Keynote: From group collaboration to large scale social collaboration
François Charoy

To cite this version:
François Charoy. Keynote: From group collaboration to large scale social collaboration. 25th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE-2016), Jun 2016, Paris, France. hal-01342751

HAL Id: hal-01342751
https://inria.hal.science/hal-01342751
Submitted on 6 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
From group collaboration to large scale social collaboration

François Charoy - @charoy
Inria/University of Lorraine
Coast Team

Inria/LORIA research team (15/20 people)
University of Lorraine
Research topic: Web Scale trustworthy collaborative systems

teams.inria.fr/coast
@InriaLORIACoast
Outline

Origin and (fly over) state of art of Computer Supported Collaboration

Large scale collaboration (does it exists ?)

Perspectives and research issues
Dimensions of collaboration

Two or more people working together toward a shared goal

Communication
Coordination
Sharing
Implicit vs Explicit Coordination
Mother of all demos – 12/9/1968
Doug Engelbart Augmentation Project

https://archive.org/details/dougengelbartarchives
Clarence Ellis 5/1943 – 5/2014

- Pionneer in collaboration software and workflow systems
## Taxonomy of Groupware

<table>
<thead>
<tr>
<th>Same Place</th>
<th>Different Places</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Same Time</strong></td>
<td><strong>Different Times</strong></td>
</tr>
<tr>
<td>face-to-face interaction</td>
<td>asynchronous interaction</td>
</tr>
<tr>
<td>synchronous distributed interaction</td>
<td>asynchronous distributed interaction</td>
</tr>
</tbody>
</table>
Office Automation (1979)

- SCOOP (Wharton and Sisman)
- OfficeTalk (Skip Ellis – Xerox Park)
Challenges at the time

- A programming language that can be used by a Clerical worker
- Software engineering – modeling and validation
- Operating Systems and database – distribution of data, consistency and security
- Measurement and Evaluation
- Communication - teleconferencing
- Other area – psychology, social implications
Collaborative Editor

- GROVE Group Editor

users ask “Isn’t it chaotic to all edit in the same document, even the same paragraph, at the same time?” and “Why would a group ever want to edit in the same line of text at the same time?” Indeed, this editor is at
Design issues (at the time)

- WYSIWIS (What you see is what I see)
  - Relaxed vs strict
- Group Processes (Organising the collaboration)
- Concurrency Control (Optimist vs Pessimistic)
- Over issues
BPM systems

- Model and automate office processes
- Integrate Information Systems
Collaborative BPM
Social BPM

Support collaboration between people and organisations

More flexible/supports change

Allow more interactions between participant

Support for inter organisational process
BonitaBPM V6 (2015)
Inter organisational processes

The view-based approach to dynamic inter-organizational workflow cooperation

Issam Chebbi \textsuperscript{a}, Schahram Dustdar \textsuperscript{b,*}, Samir Tata \textsuperscript{a}

Fig. 4. Public vs. internal processes.
Coordination tool

Trello used by OpenShift
Data sharing and collaborative editing

- Version and Distributed version control Systems
- Wiki
- Synchronous collaborative editors
- File Sharing
Version Control Systems and File sharing

- CVS, SVN, Git, …
Workspace Awareness

• What are the other doing

Figure 1. GroupSketchpad, a real-time collaboration tool for large-scale social collaboration

A Descriptive Framework of Workspace Awareness for Real-Time Groupware

CARL GUTWIN & SAUL GREENBERG
Example: the Radar view

- Where are the other – what are they doing

*Figure 4.* Three versions of the GroupKit radar view. Version 4a shows object movement only; 4b adds location information by showing each person’s main view as a shaded rectangle; 4c adds photographs for participant identification.
Real time collaborative editing

Maintaining replicas in real time
Ensure eventual consistency
Take care of user intention

Google Doc/EtherPad
Writing workshop on Framapad

https://vimeo.com/117890611
Operational Transformation

Enforce eventual consistency
Delays in GoogleDocs

Typing speed:
2 chars/sec
Conflicts management

How to avoid them when the number of users grows
What about today’s topic
About Large SCALE

- Hundreds or thousands of people
  - Crowds, communities, social networks
- Hundreds of organisations
- Thousands of services
users ask “Isn’t it chaotic to all edit in the same document, even the same paragraph, at the same time?” and “Why would a group ever want to edit in the same line of text at the same time?” Indeed, this editor is at
Pad strasbourg nuit debout

More than 70 participants

https://strasbourg.nuitdebout.fr/pad/p/Actions/timeslider#23979
Real Time Wikipedia
Wikipedia edit conflicts

900 edit/mn

5 failure conflict/mn

https://grafana.wikimedia.org/dashboard/db/edit-count
Centralized architecture?

- Does it scale?
- Who controls the server?
- Who has access?
Abstract/Configurable p2p architecture

- Each partner decides of the control and of the contracts
MUTE v2 – No central server
The Mute P2P editor

- Based on CRDT Algorithm (LogootSplit)

Delays in MUTE

2 car./sec./utilisateurs
1 à 40 utilisateurs
Who is doing what?

- Awareness at a large scale
A complex challenge: security and trust

- Who are the people I’m working with
- Collaboration data can be encrypted and not shared on the cloud (being investigated).
- Provides guarantees to the participants
Coordination at a large scale
Haiti Earthquake 2006

- 7.0 Magnitude earthquake – 3.5M people affected – 220,000 death (est.)
- Hundreds of responding organisations
Unexpected event at an unexpected scale

- First responder: people
- Overwhelming answer from the international community (hundreds of NGO’s)
- First important use of social network
- Volunteer based crisis mapping
Hurricane Katrina 2005

- Category 5 Hurricane
- Caused breaches to flood protections structures
- 80% of the city of New Orleans Submerged (1500 deaths)
Coordination issues

- Very wide area affected
- Several political decisions level
- Coordination issues during the evacuation
The Danube floods

- Very common event – cross international borders
- Requires a dedicated organisation
Coordination at a large scale

- Combination of human and services activities
- Composition of hundred of services
- Duration, kind of contribution, governance
- Various domains: logistic, crisis management, software development,
“we are exchanging text information real-time through either cell or mobile units or satellite”

[...] “obvious the more complex a situation gets, for instance where you may have many organizations working collectively in order to accomplish a goal, but at the same time working independently in terms of their own interests, where you get in those in situation like that, you know it get to the point, to be so complex that it would be overwhelming”
Complex BP Modeling

- Outcome of a workshop with Firefighters
Coordination awareness

- Activity awareness?
- Coordination without a central authority
- Distribution of control
- Privacy

Still an issue today
Crowdsourcing

Coordinating very large scale/long running simple activities
Combining human task and service in a process at a large scale
Centralised process with very distributed unknown performers
Haiti Crisis Mapping

4636 project volunteers translated 25,186 SMS messages and numerous e-mail, Web, and social media communications, resulting in 3,596 reports that were actionable and included enough relevant information to be mapped on Ushahidi.\(^\text{31}\)

*The Fletcher Crisis Mappers at work in Boston. Photo courtesy of Carol Waters.*
Sandy Damage Assessment (2012)

- Crowdsourcing satellite pictures
  - Results with short term validity
Industrialisation of crowdsourcing

BPMN Task Instance Streaming for Efficient Micro-Task Crowdsourcing Processes

Stefano Tranquillini¹, Florian Daniel¹,², Pavel Kucherbaev¹, and Fabio Casati¹
BPM for crowdsourcing
Large scale real time coordination

Unsolicited individual data

Collaborative data

Twitter – 35000 geolocalised tweet/day
Flickr – 3000 geolocalised photos/day
10 photos/s on Instagram during sandy

Wikipedia 1500 edit to the main page between the 28 and the 30th of October, not mentioning the talk or the area related pages

Citizen Sourcing Data

Dedicated app can receive notification and but used to ask for data

FEMA dedicated app

INPUTS

Process

Computers
Experts
Crowd

WETICE 2016 - From group collaboration to large scale social collaboration
Human/Machine composition example

Cluster post and tweets

Paid crowd

No decision

Volunteer crowd

Specialised app (FEMA Disaster Reporter)

Call the walking crowd

Experts

Missing data?

WETICE 2016 - From group collaboration to large scale social collaboration
The Refugee crisis

- A very tricky kind of crisis
- Affect international relationships
- Difficult agreement on the response
Citizen coordination

- International scale
- Citizen create Facebook pages, maps, dedicated tools to coordinate
- But privacy, security, reliability, trust issues
Internet wide service composition
Behind the scene

> 1000 renting web sites

> 100000 rooms

confidence is required

> 100 pays
Conclusion

- Large scale collaboration exists
- Large scale collaboration helps to solve large scale problems
- Only very solution based ad-hoc support
Large scale collaboration issues

- A programming language that can be used by a Clerical worker
- Software engineering – *modeling* and *validation*
- Operating Systems and database – *distribution of data, consistency and security*
- Measurement and Evaluation - *prediction*
- Security, *privacy and trust*
- WYSIWweS or What ou see is what we see
- Community Processes
- Concurrency Control
- Other area – psychology, social implications
QUESTIONS ?
Credits

- Some slides have been prepared by
  - Jorn Franke (Activity management)
  - Gérald Oster (P2P sharing)

Most of the work presented here are contribution made by members of the Coast team and the preceeding ones

https://team.inria.fr/coast/

Original photos are from Christian Morel (DR)