**Workshop: De-identification of Narrative Clinical Documents**
Stakeholders' perspectives and acceptance of machine-based de-identification
National Cancer Institute, 9609 Medical Center Drive, Rockville, MD

# Agenda

| DAY 1 | Tuesday, February 25, 2020 | Room TE406 Terrace level |
|---|---|---|
| 8:00-9:00 am | Arrival, security check, breakfast on your own in cafeteria on T-level | |
| 9:00-9:05 am | Welcome and Logistics | Lynne Penberthy Annelie Landgren |
| 9:05-9:10 am | Workshop goals and objectives | Valentina Petkov |
| 9:10-10:15 am | *Session 1: De-identification of clinical information for research use: Regulatory landscape*<br>1. Coping with Evolving Standards for De-Identification of Clinical Information (30 min)<br>2. Issue Paper on Disclosure Review for Information Products with Qualitative Research Findings (15 min)<br>*Timing: Questions and discussion 20 min* | Presenters<br>1 Kristen Rosati<br>2 Joanne Pascale |
| 10:15 -10:30 am | Break | |
| 10:30 -11:45 am | *Session 2: Current status, performance, and future development of clinical text de-identification systems*<br>1. NLM scrubber<br>2. CliniDeID<br>3. NeuroNED<br>4. MIST - retrained model<br>*Timing: 15 min presentations + 15 min at the end questions/ discussion* | Presenters<br>1 Mehmed Kayaalp<br>2 Stephane Meystre<br>3 Ozlem Uzuner<br>4 Guergana Savova |
| 11:45 -12:45 pm | Lunch on your own (NCI cafeteria, level T) | |
| 12:45 -1:45 pm | *Session 3: Real world experience with de-identification of narrative clinical documents for research: Importance, challenges, data sharing policies and practices*<br>1. UPMC experience<br>2. ORNL experience<br>3. MIMIC experience<br>*Timing: 15 min presentations + 15 min at the end questions/ discussion* | Presenters<br>1 Jonathan Silverstein<br>2 Ioana Danciu<br>3 Alistair Johnson |
| 1:45-3:15 pm | *Session 4: Important aspects to consider in de-identification of narrative information*<br>1. Effectiveness of human redaction of clinical text documents<br>2. Current challenges and state-of-the art in text de-identification<br>3. Best practices in de-identification and sharing of narrative documents for research<br>4. Managing the risk: Assessment of re-identification<br>5. Ethics, software engineering and de-identification: Titrating protections to serve stakeholders' interests<br>*Timing: 15 min presentations + 15 min at the end questions/ discussion* | Presenters<br>1 David Carrell<br>2 Simson Garfinkel<br>3 Daniel Barth-Jones<br>4 Bradley Malin<br>5 Kenneth Goodman |

NOTE: Entering the building can take 20 minutes or more. Please arrive at the guard station no later than 8:30 to allow for the "airport" style security scan. You need a U.S government issued ID card or a passport.
There will be a chaperone waiting for you at the guard station between 8 and 9 am each day. If you need to exit the building during the day, please let us know, since you will have to go through security and be chaperoned to the meeting room.

| | | |
|---|---|---|
| 3:15- 3:30 pm | Break | |
| 3:30-4:45 pm | *Session 5: Stakeholders perspective and acceptance of machine text de-identification: What do we need to consider?*<br>1. Acceptability of automated de-identification (IRB survey)<br>2. Researcher perspective: Does it matter how the PII is obliterated in the clinical text?<br>3. Commercial/Academic IRB-perspective<br>4. Data owner perspective<br>5. Data owner perspective<br>6. Patient advocate perspective<br>*Timing: 10 min presentations + 15 min at the end questions/ discussion* | Presenters<br>1 David Carell<br>2 Guergana Savova<br>3 Michele Russell-Einhorn<br>4 Eric Durbin<br>5 Charles Lynch<br>6 Cassadie Moravek |
| 4:45 -4:55 pm | *Setting the stage for the round table discussion on Day 2*<br>(Discussion questions to be distributed prior to the workshop) | Valentina Petkov |
| 4:55-5:00 pm | Wrap-up of Day 1 | Valentina Petkov |
| 6:30 pm | No host dinner at Yard House, 211 Rio Blvd, Gaithersburg, MD 20878 https://www.yardhouse.com/menu-listing/eat. Happy hour until 6pm. | Headcount will be collected after lunch |
| **DAY 2** | **Wednesday, February 26, 2020** | **Room 2E 908**<br>**2nd floor** |
| 9:00-12:00 am | *Session 6 Round table discussion:* Outline of white paper on clinical text de-identification for research use and sharing | Moderator: TBD |

## Contact persons at the venue:

Annelie Landgren, phone (no text): 240 276 7368, email: landgrena@mail.nih.gov
Ella Saccon, phone or text: 914 924 7697
Valentina Petkov, phone or text (Ella or Annelie will answer): 240 760 0493

## Additional contact person:

Shantala Basavappa, phone: 240 276 6450, email: basavappasr@mail.nih.gov

## Shuttle Service from the hotel to the event

Tuesday, Feb. 25th
Pick-up location:     Courtyard by Marriott Gaithersburg Washingtonian Center (Front of Hotel)
Shuttle Arrive:       7:45 am - (The shuttle will depart the Courtyard at precisely 8:00 AM and will not make a second trip)
Depart location:      9609 Medical Center Drive (Front Entrance)
Shuttle Depart:       5:15 pm - (The shuttle will depart NCI at precisely 5:30 PM (no second trip)
Wednesday, Feb. 26th
Pick-up location:     Courtyard by Marriott Gaithersburg Washingtonian Center (Front of Hotel)
Shuttle Arrive:       7:45 am - (The shuttle will depart the Courtyard at precisely 8:00 AM (no second trip)