**Data and Tools**

**I. New 2016 datasets for Subtasks A, B, C, D, E** (either one can be used for both training and tuning).

* [TRAIN + DEV+DEVTEST](http://alt.qcri.org/semeval2016/task4/data/uploads/semeval2016-task4.traindevdevtest.v1.2.zip) (v1.2 -- updated Nov. 5, 2015)
* [**\*\*\*TEST data\*\*\***](https://docs.google.com/forms/d/1SxDxPKbLSiKMI9XnaFWhRGTcnWgZsljATCvvw8mHRAQ/viewform)**(requires registration -- please submit for subtasks A, C, and E only! Submission for subtasks B and D is postponed;**[**see here for detail**](https://groups.google.com/forum/#!msg/semevaltweet/t9r586LWnhw/7ysBEgbkAQAJ)**)**

**II. Scorers for all subtasks:**

* [all subtasks](http://alt.qcri.org/semeval2016/task4/data/uploads/semeval2016-task4-scorers.zip)

**III. Older data**(can be used for both training and tuning).

**1. For Subtask A**

* The [TRAIN data](http://alt.qcri.org/semeval2014/task9/data/uploads/semeval2013_task2_train.zip) for SemEval-2013 task 2.
* The [DEV data](http://alt.qcri.org/semeval2014/task9/data/uploads/semeval2013_task2_dev.zip) for SemEval-2013 task 2.

**2. For Subtasks C, D, and E**

* The 2016 [trial data](http://alt.qcri.org/semeval2016/task4/data/uploads/semeval2016_task4_trial_c_d_e.zip)

**IV. Older data that can be used as dev-test only, but NOT for training or tuning**

**For Subtask A**

* [SemEval-2013 Task 2 development-test](http://alt.qcri.org/semeval2014/task9/data/uploads/semeval2013_task2_test_fixed.zip): (i) Twitter-2013 and (ii) SMS-2013 messages (**CANNOT**be used for training or tuning!)
* [SemEval-2014 Task 9 development-test](http://alt.qcri.org/semeval2014/task9/data/uploads/semeval2014-task9-subtaskab-test-to-download.zip): (i) Twitter-2013, (ii) SMS-2013 messages, (iii) Twitter-2014, (iv) Twitter-2014-sracasm, and (v) Live Journal-2014 (**CANNOT** be used for training or tuning!)
* SemEval-2015 Task 10 development-test: Twitter-2015 (**CANNOT** be used for training or tuning!)

**NOTES:**  
1. Participants can choose to download the test data at any moment during the evaluation period (January 10-31, 2016). Regardless of the time of download, results are to be submitted by January 31, 2016, 23:59 hours. The time zone is Midway, Midway Islands, United States: see<http://www.timeanddate.com/worldclock/city.html?n=1890>).  
2. The submission will be done using the SemEval START website:  
<https://www.softconf.com/naacl2016/SemEval2016/>  
3. Participants can make new submissions, which will substitute their earlier submissions on the START server multiple times, but only before the deadline (see above). Thus, we advise that participants submit their runs early, and possibly resubmit later if there is time for that (START was not closed).  
4. Participants are free to participate for a single subtask or for any combination of subtasks.

5. We allow a single run per subtask.

6. The test datasets from previous years cannot be used for training or tuning; they can be used for development-time testing only. This includes the following: (i) Twitter-2013, (ii) SMS-2013 messages, (iii) Twitter-2014, (iv) Twitter-2014-sarcasm, (v) Live Journal-2014, and (vi) Twitter-2015

7.  Participants are free to use any other data (except for what is listed in 6 above): we will not distinguish between closed (that only use the provided data) and open (that also use additional data) runs. However, they will need to describe the resources and tools they have used to train their systems in the Web form they have recieved by email.