



CORSO S.I.G.A.

“METODI ED APPLICAZIONI DI INFORMATICA PER L'ANALISI DI DATI BIOLOGICI”

Salsomaggiore Terme, 06-10 Maggio 2013

La Società Italiana di Genetica Agraria comunica l'organizzazione del Corso "Metodi ed applicazioni di informatica per l'analisi di dati biologici" che si svolgerà nel mese di Maggio 2013 da Lunedì 06 (pomeriggio) a Venerdì 10 (mattina) a Salsomaggiore Terme (PR).

Il corso, aperto ad un massimo di 35 partecipanti, è rivolto a ricercatori del settore pubblico e privato, a dottorandi, borsisti ed assegnisti di ricerca che siano interessati ad un ambiente computazionale moderno per la gestione e l'analisi di grandi quantità di dati.

L'enfasi verrà dunque posta nel presentare tecniche di elaborazione efficiente attraverso strumenti semplici, di rapida comprensione e facile utilizzo.

Il ricercatore sarà messo in grado di affrontare al meglio i propri problemi quotidiani di gestione dei dati grazie ai concetti ed agli strumenti appresi, che costituiranno una sorta di "coltellino svizzero" nel proprio bagagliaio di risorse.

In particolare, verrà presentata la shell di Linux come ambiente minimale ma potente e flessibile per l'elaborazione sequenziale di dati in formato testuale strutturato o semistrutturato assieme alle cosiddette "espressioni regolari" come tecnica di base per la realizzazione di filtri evoluti.

Verrà poi svolta un'introduzione pratica al linguaggio di scripting Python con particolare attenzione alla costruzione di programmi che sfruttino l'elevato parallelismo dei server moderni, riducendo quindi considerevolmente i tempi di elaborazione.

Non sono richieste conoscenze pregresse dell'ambiente Linux né di linguaggi di scripting.

Il corso sarà organizzato come da programma allegato.

L'alloggio ed il corso sono previsti presso l'Hotel Primarosa (Via Valentini 12, tel 0524 575549, fax 0524 573954, <http://www.hotelprimarosa.it/>) a Salsomaggiore Terme (PR), località facilmente raggiungibile con i trasporti pubblici. Gli interessati a partecipare al Corso sono pregati di compilare l'allegata scheda di pre-registrazione e di inviarla via e-mail entro il 31 Marzo 2013 alla Segreteria della S.I.G.A. (segretario@geneticagraria.it), specificando "Corso S.I.G.A." nell'area subject del mail. Entro il 10 Aprile 2013 i candidati ammessi al Corso, selezionati in base alla data di trasmissione della scheda di pre-registrazione, all'appartenza alla S.I.G.A. o a gruppi di ricerca diretti da Soci S.I.G.A. ed al curriculum, saranno contattati dalla Segreteria S.I.G.A. per le istruzioni relative al versamento della quota di partecipazione. La domanda di ammissione al Corso è scaricabile anche dal sito della Società all'indirizzo <http://www.geneticagraria.it>. Con la domanda di ammissione, ogni partecipante si impegna, per l'intera durata del corso, a frequentare regolarmente e attivamente tutte le lezioni e le attività didattiche previste. Alla fine del Corso sarà consegnato il diploma di partecipazione a firma del Presidente S.I.G.A.

La quota di registrazione al Corso è di Euro 350,00, compresa di spese per l'alloggio in regime di pensione completa. Per chi desiderasse non usufruire dell'alloggio in regime di pensione completa, la quota è di Euro 200,00. Per i Soci S.I.G.A. in regola con il versamento delle quote sociali sino al 2012 incluso, la quota di registrazione è rispettivamente di Euro 300,00 e di Euro 200,00. I partecipanti al Corso che non desiderano alloggiare possono usufruire dei pasti presso la sede del Corso al costo di 15,00 ciascuno da pagare direttamente in loco.

Il Direttore del Corso
Dott.ssa Primetta Faccioli

Il Segretario
Prof. Edgardo Filippone

Corso S.I.G.A. "Metodi ed applicazioni di informatica per l'analisi di dati biologici"

Programma generale

0. Dove elaborare dati scientifici: alcune piattaforme di riferimento. L'enorme diffusione di strumenti di informatica individuale offre l'opportunità di poter analizzare dati in scenari all'apparenza insoliti ma davvero utili. Esamineremo il caso delle applicazioni windows portabili, installabili ed utilizzabili cioè su e da una semplice chiavetta usb ed il meccanismo delle macchine virtuali, dove un sistema operativo completo viene fatto funzionare all'interno, come ospite, di un altro sistema operativo – evitando così difficili installazioni.
1. Linux, a riga di comando. L'utilizzo di Linux viene introdotto attraverso la spiegazione della sua interfaccia tradizionale, la cosiddetta riga di comando (shell), con i suoi meccanismi e programmi di base. Viene poi spiegato il filesystem: come orientarsi tra file, directory e link vari. Infine viene fatto un cenno a git, un sistema di controllo versione che permette di tenere sistematicamente traccia dei propri lavori.
2. Le espressioni regolari (regex). Uno dei problemi più frequenti è quello di ricercare, all'interno di un file, solo quelle righe che contengano determinate informazioni. Le espressioni regolari, descritte in modo introduttivo, aiutano a risolvere questo problema in modo elegante e conciso.
3. Il linguaggio Python, un'introduzione. La programmazione scientifica è stata nel tempo effettuata con tutti i linguaggi che si sono via via affermati al pubblico: nella corrente dei linguaggi di scripting (Perl, Python, Tcl tra gli altri), che predilige la produttività di chi scrive codice rispetto all'efficienza computazionale assoluta, Python è emerso in virtù della sua eleganza nei costrutti di base e della sua efficacia. Per affrontare problemi di derivazione scientifica sono state scritte apposite estensioni al linguaggio che rendono fattibili e semplici attività normalmente ritenute complesse.
4. Elementi di programmazione parallela. Le moderne CPU presentano un parallelismo intrinseco che permette di svolgere più lavori contemporaneamente. Verrà spiegato come trarre vantaggio di questa potenza nei problemi di elaborazione di grandi quantità di dati.
5. Programmare (anche) per gli altri: infine vengono presentati cenni di programmazione professionale. quando cioè si desidera che altri usino i programmi che scriviamo.