



Northeastern

# OUTLINE

---

- ✦ schedule

- ✦ project scope: peer-assessment of scientific publications

  - ✦ publications

  - ✦ review process

  - ✦ goals

- ✦ deliverables

# SCHEDULE

---

	topic	start	end	note
Phase 1	requirements & use cases	Jan. 24	Feb. 3	individual
Phase 2	UML design & Java interfaces	Feb. 3	Feb. 21	team-based
Phase 3	implementation	Feb. 24	Mar. 24	team-based
Phase 4	adaptation & extension	Mar. 28	Apr. 18	team-based

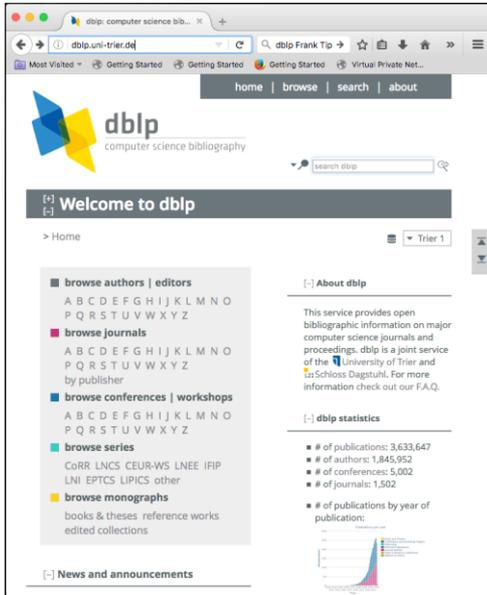
- ✦ no collaboration allowed in Phase 1 !
- ✦ the instructors will assign you into teams for Phase 2-4
- ✦ team composition will be changed during the project!

# PUBLICATIONS

---

- ✦ computer scientists publish papers about their research
  - ✦ in **conferences** and **journals**
  
- ✦ all publications have at least:
  - ✦ title
  - ✦ author(s)
  - ✦ page numbers
  - ✦ year of publication
  
- ✦ conference publications have:
  - ✦ name (e.g., ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications)
  - ✦ acronym (e.g., “OOPSLA”)
  
- ✦ journal publications have:
  - ✦ name (e.g., ACM Transactions on Software Engineering and Methodology)
  - ✦ acronym (e.g., “TOSEM”)
  - ✦ volume
  - ✦ issue

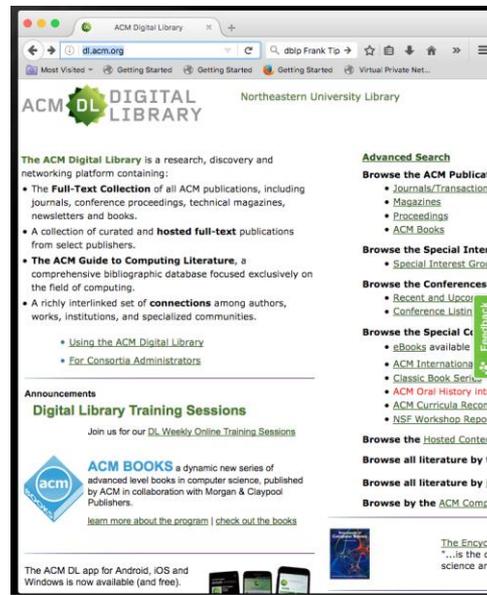
# ACCESSING PUBLICATION DATA



The screenshot shows the DBLP (Computer Science Bibliography) homepage. The browser address bar displays 'dblp.uni-trier.de'. The page features a navigation menu with 'home', 'browse', 'search', and 'about'. A search bar is located below the navigation. The main content area is titled 'Welcome to dblp' and includes a 'Home' button. There are several categorized lists for browsing: 'browse authors | editors', 'browse journals', 'browse conferences | workshops', and 'browse series'. A sidebar on the right contains 'About dblp' and 'dblp statistics' with a small line graph. The bottom of the page has a 'News and announcements' link.

**DBLP**

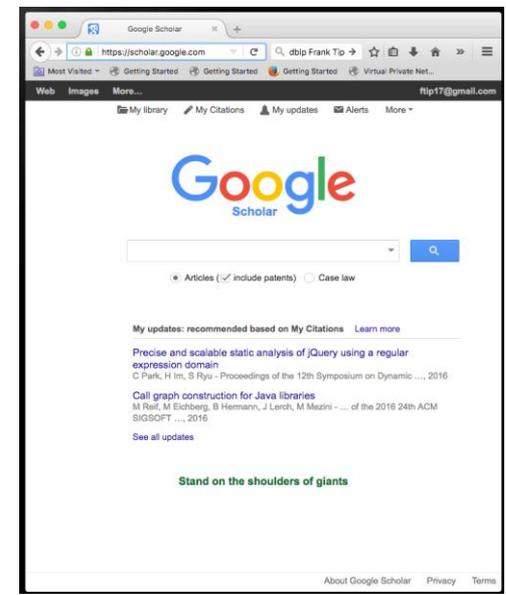
<http://dblp.uni-trier.de/>



The screenshot shows the ACM Digital Library homepage. The browser address bar displays 'dl.acm.org'. The page header includes the 'ACM Digital Library' logo and 'Northeastern University Library'. The main content area describes the library as a research, discovery, and networking platform. It lists key features: 'The Full-Text Collection of all ACM publications', 'The ACM Guide to Computing Literature', and 'A richly interlinked set of connections'. There are several navigation links and sections: 'Advanced Search', 'Browse the ACM Publications', 'Browse the Special Interest Groups', 'Browse the Conferences', 'Browse the Special Collections', 'Announcements', 'Digital Library Training Sessions', 'ACM BOOKS', and 'Browse all literature by type'. The bottom of the page mentions the availability of the ACM DL app for Android, iOS, and Windows.

**ACM Digital Library**

<http://dl.acm.org/>



The screenshot shows the Google Scholar homepage. The browser address bar displays 'https://scholar.google.com'. The page features the Google Scholar logo and a search bar. Below the search bar, there are options for 'Articles (checked)' and 'include patents', and a 'Case law' option. The page also displays 'My updates: recommended based on My Citations' and a section for 'Precise and scalable static analysis of jQuery using a regular expression domain' by C Park, H Im, S Ryu. The bottom of the page includes a 'Stand on the shoulders of giants' quote and links for 'About Google Scholar', 'Privacy', and 'Terms'.

**Google Scholar**

<https://scholar.google.com/>

# PUBLICATIONS

---

- ✦ computer scientists publish papers about their research
  - ✦ in **conferences** and **journals**
  
- ✦ all publications have at least:
  - ✦ title
  - ✦ author(s)
  - ✦ page numbers
  - ✦ year of publication
  
- ✦ conference publications have:
  - ✦ name (e.g., ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications)
  - ✦ acronym (e.g., “OOPSLA”)
  
- ✦ journal publications have:
  - ✦ name (e.g., ACM Transactions on Software Engineering and Methodology)
  - ✦ acronym (e.g., “TOSEM”)
  - ✦ volume
  - ✦ issue

# PEER REVIEW

---

- ✦ publications are **reviewed** by peers
  - ✦ researchers who published papers on similar topics
  - ✦ typically, only 15-25% of submitted papers is accepted
- ✦ conference reviewing
  - ✦ reviewing is done by a **Program Committee (PC)**
  - ✦ each PC member writes a review
  - ✦ managed by a **Program Committee Chair (PC Chair)**
  - ✦ a new PC is formed each year, for each instance of a conference
  - ✦ for most conferences, members of a PC participate in a physical meeting to discuss all submitted papers & make accept/reject decisions
- ✦ journal reviewing
  - ✦ overall decisions made by **Editor-in-Chief**, who manages an **Editorial Board of Associate Editors**
  - ✦ each submitted paper is assigned to an Associate Editor, who solicits reviews from a number of subject-matter experts
  - ✦ Associate Editors typically serve for a fixed term on an Editorial Board (e.g., 3 years)

# THE PROBLEM

---

- ✦ finding good PC members or journal reviewers is a challenge:
  - ✦ ideally, a PC contains experts on all topics covered by the submitted papers
  - ✦ leading experts are often busy and may not have time for reviewing
  - ✦ relying too much on a PC chair's personal network may lead to real or perceived bias in the reviewing process
  - ✦ need reasonable coverage of all areas
- ✦ most conferences impose constraints on the composition of program committees
  - ✦ geographic diversity
  - ✦ reasonable mix of male/female, academia/industry PC members, etc.
  - ✦ reasonable mix of junior/senior researchers
  - ✦ size is usually constrained to 30-35 people
  - ✦ the same person should not serve on the PC for a conference for more than 2 consecutive years
- ✦ similar constraints for journal reviewing

# PROJECT GOALS

---

- ✦ build a tool that can assist PC Chairs & Associate Editors with the identification of suitable candidates
- ✦ the application should have an interactive user-interface to execute queries against publication data from sources such as DBLP and ACM DL such as:
  - ✦ show the names of authors who published more than one paper in OOPSLA since the year 2010.
  - ✦ show the names of authors who published at least one paper with the words “pointer” and “analysis” in the title.
  - ✦ show the names of authors who published at least two papers in OOPSLA or ECOOP and did not serve on the committee during the last two years.
  - ✦ identify authors who have “similar” profiles to a given author.
- ✦ initially, you’ll be working with data from DBLP. Other sources of data will be considered in phase 4

# REQUIRED INFRASTRUCTURE

---

<b>implementation language</b>	Java
<b>development environment</b>	Eclipse
<b>automated testing</b>	JUnit
<b>version control</b>	NEU GitHub
<b>issue tracking</b>	JIRA
<b>requirements/documentation</b>	Confluence
<b>continuous integration</b>	Jenkins
<b>messaging</b>	Slack (optional)
<b>integration environment</b>	AWS (preferred)

# PHASE 1: DELIVERABLES

---

- ✦ Write 5-10 use cases that describe the system's functionality
- ✦ The clients (aka your instructors) will be present in class on January 27 for a Q&A session
- ✦ Your deliverable must be submitted as a PDF file in BlackBoard by February 3 at 23:59pm.
- ✦ Your submission must be no longer than five (5) pages.

# NEED HELP WITH THE PROJECT?

---

- ✦ Ask your team members first
- ✦ ...If you still need help, ask one of the TAs
- ✦ ...If you still need help, ask the instructor
- ✦ Note: questions of a general nature can be asked and answered on Piazza