Report from Dagstuhl Seminar 13402

Physical-Cyber-Social Computing

Edited by

Amit P. Sheth[1], Payam Barnaghi[2], Markus Strohmaier[3], Ramesh Jain[4], and Steffen Staab[5]

1 **Wright State University – Dayton, US,** amit@knoesis.org
2 **University of Surrey – Guildford, GB,** p.barnaghi@surrey.ac.uk
3 **Universität Koblenz-Landau, DE,** strohmaier@uni-koblenz.de
4 **Univ. California – Irvine, US,** jain@ics.uci.edu
5 **Universität Koblenz-Landau, DE,** staab@uni-koblenz.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13402 "Physical-Cyber-Social Computing".

Seminar 29. September – 04. October, 2013 – www.dagstuhl.de/13402

1998 ACM Subject Classification C.2 Computer-communication Networks, K.4 Computers and Society, B. Hardware

Keywords and phrases Semantic sensor networks, Web of Things, Internet of Things, Semantic interoperability and context, Social computing, Collective intelligence, Situation-awareness, Human and machine perception, Computing for human experience, Social life networks

Digital Object Identifier 10.4230/DagRep.3.9.245

Edited in cooperation with Rajendra Akerkar (Vestlandsforsking, NO), Josiane Xavier Parreira (National University of Ireland, IE), Pramod Anantharam (Wright State University, US)

1 Executive Summary

Amit P. Sheth
Payam Barnaghi
Markus Strohmaier
Ramesh Jain
Steffen Staab

Miniaturisation, progress with energy issues and cost reductions have resulted in rapid growth in deployment of networked devices and sensing, tightly connecting the physical word with the cyber-world as well as interconnected humans bringing along them virtual social interactions.. The number of devices connected to the Internet already exceeds the number of people on earth and is estimated to grow to 50 billion devices by 2020. The resulting system called Internet of Things (IoT) incorporates a number of technologies including wireless sensor networks, pervasive computing, ambient intelligence, distributed systems and context-aware computing. With growing adoption of smart-phones and social media, citizens or human-in-the-loop sensing and resulting user generated data and data generated by user carried devices have also become key sources of data and information about the physical world and corresponding events. Data from all these sources will result in tremendous volume, large variety and rapid changes (velocity). The combination of cyber-physical and social data can help us to understand events and changes in our surrounding environments better, monitor and control buildings, homes and city infrastructures, provide better healthcare and elderly care services among many other applications. To make efficient use of the physical-cyber-social data, integration and processing of data from various heterogeneous sources is necessary. Providing interoperable information representation and extracting actionable knowledge from deluge of human and machine sensory data are the key issues. We refer to the new computing capabilities needed to exploit all these types of data to enable advanced applications as physical-cyber-social computing.

2 Table of Contents

3 Overview of Talks

3.1 From smart meters to smart behaviour

Harith Alani (The Open University – Milton Keynes, GB)

It is becoming clear to policy makers and environment agencies that changing the behaviour of energy consumer is vital for battling climate change. Much work has being done recently to support behaviour change in this domain, for example with smart meters, refined energy consumption reports, and information campaigns. However, many questions remain unanswered, such as which combination of incentives work best, and for which scenarios and demographic areas, which behaviour change strategies are short/long term, and how much behavioural change can be expect from which strategy. Answering these important and complex questions requires the close collaboration of computer scientists, sociologists, energy meter producers, and environment campaigners, to experiment with, and evaluate, the impact of different approaches in changing the behaviour of energy consumers in households and businesses.

3.2 Beyond Factual Question Answering

Pramod Anantharam (Wright State University – Dayton, US)

Question Answering (QA) systems have been around for decades but the most notable breakthrough in QA systems was achieved by IBM Watson, the machine that defeated Ken Jennings, a champion who had won Jeopardy! 72 times in a row. The Jeopardy! game was broadcasted on national television in February 2012. Such QA systems will alleviate the challenge of decision making in data intensive environments. This type of question answering is called factual question answering as the answers are synthesized by facts in textual documents. While this has been the state-of-the-art in QA systems, Physical-Cyber-Social systems offer a unique set of challenges and this short presentation proposes a provocative idea of building a cognitive system that goes beyond factual QA. By presenting an asthma example, the challenge of combining observations from multimodal and multi-sensory sources is demonstrated.

3.3 Social networks across the digital-physical boundary

Ciro Cattuto (ISI Foundation, IT)

The advances in mobile technologies and wearable sensors allow to quantify human behavior at unprecedented levels of scale and detail. Wearable sensors, in particular, are opening up a new window on social behavior at the finest resolution of individual interactions, impacting diverse research areas such as social network analysis, organizational science and infectious disease dynamics. I will summarize recent efforts on measuring and analyzing social networks from spatial behavior, and highlight important structural properties of the empirical data collected in real-world environments. I will discuss specific challenges in data cleaning, curation and integration of social network data from sensors. Finally, I will reflect on a few outstanding challenges in using these data sources to achieve impact in specific domains such as infectious disease dynamics.

3.4 A few thoughts on engineering social-computational systems

Markus Strohmaier (Universität Koblenz-Landau, DE)

Social computational systems are integrated systems of people and computers. What distinguishes social computational systems from other types of software systems – such as software for cars or air planes – is the unprecedented involvement of data about user behavior, -goals and -motivations into the software system's structure. In social computational systems, the interaction between a user and the system is mediated by the aggregation of explicit or implicit data from other users. This is the case with systems where, for example, user data is used to suggest search terms (e.g. Google Autosuggest), to recommend products (e.g. Amazon recommendations), to aid navigation (e.g. tag-based navigation) or to filter content (e.g. Digg.com). This makes social computational systems a novel class of software systems (as opposed to for example safety-related software that is being used in cars) and unique in a sense that potentially essential system properties and functions – such as navigability – are dynamically in uenced by aggregate user behavior. Such properties can not be satisfi ed through the implementation of requirements alone, what is needed is regulation, i.e. a dynamic integration of users' goals and behavior into the continuous process of engineering.

3.5 Citizen Actuation For Lightweight Energy Management

Edward Curry (National University of Ireland – Galway, IE)

Cyber-Physical Energy Systems (CPES) exploit the potential of information technology to boost energy efficiency while minimizing environmental impacts. CPES can help manage energy more efficiently by providing a functional view of the entire energy system so that energy activities can be understood, changed, and reinvented to better support sustainable practices. CPES can be applied at different scales from Smart Grids and Smart Cities to Smart Enterprises and Smart Buildings. Significant technical challenges exist in terms of information management, leveraging real-time sensor data, coordination of the various stakeholders to optimize energy usage.

In this talk I describe an approach to overcome these challenges by re-using the Web standards to quickly connect the required systems within a CPES. The resulting lightweight architecture leverages Web technologies including Linked Data, the Web of Things, and Social Media. The talk describes the fundamentals of the approach and demonstrates it within an Enterprise Energy Management scenario smart building.

3.6 Principles of Elastic Systems Towards building Cyber-Physical-Social Systems

Schahram Dustdar (TU Wien, AT)

In this talk I present the fundamental models, algorithms and engineering methods and programming abstractions which enable software engineers to model, design, and execute a novel class of software systems: Elastic Systems on a Cloud computing infrastructure. These systems are compositions of the Internet of Things, People, and Software services (including compute, storage, network units).

3.7 Human-Machine Cooperation in Research

Michael Granitzer (Universität Passau, DE)

While Machines process enormous processing capabilities, the lack Humans creativity, intuition and common sense background knowledge. A circumstance unlikely to be changed within the next few years, if ever. In this talk i aim to briefly highlight recent developments in machine learning and outline the need for a tighter integration of machines and humans for upcoming data challenges. Particular emphasis will be placed on data challenges in research and the Giant Global Graph as all encompassing database.

3.8 Understanding and shaping human behavior

Vivek K. Singh (MIT Media Lab, US)

With the growth trends in sensing and information sharing, we can soon expect personal behavioral data (e.g. calls, Bluetooth, heart rate, other quantified-self sensing) to become readily accessible via the ubiquitous Internet-of-Everything. This will allow computational systems to go beyond cyber trails/partial reports of people's actions (e.g. tweets, surveys, or yearly medical checkups) and actually work with real-world signals, coming in real-time, from the real world. Analyzing and utilizing this data is important because: 1) it can answer epistemological questions on human behavior in a data-driven manner, and 2) provide actionable guidelines on how to persuade humans in real world social scenarios. This talk gives a short review of the recent approaches at understanding and shaping human behavior using such data.

3.9 Big Money and NSA^2 – The Future for Physical-Cyber-Social Computing?

Manfred Hauswirth (National University of Ireland – Galway, IE)

It is an established fact that we produce enormous amounts of static and dynamic information. This information is exploited to a certain extent already. Research focused on making this information accessible in a simple fashion (infrastructures), ways of scalable data integration in open environments (Linked Data, ontologies) and putting data to use (analytics, smart cities, etc.). A lot of business opportunities are predicted in this area. However, the systems are not mature enough yet and a lot of research is still required. Additionally, if successful, the flipside of this success will be that we may completely lose any privacy as we can then be monitored comprehensively in the real and in the online worlds. The question is how we can do good research while not making the job of the NSA even easier as it is already. Everyone talks about privacy and actually offer and apply it. In this talk, I will quickly review the existing state of the art, some of the driving requirements (in my opinion) and issues the research community must turn their attention to (again my personal opinion).

3.10 Social and Sensor Information – Two Views on the World

Andreas Hotho (Universität Würzburg, DE)

In the last decade the social web emerged and had a strong influence on everyone's daily live. Today, most of the newly bought mobile phone are smartphones which have a bunch of additional sensors on board. Using this novel combination of sensor information and opinions of users uttered in the social web will lead to a new level of information quality. This talk will discuss this emergent new area along examples from the EveryAware project. We will use the results to illustrate and explain the future changes and challenges.

3.11 Weaving the Social Web into User Modeling and Adaptation

Geert-Jan Houben (TU Delft, NL)

The social web is having a clear impact in the field of user modeling and adaptation. On the social web a large source of data s generated by users themselves, often for different purposes, and that provides an unprecedented potential for systems to understand their users and to adapt based on that understanding. As we can see from researchers and projects in a number of relevant fields, data on various manifestations of what users do socially on the web brings new opportunities. Exciting ideas are generated and first explorations show promising results. In this talk we aim to understand the impact on methods and techniques for user modeling and adaptation. We also look forward by identifying challenges that can drive our research.

3.12 Smart Social Systems

Ramesh Jain (University of California – Irvine, US)

Availability of enormous volumes of heterogeneous Cyber-Physical-Social (CPS) data streams may allow design and implementation of networks to connect various data sources to detect situations with little latency. In fact, in many cases it may even be possible to predict situations well in advance. This opens up new opportunities in designing smart social systems for specific tasks. Such systems may be very useful for many important problems at local as well as regional and even global level. We believe that such systems offer many novel challenges to researchers in multimedia, particularly in social and cross-modal media systems. We will present our ideas and challenges derived from our early experience towards building smart social systems.

3.13 Future of End-User Configuration of IoT : "Do It Yourself" or (only) "Choose It Yourself"?

Artem Katasonov (VTT – Espoo, FI)

This talk aims at initiating a discussion about what the future holds with respect to end-user configuration of IoT environments, such as a smart home. Will users "program" their homes using services like IFTTT.com, or they will prefer to browse through thousands of applications published in an online app store and try some of them out? Both approaches pose a number of research issues, and it is argued that the semantic technology can help to approach most of them. In addition, the talk briefly describes three related research efforts at VTT: Smart Modeller (for do-it-yourself future) and Semantic Smart Gateway Framework and Semantic Agent Programming Language (for app store future).

3.14 Using Insights from Social Computing to Augment Automotive Sensory Data

Claudia Müller-Birn (FU Berlin, DE)

Social computing systems provide added value by processing user information created by social interactions. Based on a couple of examples we show how these added values are realized in the Web and what challenges exist to tape the whole potential of these systems. Nowadays, ideas from social computing are entering completely new areas such as in the automotive industry. The car manufacturer Ford, for example, provides freely and at no cost its Sync AppLink to any automaker. The idea is to have a shared platform for developers to create novel apps based on sensory data. We enter this emerging area by presenting first results from a project on car data we have carried out recently and show, how car data can provide benefit to humans or web services. We conclude the talk by highlighting existing drawbacks and discussing implications for future research.

3.15 Towards Linked Closed Data?

Axel Polleres (Institute for Information Business of WU Wien, AT)

In the current trend for open Data, a lot of optimism is join into the belief that efforts like Linked Open Data from public sources will enrich and enable the usage of closed sensing data from all kinds of sources, and that aggregated dynamic sensing data will again be potentially published openly. However, various variables are unbound in this equation: How private can data in physical-cyber-social computing be? Can linked open data be trusted? How can physical-cyber-social-data be protected? How can data be charged and what's the value of aggregated data? I don't have answers to these questions but I'd like to discuss these issues in the workshop along with a roadmap and strategies on enabling technologies to answer them.

3.16 Identity: Physical, Cyber, Future

Matthew Rowe (Lancaster University, GB)

Social web systems offering communication functionality allow users to form groups, make connections and shape their identity over time. The development of identity, and the theoretical underpinnings that currently explain such developments, are based on psychoanalysis grounded by real-world, physical experiences. In this talk I will explain how such theories transcend physical-cyber boundaries and that users also exhibit identity crises in social web systems when interacting in a cyber environment. Such a transcendent phenomena leads to questions such as: how can identity be defined in the future? Does behaviour diffusion occur between the cyber and physical worlds? And how can we pre-empt physical decisions through cyber-based analyses?

3.17 Approximate Services in the Internet of Things

RangaRao Venkatesha Prasad (TU Delft , NL)

With the advent of newer technologies and highly miniaturized and computationally capable communicating devices, many possibilities of service provisioning is opening-up. The ICT (Information and Communication Technology) devices are now getting into our everyday life without our notice. Now, with the advent of Internet of Things (IoT), many newer possibilities of service provisioning are opening up. Since the needs of a person are different from another and so many different situations have to be dealt with, an exact service could not always be offered. To deal with this, approximate services paradigm is proposed. The idea is to find services that are close to the required with whatever the surrounding devices could provide opportunistically. We provided an example to motivate towards such a paradigm.

We identify some structural components of such a service. We expect a truly Google like searching for objects in our daily life is not ruled out in near future. We proposed many aspects of approximate service provisioning. We discussed briefly the notion and concept of approximate services and also the ideas of achieving such a paradigm. Further, since we know that there could be enormous number of devices in the future and in particular in IoT, a centralized solution seems impractical. Thus the use of distributed approximate services is the target. Here is a list of articles that deal with approximate services.

3.18 Physical-Cyber-Social Systems, Challenges and Opportunities

Amit P. Sheth (Wright State University – Dayton, US)

Main reference A. P. Sheth, P. Anantharam, C. Henson, "Physical-Cyber-Social Computing: An Early 21st Century Approach," IEEE Intelligent Systems, Vol. 28, No. 1, pp. 78–82, 2013.
URL http://dx.doi.org/10.1109/MIS.2013.20

The proper role of technology to improve human experience has been discussed by visionaries and scientists from the early days of computing and electronic communication. Technology now plays an increasingly important role in facilitating and improving personal and social activities and engagements, decision making, interaction with physical and social worlds, generating insights, and just about anything that an intelligent human seeks to do. Increasing number of exciting and important applications that technology have started to enable now include three interacting components:

- physical component: data about the physical world as captured by the sensors/devices and Internet of Things (IoT),
- social component: social interactions as enabled by social networks as well as the use of sensors to capture human interactions, and
- cyber component: by the use of massive amount of background knowledge often created through collective processes (e.g., Wikipedia), and other factual data (such as those becoming part of the Linked Open Data).

Given that the cyber component also provides computing needed to bridge physical world to the social world made up of humans, and whom any technology ultimately seeks to serve, we call this emerging computing paradigm as Physical-Cyber-Social Computing.

Related papers and talks/keynote on this topic appear at:
http://wiki.knoesis.org/index.php/PCS

3.19 Physical-Cyber-Social Agriculture

Kerry Taylor (CSIRO – Canberra, AU)

Many of the major challenges facing the world in the decade ahead are focused on food production: feeding the growing population; preserving biodiversity; mitigating and responding to climate change. As scientists, we dedicate ourselves to reducing uncertainty, whereas practicing farmers are daily experts in decision-making under uncertainty. How can we improve the precision and reliability of information for farm management? How can we make that information more directly actionable in the farmer's knowledge-intensive world? And are the future technology developments a threat to the best of traditional rural lifestyle and culture?

4 Working Groups

Ramesh Jain, Amit P. Sheth, Steffen Staab, Markus Strohmaier, Payam Barnaghi, Rajendra Akerkar, Pramod Anantharam, Josiane Xavier Parreira, Harith Alani, Ciro Cattuto, Benoit Christophe, Edward Curry, Emanuele Della Valle, Schahram Dustdar, Frieder Ganz, Michael Granitzer, Manfred Hauswirth, Laura Hollink, Andreas Hotho, Geert-Jan Houben, Mohan S. Kankanhalli, Artem Katasonov, Claudia Müller-Birn, Axel Polleres, RangaRao Venkatesha Prasad, Matthew Rowe, Vivek K. Singh, Kerry Taylor, and Koji Zettsu

4.1 Physical-Cyber-Social Computing

The discussions were followed in three main working groups focused on data, semantic and social aspects of Physical-Cyber-Social. Several joint discussions and talks, that are included in this report, also addressed inter-relations between these topics. The following sections summarise some of the key issues, challenges and solutions that were discussed in the working groups.

4.2 Working Group on Data

Physical-Cyber-Social (PCS) applications are characterised by the fact that information is represented through different types of data such as text (alphanumeric), images, audio, video and various data types such as like temperature, sound and light. The abundance of physically different data types or data heterogeneity is driven both by the increasing capabilities of computational systems as well as the increasing sophistication and ease of use of digital sensor technologies.

Some important issues with regards to the data in such systems include:

- Volume of data is growing by orders of magnitude every year.
- Different data sources provide facets of information which have to be combined to form a complete picture.
- There is a need to describe the dynamic nature of data and its refinement over time
- In several applications, real-time data processing is crucial
- Data assimilation in Physical-Cyber-Social system increasingly requires to consider the spatio-temporal characteristics of the data.

The above issues result in some challenges that are listed below:

1. How to trust the quality of data?
2. How to find useful data sets out of available data?

3. We need a mechanism to reason what the data pipelines do.
4. How to wrap up data, coming from multiple users and heterogeneous devices, into a common format and make accessible to the system?
5. How to handle multiple issues of scale, real-time processing and indexing of the physical organisation of data?
6. How to understand the power of data and a data pipeline? Determine the expressiveness of operators.
7. How to translate data from localised sensor/human input to higher level situational abstractions?

The Data Working Group followed several discussions with regards to the above issues and challenges. The following provides a summary of the discussions.

There are a lot of open issues which need to be solved to access wide variety of data with Physical-Cyber-Social computing. In order to make the data available for various tasks in Physical-Cyber-Social computing we need to describe the data. Web3.0 approaches provide appropriate solutions to describe sensor data in an abstract way. A more abstract and machine readable description of data allows for a better use of data. However, to make this vision a reality, we have to solve problems resulting from limited resources, limited bandwidth and/or missing and contradicting data. To sense and report the real world observations and measurements, one facet is using citizen sensing. Citizen sensing has many advantages to machine sensing. Machines are good at symbolic processing but are limited with respect to perception, which is the act of transforming sensory information into symbols/words that are expressive to humans. Furthermore, humans are better at contextualising data, filtering across multiple modalities, and capturing the resulting observations for future symbolic processing by machines. The Physical-Cyber-Social data can be combined with other data to create different abstractions of the environment, or it can be integrated into the data processing chain in an existing application to support context and situation awareness. It is necessary that heterogeneous data can be effectively integrated or one type of data can be combined with other physical, cyber, or social world data. Here, enhancing data interoperability (semantic interoperability), between different sources, by means of standardisation (common models) and benchmarks (describing the best scientific approach to describe the data) is required to facilitate the PCS data integration with other existing domain knowledge.

The cyber data can help to interpret/enrich the physical world data. There are many open issues. There is a need for a coherent implementation-independent description of the real world data by including rich semantics. There is also a need to describe the dynamic nature of data and its refinement over time. Finally, there is a need to provide patterns and best practices that describe how to implement such data descriptors in resource constrained environments.

4.3 Working Group on Semantics

The discussions in this working group focused on how the semantic Web standards should be extended and/or adapted to make them more suitable for Physical-Cyber-Social computing. One of the main questions was if we need upper ontologies for sensors. The participants argued that probably not; there are existing (partial) ontologies that cover important aspects of the requirements for the current use-cases. The group discussed whether concrete extensions and best practices of standard in terms of how to model "context" would be necessary? For example the W3C SSN ontology [SSN1], [SSN2] can be extended with concrete (recommended?) ontologies for modelling temperature, units, etc.

The following summarises the discussions on various concepts regarding the role and issues of using semantics in Physical-Cyber-Social systems that were discussed.

Provenance. Provenance is a key element in understanding where the data comes from the physical world (e.g. in what conditions the measurements were recorded, at which location and at what time). This is particularly important to derive and understand the different contexts in cases that the data is used in different distributed systems and applications and also when the data is aggregated (e.g. certain averages might be location or time dependent).

The W3C SSN ontology already provides a set of properties that allow recording provenance information. What is left to be done is to analyse how existing sensor ontologies need to be extended and how they align with other provenance ontologies so that the provenance for sensor data can be defined. This is an opportunity that needs to be explored.

Data linking. There are several existing efforts on integrating sensor data with static data at the metadata level. There might be an opportunity to investigate how to link dynamic data to static data via semantics. This will require to analyse use-cases and specific challenges in order to provide dynamic links and update the associations between the linked resources.

Data annotation / Standards for data representation and processing. The key question related to annotation is whether having standards are always the best choice. An alternative approach to common standardisation can be providing data structures that are better suited to certain applications, e.g. annotations for sensor data, as opposed to using properties. This also applies to the query languages. Named graphs are not suitable for such annotations. This is particularly true for annotations that are associated with certain "logics".

An important concept to note is that the linking is not an "either or" decision; time and other attributes might require to be stored in annotations for faster processing, whether other types of provenance (e.g. description of a procedure) can take advantage of the provenance ontologies. Existing query processing solutions would need to be revised to account for the annotations. Another opportunity is using the temporal logic for reasoning data in Physical-Cyber-Social computing.

Emerging semantics. There are existing solutions that describe containers (e.g. representation models to express Resource A is a sensor), but the same is not true for the content (e.g. is the weather hot or cold). This applies to both physical and citizen (social) sensing. However, it is even more challenging when it comes to social sensors, as their content often do not have a defined domain (e.g. what does a Tweet mean? Is it an event an observation or a feeling?)

One of the challenges is how to model social sensors. Extracting semantics from text only works in static environments. In social sensors the semantics change. This is different from physical sensors because on its own an individual sensor observation is rather meaningless. How should this be attacked? How can such social data be correlated with physical data?

Another question is related to tools and whether the existing tools are adequate to extract information from social sensors. The semantics might influence the data extraction procedure. The interpretation of social data may require a "longer pipeline" compared to physical data. Context description becomes even more challenging in social semantics, where the meaning of a particular post or tweet, and how this relates to other entities can be very subjective.

4.4 Working group on Social Systems

Social science research for decades has been mostly studying social interactions in the physical world by formulating social science theories and conducting scientific experiments. These experiments are focused on the social interactions (interpersonal) in the physical world. Social networks on the cyber world have lead to unprecedented connectivity for information sharing and social interactions extending beyond the social interactions in the physical world. Citizens behave as sensors outside the context of social interactions and this provides a valuable complementary source of information in a Physical-Cyber-Social system.

With Physical-Cyber-Social computing there is an opportunity to understand two fundamental questions:

Q1. How does the emergence of physical-cyber-social systems enable new forms of social science research? Social scientists have access to the social networks on the cyber world which serves as a valuable source of information to understand social interactions. Social networks consist of social connections and interactions in a dynamic setting between people in the network. Social scientists are no longer limited to social interactions in the physical world. They can study interactions in the social networks to validate existing theories and to propose new theories. The validation of social theories is conventionally done by conducting surveys and by employing people to participate in a study. Social scientists have now access to the data from the real-world social interactions in the cyber world. These interactions provide a sample (incomplete) picture of the real-world social interactions but on a massive scale. The social science theories need to be adopted to deal with this incompleteness and data biases in a physical-cyber-social system.

Q2. How can we use social theories to inform the design of novel Physical-Cyber-Social systems? Understanding social theories is the first step toward building systems that interact smoothly with people. These theories are crucial in the design and development of physical-cyber-social systems. For example, a system that interprets various social interactions can be used to capture images/videos. One of the key challenges in dealing with social systems is how to maintain the privacy of participants. Privacy becomes an essential component and it is crucial for wide adoption of physical-cyber-social systems. This is partly due to the fine-grained information collected from sensors and its correlation with behavior patterns that would reveal personal information which may be misused. For example, a smart-meter installation may result in revealing the occupancy of a house.

Some of the key challenges when using social data, or in general any data processed by Physical-Cyber-Social computing are listed below.

- Social bots: there are attempts to simulate and flood the social data generated automatically by programs that try to emulate human behaviour. Such a source of information should be used carefully and separated from rest of the social data.
- 'Twitter' data is not always reliable (i.e. requires careful consideration): while social data is available in massive scale (e.g., around 500 million tweets a day), the data is often very noisy, informal, and unevenly distributed.
- Assessing the relevance of Twitter to a problem: not all studies can be done on Twitter data since the nature of data and the social behaviours have a great variance on Twitter.
- Sometimes there is an assumption that data is available at all times: theoretically, there is data available related to various events. However in reality, it may be very hard to find sensors and their observations on Twitter and the Web in general. Choosing the

appropriate data source is an important challenge in the context of Physical-Cyber-Social systems.

- Understanding the feedback mechanism: social scientists need to understand the feedback mechanism that exists between the physical world and the social world interactions. This is a challenging and important task to gain insights into systems that involve the social component.
- Data biases are crucial: social scientists should consider data biases carefully with the availability of massive data from social networks such as Twitter.
  - Tweeter vs. non-Tweeter: for example, not all the social interactions are present on Twitter and only a group of people may use Twitter.
  - people who can read vs. cannot read: such biases have been around even in the past when social scientists has to deal with the physical world for conducting their studies.
- Combining reactive vs. non-reactive data: reactive data are those collected by social scientists through surveys and questions. Non-reactive data are those collected by sensors on a continuous basis.
  - collecting sensor data: this is a type of non-reactive data and it is available continuously without active involvement of people.
  - how do we combine survey data and log data?: this is a challenge as the granularity of both of them are different. Also, combining textual data (e.g., emotion) with sensor data (e.g., high heart rate) may be particularly difficult.
  - understanding the differences between reactive vs. non-reactive data enable us to design methodologies for dealing with them in a single framework.
  - Can the theories from social science be translated into graph formalisms and whether this will help expose social theories?

5 Open Problems

As the Web provided useful mechanisms to access and use new types of resources, techniques increasingly moved from syntactic and structural to semantic representations. There is also a recent resurgence of research towards Computing for Human Experience [Sheth10]. This line of work has a long lineage, starting in part with Vannevar Bush's Memex through Mark Weiser's "Computing in the 21st Century" and others. But the essence of the vision incorporates technology that serves human needs without explicit human effort. This multidisciplinary seminar seeks to develop a vision of a new class of 21st century systems involving machines (physical), computing and communications (cyber) and human-centric and social systems (social). Several challenges arise in this context, which can be categorized into the following major topics that were discussed at the seminar:

1. emerging manifestations of physical-cyber-social systems, including applications in industry markets such as health and transportation. Specifically investigation of new capabilities in terms of improving human experience and serving societal causes in ways that have not been possible earlier.
2. How are emerging physical-cyber-social systems different than integration of cyber-physical systems with social systems as conceived today?
3. How do physical-cyber-social systems utilise or benefit from Web of Things/ Internet of Things, Semantic Web, Crowd sourcing and Semantic Social Web, Semantic Sensor Web, Intelligent/natural Interfaces, Ambient Intelligence and other technologies/ advances during the last decade?

4. What novel disruptive applications are likely to result due to emerging technology in Physical-Cyber-Social systems after the initial obvious applications are addressed?
5. What is the role of semantics in physical-cyber-social systems, and how do existing semantic technologies accelerate or constrain the emergence of Physical-Cyber-Social computing?
6. How do physical-cyber-social systems transform traditional perceptions of physical objects, online engagement and social interactions?
7. What implications will the confluence of physical-cyber-social systems have on societies, including aspects such as citizen participation, democracy, open government, open government data and others?

6 Panel Discussions

Soon there will be tremendous volumes of data collected from the physical world in addition to the growing data collected from the cyber and social world. The combination of this cyber-physical and social data can help us to understand events and changes in our surrounding environments better, monitor and control buildings, homes and city infrastructures, provide better healthcare and elderly care services among many other applications. To make efficient use of the physical-cyber-social data, integration and processing of data from various heterogeneous sources is necessary. Providing interoperable information representation and extracting actionable knowledge from deluge of human and machine sensory data are the key issues.

The role of semantics for interoperability, integration, and improved querying has been investigated for over four decades. The 'Semantic Web' movement brought focus to using semantics and metadata initially to the Web documents. As the Web provided useful mechanisms to access and use new types of resources–richly represented data, services, user generated content and other social data, sensor and devices (Web of Things – WoT) data–techniques increasingly moved from syntactic and structural to semantic ones. Compared to the semantic systems built using Semantic Web languages, standards and mainstream Semantic Web technologies that employ formal representation of semantics, however, more systems are being built using informal and implicit forms of semantics. One reason is that the role of the Web is increasingly becoming diffused and incidental (e.g., more people access content through applications compared to the Web browsers). The second reason is that lighter-weight approaches have led to better developer and user engagements, and have become a lot more scalable. Apple Siri, IBM Watson, and Google Knowledge Graph are examples of using semantics at scale, but where the formal form of semantic representation or RDF/SPARQL have not found a place. All these lead us to think that 10 years from now, Semantic Web would be thought of as something that popularized the core value proposition of semantics – better search, interoperability/integration and analysis – to deal with and exploit a vast variety of things that the Web (and its on going transformations) interconnects. An analogy that comes to mind is that of Object Oriented Databases which generated huge excitement in the 1980s, and indeed had a number of secondary impacts, it only remained a niche technology, product class and market. Simultaneously, Semantic Web is increasingly merging with other powerful technologies that support semantics, including Machine Learning, NLP, and Knowledge-based systems where background knowledge is applied. Consequently, what we think of as rather distinct Computer Science areas today will not retain strong distinctions, but will broadly incorporate semantics.

References

1 Amit Sheth, "Computing for Human Experience: Semantics-Empowered Sensors, Services, and Social Computing on the Ubiquitous Web," IEEE Internet Computing (special issue on Internet Predictions), 14 (1), January/February 2010, pp. 88–91.

2 Laurent Lefort *et al*, Semantic Sensor Network XG Final Report, W3C Incubator Group Report 28 June 2011, http://www.w3.org/2005/Incubator/ssn/XGR-ssn-20110628/

3 Michael Compton *et al*, The SSN ontology of the W3C semantic sensor network incubator group, Journal of Web Semantics, Volume 17, December 2012.

4 Moshe Vardi, Computing for Humans, Communications of the ACM, Vol. 54, No. 12, December 2011, page 5.

Participants

- Rajendra Akerkar
Vestlandsforsking – Sogndal, NO
- Harith Alani
The Open University – Milton Keynes, GB
- Pramod Anantharam
Wright State University – Dayton, US
- Payam M. Barnaghi
University of Surrey, GB
- Ciro Cattuto
ISI Foundation, IT
- Benoit Christophe
Bell Labs – Nozay, FR
- Edward Curry
National University of Ireland – Galway, IE
- Emanuele Della Valle
Polytechnic Univ. of Milan, IT
- Schahram Dustdar
TU Wien, AT
- Frieder Ganz
University of Surrey, GB
- Michael Granitzer
Universität Passau, DE
- Manfred Hauswirth
National University of Ireland – Galway, IE
- Laura Hollink
Free Univ. of Amsterdam, NL
- Andreas Hotho
Universität Würzburg, DE
- Geert-Jan Houben
TU Delft, NL
- Ramesh Jain
Univ. of California – Irvine, US
- Mohan S. Kankanhalli
National Univ. of Singapore, SG
- Artem Katasonov
VTT – Espoo, FI
- Claudia Müller-Birn
FU Berlin, DE
- Axel Polleres
Wirtschaftsuniversität Wien
- RangaRao Venkatesha Prasad
TU Delft, NL
- Matthew Rowe
Lancaster University, GB
- Amit P. Sheth
Wright State University – Dayton, US
- Vivek K. Singh
MIT, US
- Steffen Staab
Universität Koblenz-Landau, DE
- Markus Strohmaier
Universität Koblenz-Landau, DE
- Kerry Taylor
CSIRO – Canberra, AU
- Josiane Xavier Parreira
National University of Ireland – Galway, IE
- Koji Zettsu
NICT – Kyoto, JP