOOCs are positioned on the border between formal and informal learning¹⁴”. According to¹⁵, “the learning cycle is an ongoing process that is designed to improve the quality of, as well as collaboration among learners.”

With the advent of MOOCs, it has become possible to track the activity of every individual and based on that, courses can be recommended to him as per his interest. MOOCs provide great platform to under-privileged populations to get the opportunity to experience a unique educational environment¹⁶ from reputed universities. At the same time, due to unbounded volume of MOOC participants¹⁷, it becomes infeasible for the instructors to evaluate every participant in given time, even with the help of TAs¹⁸. Moreover, online evaluation systems like *Moodle* may not always be suitable except for MCQ based quizzes. Other assignments demanding textual or long answers may need use of NLP and machine learning algorithms for automatic evaluation⁴, which is difficult to put into practice. This is where human intelligence can be adopted for best results¹⁸. According to Mr Miller, an MIT associate professor, “The human review is essential, because people can detect things that computers can’t, like hidden bugs or poor design¹⁹.” As stated in²⁰, “Human grading more easily provides context-appropriate responses and better handles ill-specified constraints”.

Dropout rates are much higher in MOOCs than in physical education as students may not get adequate time to submit assignments within deadline as most of them are not full time students usually and not serious about successful course completion. It is reported in¹⁵ and²¹ that average completion rate in MOOCs is around 10%, which is very poor. This issue questions the sustainability of MOOCs due to trade-off between return of investment(ROI) and matter of prestige for universities²². Solutions have been proposed in literature to reduce the dropout rates. The authors in¹² suggest to integrate online courses with a student’s social network. Informal face-to-face meetings at different locations at regular interval can also be a good solution to stop students from dropping out of the course¹⁴. They propose a tool etiquetAR, which enables teachers to structure the discussion and issues can be resolved there. *Piazza* is another such tool which is widely used among many universities across the globe. Some universities have started offering MOOC based special audit courses to encourage students for participation, and rewarding them with certificates on completing the course successfully.

3. IMPACT OF CROWDSOURCING IN MOOCS

This section discusses role of Crowdsourcing and its impact in various critical tasks carried out during typical MOOC offerings.

3.1. Peer Assessment

Crowdsourcing has gained enormous popularity in applications ranging from rating and categorizing online contents as well as judgment in reality shows^(23,9). Crowdsourcing provides several benefits to deal with issues such as assignment evaluation in MOOC scenario. The authors in ²⁰, ²⁴ and ²⁵ discuss the crowdsourcing model for peer assessment for improving assessment methodology. According to (24), “peer assessment involving evaluators from differing culture, demographics and expertise level enable the assessment evolve from an evaluative act to a learning act.”

Table 1
Tools assisting in improving MOOC activities

<i>Tool</i>	<i>Major Functionality</i>
PEAR ⁵	Peer Evaluation
Spark PLUS ⁶	Peer Evaluation
Piazza	Discussion Forum
Pick-A-Crowd ²⁶	Assign tasks to Crowd based on Push Methodology
vidWiki ²⁷	On-screen annotation of educational videos to improve their visual quality and localize their language
etiquetAR ¹⁴	Discussion Forum
Moodle	Learning Management System

3.1.1. Issues and Challenges

Peer assessment promotes anonymity to avoid bias in evaluation⁴, still it faces some issues as mentioned below:

1. How to determine contribution given by each member in case of an assignment where forming group of students (team) is permitted? It may happen that a student putting maximum efforts in a group assignment and one with almost no effort, get the same points for that assignment. An ideal evaluation system should be able to distinguish between students based on their performance.
2. Free riders and plagiarizers receive grades which they do not deserve, and deserving hard workers may get less grades ⁴, because in anonymous peer assessment, the assessor has no knowledge about identity or background of the participant.
3. Moreover, as reported in ⁴, “Diversity in intelligence and domain knowledge, among participants, creates an unavoidable problem of ensuring fairness in evaluation.” However, an effort has been made to address this issue in ²⁸ by proposing development of effective and standard rubrics and assignments which can be easily interpreted by participants.
4. Is the degree of agreement between the assessments provided by a domain expert (instructor or TAs) and crowdsourced workers (fellow participants) acceptable, or there is a major difference between scores given by them to the assignments ²⁹?

5. An interesting issue has been highlighted in ³⁰ as follows. If a single assignment is rated by multiple assessors, then how to aggregate the score given by them? It is highly likely that sincere users give accurate score, whereas non-serious users may not be honest to evaluate the task earnestly, and are in a hurry to finish the task anyhow²¹. There is no mechanism for the requesters to directly observe the effort put in by the crowd workers ¹. As indicated in ³¹ and ¹¹, output provided by crowd is prone to noise and ambiguity, hence there should be a mechanism to properly aggregate the scores received by each crowd worker for a single task or assignment.
6. It is contradictory that how a peer who is also studying the same course and possibly has obtained the same knowledge can assess your assignment. Hew et. al.²¹ refer to this phenomenon as a “blind leading blind” approach.

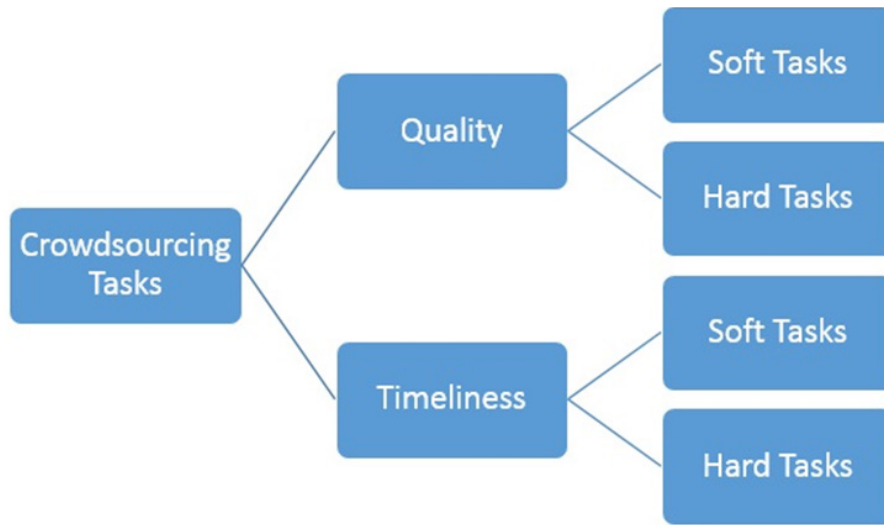


Figure 2: Critical Parameters for Crowdsourcing Tasks Solutions

Figure 2 shows the crucial parameters important for Crowdsourcing problem solutions. Quality of the task and timeliness to perform it are two major concerns in Crowdsourcing problems ^(1,32). Soft tasks are those which may allow a little compromise in terms of quality and time; whereas hard tasks must be strictly accomplished in a given deadline with utmost quality. JSMAPReduce ⁽³³⁾ is an implementation of MapReduce, which enables data parallelism and can be easily migrated with today’s e-learning platforms to achieve speed. Table 1 lists some of the tools extensively used in Education domain.

3.1.2. Solutions

Some attempts have been reported in literature to tackle such issues. In this section, a few existing solutions are discussed.

3.1.2.1. Reputation Scores

In order to resolve the issue of satisfactory ratings given by the scorer as discussed earlier, assigning reputation score to individual workers or a subset of crowds was proposed in ¹. Here, workers whose contributions are accurate, are assigned a high reputation score. Ill-minded responders are assigned low score so as to differentiate them from the genuine responders.

3.1.2.2. Pull vs Push Methodology²⁶

It has also been observed that if a peer evaluator is interested in a specific subset of a larger domain, then he would be more competent at solving assignments pertaining to that domain. Hence it is always a good choice to assign domain relevant tasks to a worker to achieve good quality results. The authors in ²⁶ propose a novel approach, Pick-A-Crowd, which captures user profile from social networks and assigns tasks matching their interest. The relationship between worker characteristics and quality of their work has been investigated in ⁸. In ⁴ also, it is suggested that users, when assigned tasks based on their priority or interest, perform better.

Peer assessment helps students to observe their work from perspective of their fellow participants and they are exposed to solutions and strategies addressed by others ²⁰. This further helps enhance in learning for students as they easily come to know where they need to improve. Crowd can also be used for testing plagiarism in assignments.

3.2. Collaborative Learning through Discussion Forums

In MOOC or online courses, discussion forum plays an important role since it is the major medium of communication between all the participants. Forums can be effectively used for discussing doubts, queries, feedback or sharing any significant information about the course. Most of the MOOC platforms including *Coursera* and *edX* extensively use discussion forums for constantly receiving feedback and other important notifications from their participants.

The authors in ¹⁶ categorize *speech acts* into seven classes such as question, answer, issue, issue resolution, positive or negative acknowledgment and others. These classes help the instructor understand the sentiments among the participants during the tenure of the course.

In an interesting research by ¹⁷, it has been revealed that the *superposters* in discussion forums tend to obtain above average grades than the average forum posters. Moreover, it is also found that superposters in one course, are more likely to become superposters in other courses as well. Hence, superposters can become model participants for a course discussion forum in a sense that they continuously keep on contributing high quality posts.

4. CHALLENGES AND OPEN QUESTIONS

According to Bill Gates, quoted in ¹⁵, currently the emphasis is on using MOOCs for flipping the classrooms, and this has a mixed response among the teacher community. One group of *supporters* think that it supplements the traditional teaching and improves their knowledge as well as the pedagogical skills over the time. On the other hand, there are some *protestors*, who fear that the concept of flipped classrooms will eventually diminish their role as a conventional teacher²². Moreover, MOOCs may harm less prestigious institutes²¹ due to their inability of offering such courses.

Even after promising growth of MOOCs, there exist some open questions:

1. How to measure the success of a MOOC? Is it from number of initial enrollments, or completion rate, or success rate, or diversity of enrollments, or something else?
2. Does a MOOC with rich contents in terms of discussion forums tend towards the success?
3. How can MOOCs be successful in convincing people that they are not against conventional teaching, but are to enhance it?
4. How to deal with issues such as accountability of students in clearing MOOCs by cheating and fraudulent practices? The authors in ²¹ discuss some strategies to deal with these issues but this is still an open problem which needs to be addressed at the earliest.

5. CONCLUDING REMARKS

Participating in peer assessment and discussion forums enable the users to grow as writers²⁸. This, in turn, helps students to improve their comprehension skills and inspires them to work collaboratively. In order to reduce MOOC dropout rates, students should be offered appropriate MOOCs in parallel with their regular credit courses. For example, *Data Mining* or *Data Analytics* course students can opt for MOOCs on *R Programming* and so on.

Finally, this study shows that MOOCs supplemented with Crowdsourcing are doing better and it is expected that in recent future, MOOCs will achieve huge success.

REFERENCES

- [1] Kosinski M, Bachrach Y, Kasneci G, Van-Gael J, Graepel T. Crowd IQ: Measuring the Intelligence of Crowdsourcing Platforms. In: Proceedings of the 4th Annual ACM Web Science Conference. WebSci '12. New York, NY, USA: ACM; 2012. p. 151–160.
- [2] Lukyanenko R, Parsons J. Conceptual Modeling Principles for Crowdsourcing. In: Proceedings of the 1st International Workshop on Multimodal Crowd Sensing. CrowdSens '12. New York, NY, USA: ACM; 2012. p. 3–6.
- [3] Parameswaran AG, Park H, Garcia-Molina H, Polyzotis N, Widom J. Deco: Declarative Crowdsourcing. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management. CIKM '12. New York, NY, USA: ACM; 2012. p. 1203–1212.
- [4] Agrawal R. Crowdsourcing for Educational Assessment in Large Classes. IIT Bombay; 2013.
- [5] Howe J. The Rise of Crowdsourcing. Wired Magazine. 2006 06; 14⁶.
- [6] Eickhoff C, Harris CG, de Vries AP, Srinivasan P. Quality through Flow and Immersion: Gamifying Crowdsourced Relevance Assessments. In: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '12. New York, NY, USA: ACM; 2012. p. 871–880.
- [7] Chatzimilioudis G, Zeinalipour-Yazti D. Crowdsourcing for Mobile Data Management. In: Proceedings of the 2013 IEEE 14th International Conference on Mobile Data Management - Volume 02. MDM '13. Washington, DC, USA: IEEE Computer Society; 2013. p. 3–4.
- [8] Kazai G, Kamps J, Milic-Frayling N. The Face of Quality in Crowdsourcing Relevance Labels: Demographics, Personality and Labeling Accuracy. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management. CIKM '12. New York, NY, USA: ACM; 2012. p. 2583–2586.
- [9] Gomes C, Schneider DS, Moraes K, de Souza JM. Crowdsourcing for Music: Survey and Taxonomy. In: SMC. IEEE; 2012. p. 832–839.
- [10] Väättäjä H, Vainio T, Sirkkunen E, Salo K. Crowdsourced news reporting: supporting news content creation with mobile phones. In: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services 2011 Aug 30 (pp. 435-444). ACM.
- [11] Bott M, Young G. The Role of Crowdsourcing for Better Governance in International Development. The Fletcher Journal of Human Security. 2012; 27:47–70.
- [12] Barbosa CE, Epelbaum VJ, Antelio M, Oliveira J, de Souza JM. Crowdsourcing Environments in E-Learning Scenario: A Classification Based on Educational and Collaboration Criteria. In: SMC. IEEE; 2013. p. 687–692.
- [13] Smith B, Eng M. MOOCs: A Learning Journey - Two Continuing Education Practitioners Investigate and Compare cMOOC and xMOOC Learning Models and Experiences. In: Cheung SKS, Fong J, Fong W, Wang FL, Kwok LF, editors. ICHL. vol. 8038 of Lecture Notes in Computer Science. Springer; 2013. p. 244–255.
- [14] Alario-Hoyos C, Prez-Sanagustn M, Kloos CD. Supporting Crowdsourcing in MOOC informal face-to-face meetings. In: Garca-Pealvo FJ, GonzlezMnC, Griffiths D, editors. WEILER@EC-TEL. vol. 1039 of CEUR Workshop Proceedings. CEUR-WS.org; 2013. p. 6–12.

- [15] Brahim T, Sarirete A. Learning outside the classroom through MOOCs. *Computers in Human Behavior*. 2015 Oct 31;51:604-9.
- [16] Arguello J, Shaffer K. Predicting Speech Acts in MOOC Forum Posts. In: *Association for the Advancement of Artificial Intelligence*; 2015.
- [17] Huang J, Dasgupta A, Ghosh A, Manning J, Sanders M. Superposter Behavior in MOOC Forums. In: *Proceedings of the First ACM Conference on Learning @ Scale Conference. L@S '14*. New York, NY, USA: ACM; 2014. p. 117–126.
- [18] Weld DS, Adar E, Chilton L, Hoffmann R, Horvitz E, Koch M, et al. Personalized online education - a crowdsourcing challenge. In: *In Workshops at the TwentySixth AAAI Conference on Artificial Intelligence*; 2012. p. 159–163.
- [19] Parry M. 5 Ways That edX Could Change Education; 2012. online.
- [20] Kulkarni C, Wei KP, Le H, Chia D, Papadopoulos K, Cheng J, et al. Peer and Self-Assessment in Massive Online Classes. *ACM Trans Comput-Hum Interact*. 2013 Dec;20⁶:33:1–33:31.
- [21] Hew KF, Cheung WS. Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*. 2014;12(0):45 – 58.
- [22] White B. Is “MOOC-Mania” over? In: Cheung SS, Fong J, Zhang J, Kwan R, Kwok L, editors. *Hybrid Learning. Theory and Practice*. vol. 8595 of *Lecture Notes in Computer Science*. Springer International Publishing; 2014. p. 11–15.
- [23] Dasgupta A, Ghosh A. Crowdsourced Judgement Elicitation with Endogenous Proficiency. In: *Proceedings of the 22Nd International Conference on World Wide Web. WWW '13*. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee; 2013. p. 319–330.
- [24] Clougherty Jr RJ, Popova V. Crowdsourcing to assess MOOCs: A position paper. In *MOOCs Forum 2013 Sep 1 (Vol. 1, No. P, pp. 10-13)*. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: Mary Ann Liebert, Inc.
- [25] Ora VK. Peer Evaluation through Crowdsourcing (Doctoral dissertation, Indian Institute of Technology, Bombay).
- [26] Difallah DE, Demartini G, Cudre-Mauroux P. Pick-a-crowd: Tell Me What You Like, and I'll Tell You What to Do. In: *Proceedings of the 22Nd International Conference on World Wide Web. WWW '13*. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee; 2013. p. 367–374.
- [27] Cross A, Bayyapureddy M, Ravindran D, Cutrell E, Thies W. VidWiki: Enabling the Crowd to improve the legibility of Online Educational Videos. In: Fussell SR, Lutters WG, Morris MR, Reddy M, editors. *CSCW*. ACM; 2014. p. 1167–1175.
- [28] Comer D. MOOCs Offer Students Opportunity to Grow as Writers.
- [29] Clough PD, Sanderson M, Tang J, Gollins T, Warner A. Examining the Limits of Crowdsourcing for Relevance Assessment. *IEEE Internet Computing*. 2013;17⁴:32–38.
- [30] Dalvi N, Dasgupta A, Kumar R, Rastogi V. Aggregating Crowdsourced Binary Ratings. In: *Proceedings of the 22Nd International Conference on World Wide Web. WWW '13*. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee; 2013. p. 285–294.
- [31] Dow S, Gerber E, Wong A. A Pilot Study of Using Crowds in the Classroom. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13*. New York, NY, USA: ACM; 2013. p. 227–236.
- [32] Lofi C, Selke J, Balke WT. Information Extraction Meets Crowdsourcing: A Promising Couple. *Datenbank-Spektrum*. 2012;12²:109–120.
- [33] Langhans P, Wieser C, Bry F. Crowdsourcing MapReduce: JSMapReduce. In: *Proceedings of the 22Nd International Conference on World Wide Web Companion. WWW '13 Companion*. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee; 2013. p. 253– 256.