

CoDaMine: Effiziente soziale Interaktion durch die Analyse der Online-Kommunikation im Instant Text Messaging

Tom Gross, Mirko Fetter, Julian Seifert

Fakultät Medien, Bauhaus-Universität Weimar, Deutschland

Zusammenfassung

Die heutigen Medienlandschaften bieten eine breite Palette von Diensten zur sozialen Interaktion der Benutzerinnen und Benutzer. Daraus ergeben sich neue Möglichkeiten zur spontanen Erreichbarkeit und Kommunikation. Allerdings werden die Anwendungen den sozialen und emotionalen Anforderungen nur teilweise gerecht. In diesem Artikel wird mit CoDaMine ein innovativer Ansatz zur effizienten sozialen Interaktion vorgestellt.

1 Einleitung

Die heutigen Medienlandschaften bieten eine breite Palette von Möglichkeiten zur *sozialen Interaktion* der Benutzerinnen und Benutzer. Neben dem bereits seit längerem bekannten und verwendeten Email zum komfortablen Versenden von Textnachrichten hat sich über die letzten Jahre Instant Messaging zum spontanen Führen von Online-Konversationen via Text, Audio und Video etabliert. Während zunächst primär ICQ (ICQ Inc., 2010) zum Verwalten sowie zur wechselseitigen Information über die Präsenz von Online-Kontakten verwendet wurde, wird heute Skype (Skype, 2010) insbesondere wegen der guten Qualität bei der Audio-Kommunikation über Voice-over-IP eingesetzt. Aber auch neuere Netzwerktools zur Verwaltung und Pflege von sozialen Netzwerken wie beispielsweise Facebook (Facebook, 2010) bieten heute oft integrierte Möglichkeiten zur wechselseitigen Online-Information und spontanen Text-Konversation. Diese technologischen Entwicklungen bringen große Vorteile bezüglich der spontanen Erreichbarkeit und Kommunikation mit sich (Gross & Koch, 2007).

Neben den arbeitsorganisatorischen Vorteilen der spontane Erreichbarkeit und Kommunikation gibt es wichtige *emotionale Aspekte*, die bei der Gestaltung von Technologie zur sozialen Interaktion eine Rolle spielen. Bereits in sehr frühen Analysen des Benutzerverhaltens in Instant Messaging Anwendungen wurde deutlich, dass durch diese Anwendungen die emotionale Bindung der Benutzerinnen und Benutzer untereinander positiv beeinflusst werden kann (Nardi et al., 2000). Darüber hinaus wurde festgestellt, dass es auch eine dahingehende Erwartungshaltung seitens der Benutzerinnen und Benutzer gibt; so schreiben Reeves und Nass „...we have found that individuals' interactions with computers, television, and new media are fundamentally social and natural, just like interactions in real life.“ und „everyone expects media to obey a wide range of social and natural rules.“ (Reeves & Nass, 2003, p. 5).

Die heutigen Instant Messaging Anwendungen werden diesen emotionalen Anforderungen nur *teilweise gerecht*. Auf einer syntaktischen Ebene wird Funktionalität angeboten, um die eigene Stimmung zu kommunizieren (z.B.: können über die Mood Messages in Skype kurze Kommentare zur Situation der Benutzerinnen und Benutzer eingegeben werden) und um spezifisch in Online-Konversationen Emotionen zu kommunizieren (z.B.: können Emoticons den intendierten Ton einer Textnachricht unterstreichen). Auf der semantischen Ebene gibt es derzeit nur wenige Möglichkeiten, die Bedeutung und Einordnung von Online-Kontakten und Online-Konversationen insgesamt zu unterstützen.

In diesem Artikel wird das grundlegende Konzept von CoDaMine erläutert, welches einen innovativen Ansatz zur effizienten sozialen Interaktion durch die semantische Einordnung von Online-Kontakten und Online-Konversationen bietet. Darauf aufbauend werden die Vorteile für die Benutzerinnen und Benutzer anhand einer Beschreibung der Benutzerführung in CoDaMine illustriert. Des Weiteren wird die technische Umsetzung von CoDaMine skizziert. Schließlich werden verwandte Arbeiten beschrieben und der Mehrwert von CoDaMine hervorgehoben.

2 CoDaMine Konzept

Das Konzept von CoDaMine (Communication Data Mining) ist ein innovativer Ansatz zur effizienten sozialen Interaktion durch die semantische Einordnung von Online-Kontakten und Online-Konversationen mittels Textanalyse und maschinellem Lernen in Instant Text Messaging.

CoDaMine basiert auf der zentralen Erkenntnis aus früheren Untersuchungen, dass es online zwischen Kontakten inhaltlich und emotional sehr vielschichtige Beziehungen und Konversationen geben kann, (Gross & Prinz, 2004) und dass dies zur effizienten sozialen Interaktion entsprechend unterstützt werden muss. In CoDaMine wurde dazu das Konzept von *Faces* entwickelt, welche auf den Erkenntnissen des Soziologen Erwin Goffman basieren. Goffman hat die soziale Interaktion in realweltlichen Situationen studiert und festgestellt, dass die beobachteten Personen je nach Situation — d.h. abhängig von Anlässen und anwesenden Personen — ein spezifisches Auftreten haben (Goffman, 1959). CoDaMine erlaubt es den Benutzerinnen und Benutzern durch Faces verschiedene Online-Auftritte gegenüber Online-Kontakten und in Online-Konversationen zu verwalten. Sie können festlegen, welche Informationen wie kommuniziert werden sollen. Dies betrifft sowohl die explizite Kommunikation (d.h. Instant Text Messaging), als auch die implizite Kommunikation (d.h. die durch Sensoren erfassten Informationen wie Online-Status, Aufenthaltsort usw., welche anderen Benutzerinnen und Benutzern weitergeleitet werden können).

Um die Vorteile von Faces nutzen zu können, müssen die Benutzerinnen und Benutzer den *Aufwand der Konfiguration* der Faces leisten. Dies betrifft zum einen den Aufwand beim neu einrichten der Anwendung und zum anderen den Aufwand zur Aktualisierung der Konfigurationen. Während der Einrichtungsaufwand einmalig anfällt, entsteht der Aktualisierungsaufwand fortlaufend. Die Benutzerinnen und Benutzer müssen prüfen, ob sie die richtigen Faces haben, ob sie diesen richtigen Kontakte zugewiesen haben und ob die explizite und implizite Kommunikation noch zum jeweiligen Face passt.

CoDaMine zielt daher darauf ab, durch das *FaceReco* Konzept komfortabel Faces aktuell zu halten. CoDaMine macht den Benutzerinnen und Benutzern basierend auf der Analyse ihrer sozialen Interaktion Vorschläge bezüglich der Verwaltung von Online-Kontakten. Dazu wird das Auftreten von sprachlichen Merkmalen in Online-Konversationen gemessen und als Merkmal für die Charakterisierung der Verbindung zweier Kontakte aufgefasst.

Der in CoDaMine verwendete Ansatz stützt sich auf Erkenntnisse der Soziolinguistik — insbesondere auf *Speech Communities* und *Register*. *Speech Communities* sind Gruppen variabler Größe, die sich durch gemeinsame sprachliche Merkmale auszeichnen (Fitzpatrick et al., 2002). Auch *Register* gehören dazu, wobei diese Sprachgruppen beschreiben, die auf in der Regel auf gemeinsamen Aktivitäten basieren. Auf Basis dieser sprachlichen Merkmale gelingt es, Vorschläge für Aktualisierung der Einstellungen von Faces zu generieren.

Die Konzepte Faces and FaceReco sollen anhand eines Szenarios aus der Sicht der Benutzerinnen und Benutzer illustriert werden. Ein Benutzer Klaus verwendet CoDaMine als Instant Text Messaging Anwendung für Online-Konversationen und die Verwaltung von Online-Kontakten. Er legt verschiedene Faces an und fügt diesen entsprechend Online-Kontakte hinzu. Darüber hinaus legt Klaus fest, welche Informationen den Online-Kontakten eines Faces bereitgestellt werden. Dazu fügt er den Faces Sensoren hinzu, welche selektiv Informationen über ihn erfassen und vertraulich weiterleiten können. Klaus beginnt mit seinen Online-Kontakten zu kommunizieren. Er unterhält sich mit den Kommunikationspartnern in der Regel über Sachverhalte, die dem Kontext des jeweiligen Faces entsprechen, welchem der Gesprächspartner zugeordnet ist. Klaus unterhält sich häufig mit seinem Kollegen Julian. Julian ist neu im Projekt CoDaMine, für welches Klaus bereits ein Face angelegt hat. Nachdem Klaus und Julian sich eine Zeit lang über CoDaMine unterhalten haben, erkennt das System, dass Klaus Julian in das Face CoDaMine aufnehmen könnte und schlägt dies vor. Klaus nimmt den Vorschlag an. Somit erhält der Kommunikationspartner Julian nun die Online-Kommunikation im richtigen Kontext sowie auch die Sensor-Informationen, die Klaus über dieses Face anbietet wie beispielsweise wichtige das Projekt betreffende Termine, seine Verfügbarkeit für die Projektmitglieder.

3 CoDaMine Benutzerführung

Die Benutzerführung in CoDaMine ist so gestaltet, dass die Benutzerinnen und Benutzer komplexe soziale Verflechtungen abbilden können, ohne dadurch großen Konfigurationsaufwand in Kauf nehmen zu müssen.

Beim *ersten Login* in die Instant Messaging Umgebung CoDaMine, erhalten die Benutzerinnen und Benutzer ein *Default-Face*. Dieses enthält zum Schutz der Privatsphäre zunächst weder Online-Kontakte noch Sensoren. Dieses Default-Face ist standardmäßig stets verfügbar ist und nicht löscherbar. Es lässt sich jedoch frei konfigurieren.

Die Benutzerinnen und Benutzer können diesem Face beliebige Online-Kontakte und Sensoren hinzufügen (siehe Abbildung 1 mit den Knöpfen Add a new Contact und Add a new Source) und wieder entfernen (mit den Knöpfen Delete Contact und Delete Source) bzw. zwischen Faces verschoben (mit den Knöpfen Move Contact) werden. Des Weiteren können Faces verwaltet werden (mit den Knöpfen Add new Face und Delete Face).

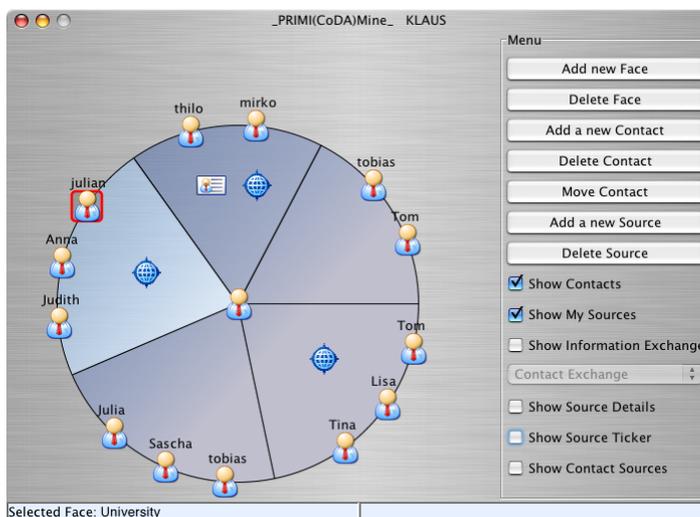


Abbildung 1. CoDaMine Hauptfenster mit dem Default-Face sowie vier weiteren benutzerdefinierten Faces. Allen Faces sind Online-Kontakte zugeordnet (durch Icons am Kreisrand dargestellt) und einigen Faces sind Sensoren zugeordnet (durch Icons im Kreisinneren dargestellt). Alle Kontakte eines Faces erhalten alle Informationen der Sensoren dieses Faces.

Wird eine *Online-Konversation* mit einem Online-Kontakt begonnen, welcher in mehreren Faces enthalten ist, dann sind die Benutzerinnen und Benutzer angehalten nach öffnen des Chat-Fensters aus einer Auswahlliste das Face auszuwählen, welchem das beabsichtigte Gespräch zugeordnet werden soll (siehe Abbildung 2; möchte der Benutzer beispielsweise mit einem Online-Kontakt im Kontext des Faces Sports über gemeinsame sportliche Aktivitäten diskutieren, so wäre das Face Sports auszuwählen).

Die Information, welches Face einer Konversation mit einem Online-Kontakt zugeordnet ist, wird von CoDaMine gespeichert und später verwendet, um einen Classifier als Basis für spätere Empfehlungen zu trainieren.

Initiiert ein Online-Kontakt eine Unterhaltung, dann ist die Information, welchem Face die Benutzerinnen und Benutzer dieser Unterhaltung zuordnen möchten, noch nicht vorhanden. Eingehende Nachrichten neuer Konversationen werden daher zunächst dem Default-Face zugeordnet. Erst wenn die Benutzerinnen und Benutzer das entsprechende Face über die Auswahlliste festlegen, werden die ein- und ausgehenden Nachrichten für das Trainieren des Systems verwendet.

Damit CoDaMine in die Lage versetzt wird, gegebenenfalls *Empfehlungen* generieren und präsentieren zu können, müssen die Benutzerinnen und Benutzer das System zunächst trainieren und immer wieder manuell für jede Konversation ein Face auswählen. Nach dieser Trainingsphase ist CoDaMine voll einsatzfähig und es entfällt der manuelle Aufwand der Face-Auswahl.



Abbildung 2. CoDaMine Instant Text Messaging Fenster. Im oberen Textfeld ist vermerkt, aus welchem Face heraus dieses Fenster geöffnet wurde. Im unteren Teil ist die Liste verfügbarer Faces zu sehen. Diese Liste kann zum wechseln des assoziierten Faces, auch während der Konversation, verwendet werden.

Erhält der Benutzer von CoDaMine eine Empfehlung, so wird dies als Vorschlag verstanden, welcher angenommen oder abgelehnt werden kann. Die Empfehlung wird in Form eines Dialogfensters präsentiert (siehe Abbildung 3).

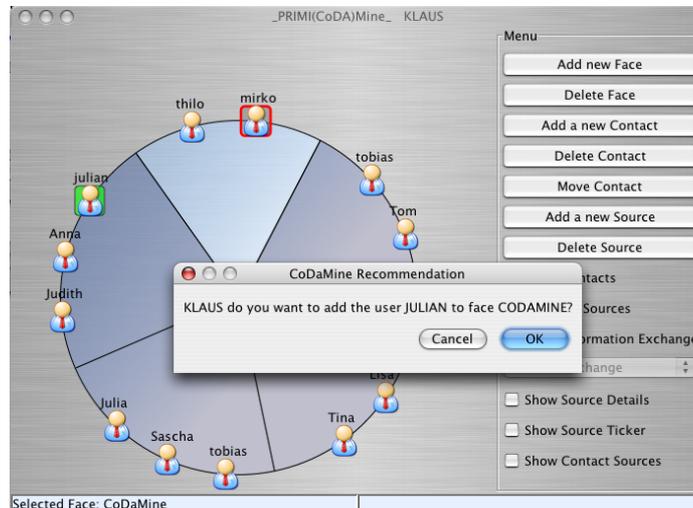


Abbildung 3. CoDaMine Hauptfenster mit Face-Empfehlung. Wird die Empfehlung angenommen, wird der vorgeschlagene Benutzer dem vorgeschlagenen Face hinzugefügt.

Sofern Benutzerinnen und Benutzer Empfehlungen annehmen, wird der jeweilige Online-Kontakt dem entsprechenden Face hinzugefügt oder aus diesem entfernt, je nachdem um was für eine Empfehlung es sich handelt. Wenn die Empfehlung nicht angenommen wird, ändert sich aus Sicht der Benutzerinnen und Benutzer nichts. Systemintern wird die Ablehnung allerdings als Feedback für das Modul verwendet, welches die Empfehlungen generiert und korrigiert dessen Verhalten.

4 CoDaMine Implementierung

CoDaMine basiert auf einer Client-Server-Architektur und besteht aus drei verteilten Hauptkomponenten (siehe Abbildung 4): dem CoDaMine Client, dem CoDaMine Server und einem Wildfire Server. Beim Wildfire Server handelt es sich um einen frei verfügbaren Instant Messaging Server, auf welchen in weiterer Folge nicht näher eingegangen wird (Ignite Realtime, 2006).

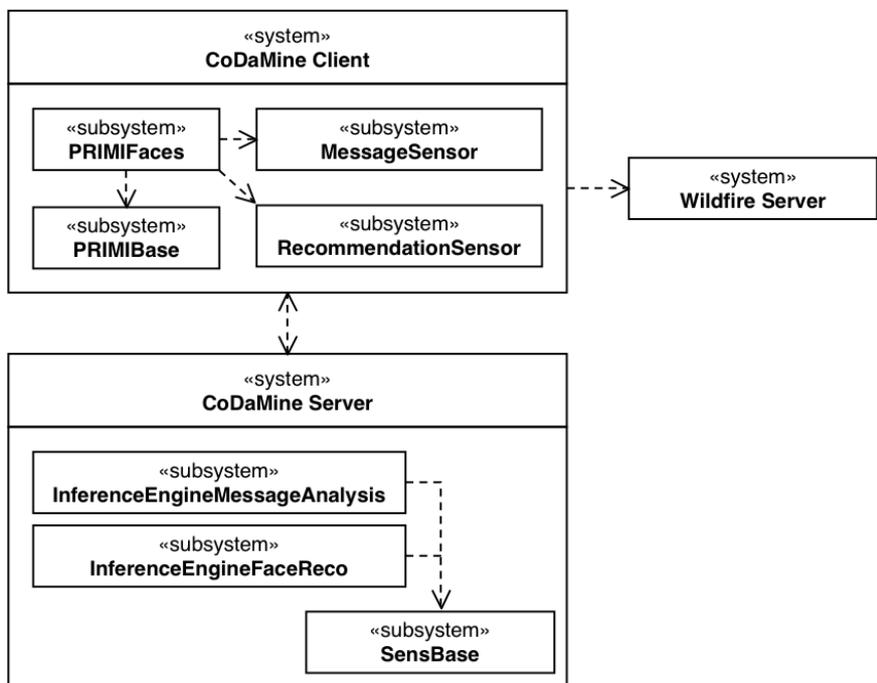


Abbildung 4. CoDaMine Komponentendiagramm.

Der *CoDaMine Client* ist ein Thin Client, der keine Daten lokal speichert und somit auf verschiedenen Rechnern verwendet werden kann. Er ist modular aufgebaut und basiert auf den PRIMIBase (Gross & Oemig, 2005) und PRIMIFaces (Gross & Oemig, 2006) Plattformen. Dies sind Plattformen für die Entwicklung von Instant Messaging Infrastrukturen, die komfortabel über Plugins erweitert werden können. Die Benutzeroberfläche des CoDaMine Clients ist als GUI-Plugin für *PRIMIFaces* implementiert. Die Kommunikation und insbesondere das Senden und Empfangen von Textnachrichten wird über ein Plugin für *PRIMIBase* realisiert. Der *MessageSensor* überträgt die Kommunikationsdaten zur Auswertung an den CoDaMine Server. Der *RecommendationSensor* prüft zyklisch ob neue Empfehlungen generiert wurden und übermittelt diese an das GUI-Plugin für die Benachrichtigung der Benutzerinnen und Benutzer.

Der *CoDaMine Server* basiert auf der SensBase Plattform zur Entwicklung sensorbasierter Infrastrukturen (Gross et al., 2006). SensBase stellt eine Vielzahl von Sensoren zur Verfügung und bietet Möglichkeiten zur Verarbeitung der erfassten Sensordaten. Die beiden Inference Engines stellen Plugins für das maschinelle Lernen dar. Die *InferenceEngineTextAnalysis* macht die Kommunikationsdatenanalyse. Es werden zunächst Basismerkmale wie die Anzahl der Schreibfehler oder die Menge der im Text enthaltenen Emoticons extrahiert. In einem weiteren Schritt wird das Face ermittelt, zu dessen Eigenschaften die aktuell verarbeitete Nachricht am wahrscheinlichsten gehört. Die *InferenceEngineTextAnalysis* wendet hierfür Techniken der Textklassifikation an, wobei die vorhandenen Faces als Klassen dienen. Dies wird mit dem Werkzeugkasten für maschinelles Lernen Weka realisiert (Witten & Frank, 2006). Die *InferenceEngineFaceReco* situation für die Empfehlungsgenerierung verantwortlich. Sie verwendet die erhaltenen Merkmale und berechnet aus

.....

Werten der Merkmale gewichtet Punkte für ein einfaches Scoring System. Für jedes Face und für jeden Online-Kontakt eines Benutzers werden diese Punkte verrechnet. Wird nach dem Eingehen neuer Punkte ein oberer Schwellwert erreicht, wird eine Empfehlung generiert, den entsprechenden Online-Kontakt einem Face hinzu zu fügen. Wird ein unterer Schwellwert erreicht, wird die Empfehlung gegeben, den Online-Kontakt aus dem Face zu entfernen.

Alle CoDaMine Komponenten kommunizieren über ein Netzwerk, wobei der CoDaMine Client zum einen mit dem Wildfire Server über das XMPP (Jabber Inc., 2010) und zum anderen mit dem CoDaMine Server über XML-RPC (UserLand Software Inc., 2010). Weitere technische Details finden sich in (Gross & Fetter, 2008).

5 Verwandte Arbeiten

In den letzten Jahren wurden einige Konzepte entwickelt, die Ähnlichkeiten zu CoDaMine haben. Wir stellen hier kurz ein Beispiel für die Analyse sozialer und zeitlicher Strukturen und ein Beispiel für die Verwaltung von Online-Kontakten vor.

Die Infrastruktur Soylent unterstützt die Exploration von Elementen der sozialen und zeitlichen Struktur von Zusammenarbeit. Soziale Strukturen beschreiben hierbei Muster des Kontaktes und der Zusammenarbeit zwischen Personen, zeitliche Strukturen versuchen zu erfassen wie sich Interaktionsmuster über Zeit entwickeln. Basierend auf Soylent unterstützt die TellMeAbout Anwendung Benutzerinnen und Benutzer bei der Analyse ihrer eigenen Situation (Fisher & Dourish, 2004). Dabei können die Benutzerinnen und Benutzer flexibel Abfragen machen, aber das System generiert keine Empfehlungen.

ContactMap analysiert ebenfalls die Kommunikationshistorie der Benutzerinnen und Benutzer auf Basis der empfangenen und geschriebenen Emails, um daraus ein soziales Netzwerk abzuleiten (Nardi & Whittaker, 2002). Anders als bei Soylent und TellMeAbout steht dabei die Verwaltung der Kontakte in einer Desktopdarstellung im Vordergrund. Im Gegensatz zu CoDaMine bietet ContactMap keine Informationen über die Präsenz und Erreichbarkeit der Online-Kontakte.

6 Zusammenfassung

In diesem Beitrag haben wir CoDaMine vorgestellt. Es bietet innovative Konzepte für das einfache Management von komplexen sozialen Strukturen.

Erste informelle Benutzerstudien mit CoDaMine zeigen große Zufriedenheit bei den Benutzerinnen und Benutzer. Für die Zukunft sind allerdings systematische Evaluationen nötig, um die Effektivität und Effizienz der Konzepte sowie die Zufriedenheit der Benutzerinnen und Benutzer beurteilen zu können.

7 Literaturhinweise

Facebook. (2010). Facebook | Home, <http://www.facebook.com>. (Letzter Zugriff 12/2/2010).

Fisher, D., & Dourish, P. (2004). Social and Temporal Structures in Everyday Collaborations. Proceedings of the Conference on Human Factors in Computing Systems - CHI 2004, Wien, Österreich. pp. 551-558.

Fitzpatrick, G., Kaplan, S., Mansfield, T., Arnold, D., & Segall, B. (2002). Supporting Public Availability and Accessibility with Elvin: Experiences and Reflections. Computer Supported Cooperative Work: The Journal of Collaborative Computing, 11(3-4). pp. 447-474.

Goffman, E. (1959). The Presentation of Self in Everyday Life. N.Y.: Doubleday Anchor Books.

Gross, T., Eglar, T., & Marquardt, N. (2006). Sens-ation: A Service-Oriented Platform for Developing Sensor-Based Infrastructures. International Journal of Internet Protocol Technology (IJIPT), 1(3). pp. 159-167.

- Gross, T., & Fetter, M. (2008). CoDaMine: Communication Data Mining for Feedback and Control in Ubiquitous Environments. Proceedings of the Sixteenth Euromicro Conference on Parallel, Distributed, and Network-Based Processing - PDP 2008, Toulouse, Frankreich. pp. 539-546.
- Gross, T., & Koch, M. (2007). Computer-Supported Cooperative Work (Computer-Supported Cooperative Work; in German). München: Oldenbourg.
- Gross, T., & Oemig, C. (2005). PRIMI: An Open Platform for the Rapid and Easy Development of Instant Messaging Infrastructures. Proceedings of the 31st EUROMICRO Conference on Software Engineering and Advanced Applications - SEAA 2005, Oporto, Portugal. pp. 460-467.
- Gross, T., & Oemig, C. (2006). From PRIMI to PRIMIFaces: Technical Concepts for Selective Information Disclosure. Proceedings of the 32nd EUROMICRO Conference on Software Engineering and Advanced Applications - SEAA 2006, Cavtat, Dubrovnik, Kroatien. pp. 480-487.
- Gross, T., & Prinz, W. (2004). Modelling Shared Contexts in Cooperative Environments: Concept, Implementation, and Evaluation. Computer Supported Cooperative Work: The Journal of Collaborative Computing, 13(3-4). pp. 283-303.
- ICQ Inc. (2010). ICQ.com - Community, People Search, and Messaging Service!, <http://www.icq.com/>. (Letzter Zugriff 12/2/2010).
- Ignite Realtime. (2006). Wildfire, <http://www.jivesoftware.org/wildfire/>: Jive Software Community. (Letzter Zugriff 10/7/2007).
- Jabber Inc. (2010). Jabber: Open Instant Messaging and a Whole Lot More, Powered by XMPP, <http://www.jabber.org/>. (Letzter Zugriff 12/2/2010).
- Nardi, B. A., & Whittaker, S. (2002). Integrating Communication and Information Through ContactMap. Communications of the ACM, 45(4). pp. 89-95.
- Nardi, B. A., Whittaker, S., & Bradner, E. (2000). Interaction and Outeraction: Instant Messaging in Action. Proceedings of the Conference on Computer-Supported Cooperative Work - CSCW 2000, Philadelphia, PE. pp. 79-88.
- Reeves, B., & Nass, C. (2003). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge, MA: Cambridge University Press.
- Skype. (2010). Make the most of Skype - Free Internet Calls and Cheap Calls, <http://www.skype.com/>. (Letzter Zugriff 12/2/2010).
- UserLand Software Inc. (2010). XML-RPC Home Page, <http://www.xmlrpc.com/>. (Letzter Zugriff 12/2/2010).
- Witten, I. H., & Frank, E. (2006). Data Mining: Practical Machine Learning Tools and Techniques. (2. Ed.). San Francisco: Morgan Kaufmann Publishers.