Astrofisica e Osservazione del Cosmo. Intervista a Claudio Gheller

di Pandora Rivista

11-06-2026

Glaudio Gheller - Spoke 3 del Centro Nazionale ICSC.

Quanto è visibile, oggi, l'impatto dell'evoluzione tecnologica - dal supercalcolo all'intelligenza artificiale - nel campo dell'astrofisica? E in che modo è cambiato, anche solo negli ultimi anni, il rapporto tra ricerca astronomica e capacità di calcolo?
Gheller: All'Istituto di Radioastronomia dell'INAF di Bologna mi occupo soprattutto di applicazioni per l'analisi e la simulazione dei dati, fondamentali per interpretare le osservazioni astronomiche nella banda radio. L'astronomia è stata probabilmente uno dei primi campi scientifici a sfruttare sistemi di calcolo molto avanzati soprattutto per quanto riguarda le simulazioni numeriche. Il motivo è semplice: a differenza di altre branche della scienza, l'astronomia non può riprodurre in laboratorio i fenomeni osservati nell'universo, come la formazione delle stelle o delle galassie. Non è possibile fare un "esperimento" nel senso tradizionale del termine. Per questo motivo le simulazioni numeriche sono diventate uno strumento essenziale, che permette di costruire modelli teorici e confrontarli con ciò che osserviamo nel cielo.
Già dagli anni Settanta, con i primi computer dotati di una minima capacità di calcolo, gli astronomi hanno iniziato a utilizzare strumenti numerici per interpretare i dati osservativi. Con il tempo, però, i modelli sono diventati sempre più complessi e dettagliati, aumentando di conseguenza anche il bisogno di potenza di calcolo. Negli anni Novanta le simulazioni sono state trasferite su sistemi paralleli, cioè computer in grado di eseguire molti calcoli contemporaneamente. Oggi siamo nell'era del calcolo accelerato, che sfrutta processori specializzati come le GPU, utilizzate anche per l'intelligenza artificiale. La prossima frontiera è il quantum computing: una tecnologia ancora sperimentale, ma che in futuro potrebbe offrire capacità di calcolo completamente nuove anche per la ricerca astrofisica. Su questi temi si è concentrato anche il lavoro svolto nello Spoke 3, dedicato al rapporto tra astrofisica e calcolo ad alte prestazioni.

Negli ultimi anni non è cambiato soltanto il modo di fare simulazioni, ma anche la quantità di dati prodotta dagli strumenti osservativi. Quanto sta incidendo questa trasformazione sul lavoro degli astronomi e sull'organizzazione stessa della ricerca?
Gheller: Nell'ultimo decennio è emersa l'esigenza di utilizzare enormi capacità di calcolo per l'interpretazione e l'analisi dei dati osservativi. I telescopi, i satelliti e gli strumenti di osservazione sono diventati sempre più grandi, precisi e accurati, e stanno producendo moli di dati che fino a pochi decenni fa erano assolutamente inimmaginabili e che renderanno indispensabile l'utilizzo dell'HPC per poter analizzare le informazioni. Penso, per esempio a Square Kilometre Array (SKA), un progetto internazionale per la costruzione del più grande e sensibile radiotelescopio al mondo, basato su una rete di centinaia di migliaia di antenne. A questo si stanno affiancando sempre di più anche strumenti basati su machine learning e intelligenza artificiale, perché è necessario avere sistemi di analisi il più possibile veloci e capaci di lavorare con una supervisione umana ridotta, che comunque resterà fondamentale. In passato un astronomo poteva raccogliere poche immagini,

analizzandole una ad una, cercare sorgenti, oggetti o fenomeni di interesse. Oggi un telescopio può invece acquisire una quantità di immagini la cui analisi e investigazione non può essere effettuata manualmente neanche da un ampio gruppo di persone.
Questa è una delle grandi frontiere attuali: affiancare al calcolo HPC strumenti di machine learning che, a loro volta, hanno bisogno di HPC per essere addestrati e poi utilizzati in fase di inferenza nel modo più rapido possibile, idealmente quasi in tempo reale. Il rischio, altrimenti, è di investire quantità enormi di risorse per costruire strumenti di osservazione capaci di produrre centinaia di petabyte di dati all'anno, per poi archiviarli in un sistema di storage senza avere davvero la possibilità di analizzarli. E questo vale anche per le simulazioni sempre più grandi che producono enormi quantità di dati, che spesso rischiano di restare dimenticati nei sistemi di calcolo.

Rimanendo sul tema della quantità enorme di dati prodotta dagli strumenti osservativi, quali sono oggi le principali difficoltà tecniche che la ricerca si trova ad affrontare?
Gheller: Una delle difficoltà più grandi dal punto di vista tecnico riguarda la gestione stessa dei dati: come trasferirli, archiviarli e conservarli quando si parla di quantità enormi di informazioni. Per le simulazioni il problema è in parte più semplice, perché il dato nasce direttamente sul supercalcolatore e lì può essere immagazzinato. Ma quando si iniziano ad avere decine o centinaia di terabyte, diventa comunque necessario trovare sistemi efficienti per archiviare i dati a lungo termine e trasferirli verso infrastrutture locali adeguate. E qui emergono tutti i limiti legati alla banda e alla rete: trasferire dataset di queste dimensioni richiede tempi lunghi e infrastrutture robuste. La situazione è ancora più complicata nel caso dei dati osservativi, perché i telescopi si trovano spesso in regioni remote del pianeta, oppure direttamente su satelliti. Il trasferimento dei dati diventa quindi una sfida enorme. L'esempio più evidente è probabilmente il già citato SKA, per cui uno dei problemi principali sarà proprio il trasferimento dei dati dai radiotelescopi installati in aree remote del Sudafrica e dell'Australia verso i centri europei di elaborazione e poi verso tutti i data center regionali distribuiti sul territorio.
Nel nostro Spoke abbiamo affrontato questi aspetti anche attraverso un work package specifico dedicato ai data lake, alle tecnologie cloud e ai sistemi di trasferimento dati, con l'obiettivo di sperimentare soluzioni efficaci per la gestione di questi flussi enormi di informazioni. Altre difficoltà sono legate allo sviluppo software. Oggi lavoriamo con sistemi eterogenei e complessi, e questo rende i modelli di programmazione estremamente articolati. Trent'anni fa il problema era sostanzialmente scegliere tra due diversi linguaggi di programmazione, Fortran e C, mentre oggi esiste una quantità enorme di linguaggi, framework e modelli. Alcuni sono più efficienti, altri più flessibili, altri ancora più adatti all'analisi dati, come Python. Questo rende complicato sia sviluppare sia mantenere i codici, perché basta un cambiamento tecnologico - per esempio il passaggio da GPU NVIDIA a un'altra architettura - per costringere a riscrivere porzioni significative del software. Il ricercatore medio dell'INAF non viene assunto per sviluppare codice ma per studiare galassie, esopianeti o fenomeni cosmologici, ma oggi questi aspetti computazionali stanno diventando sempre più centrali per fare ricerca. Il mantenimento dei codici scientifici è una sfida continua. Ed è proprio qui che il lavoro del Centro Nazionale è stato fondamentale.

Come è stato organizzato concretamente il lavoro dello Spoke 3? E che tipo di comunità scientifica e professionale siete riusciti a costruire attraverso queste attività?
Gheller: Lo Spoke 3 ci ha permesso di formare una nuova generazione di ricercatori e tecnologi capaci di lavorare sulle problematiche che ho descritto. Dal punto di vista organizzativo, il lavoro era suddiviso in diversi work package. Quattro erano dedicati alle attività scientifiche e tecnologiche - HPC, machine learning, gestione dati e visualizzazione - mentre un quinto aveva una funzione più di supporto, legata all'infrastruttura web e agli strumenti collaborativi. All'interno di questi work package

sono state sviluppate circa quaranta task differenti e ogni task aveva un referente scientifico e un referente operativo. Tutto il materiale sviluppato è stato reso pubblico attraverso i repository GitLab del Centro Nazionale. Parallelamente sono stati organizzati corsi di HPC e machine learning per formare ricercatori e tecnologi, oltre a workshop, seminari, conferenze e attività di disseminazione anche in contesti internazionali.
Una parte importante del lavoro è stata anche la promozione di queste attività in ambienti non strettamente astrofisici. Per esempio, partecipiamo regolarmente a Supercomputing Conference, la grande conferenza internazionale che si svolge ogni anno negli Stati Uniti, dove coordiniamo insieme a INFN e CINECA degli spazi dedicati alla presentazione delle attività legate al supercalcolo e al Centro Nazionale. Nel complesso è stato quindi un lavoro capillare, non solo sul piano tecnico e scientifico, ma anche su quello della formazione e della costruzione di una comunità con competenze specifiche.

Uno degli aspetti che emerge quando si parla del Centro Nazionale è la capacità di sviluppare un dialogo tra ricerca, industria e sviluppo tecnologico. Quanto è stato importante questo aspetto per lo Spoke 3? E che tipo di progetti sono nati da queste collaborazioni?
Gheller: Inizialmente lo Spoke 3 era composto quasi esclusivamente da enti di ricerca e università - INAF, INFN, Università di Tor Vergata, Università di Trieste e altri partner accademici - quindi il lavoro era orientato soprattutto verso la ricerca di base. Progressivamente, però, sono state integrate alcune attività con aziende e soggetti industriali, soprattutto attraverso gli innovation grant, ovvero progetti pensati proprio per favorire il trasferimento tecnologico tra enti di ricerca e imprese coinvolte nel Centro Nazionale. L'idea era costruire un ponte tra chi sviluppa innovazione e chi poi può applicarla concretamente nel tessuto industriale e sociale. Nel nostro caso abbiamo avuto diversi progetti importanti con collaborazioni che hanno coinvolto grandi realtà industriali e tecnologiche come Unipol, Leonardo, Thales Alenia Space, Intesa Sanpaolo e Sogei. L'obiettivo era capire se gli algoritmi e le tecnologie sviluppate nello Spoke potessero trovare applicazioni anche fuori dall'astrofisica. Un esempio significativo è stato il progetto HaMMon, coordinato da Unipol ma con un forte coinvolgimento dell'INAF e dello Spoke 3. In una delle attività che ho seguito direttamente abbiamo preso algoritmi di machine learning sviluppati originariamente per identificare sorgenti nelle immagini radioastronomiche. Gli stessi approcci sono stati poi riutilizzati per analizzare il territorio urbano. Questi metodi, basati su reti neurali e tecniche avanzate di machine learning, sono stati applicati anche all'identificazione e classificazione dell'edificato in aree urbane come Bologna e Milano, con l'obiettivo di valutare possibili vulnerabilità legate a eventi meteorologici estremi o altri rischi naturali. È stato un caso interessante perché la stessa expertise sviluppata in ambito radioastronomico è stata trasferita quasi direttamente su un problema completamente diverso. Sempre nel progetto HaMMon sono stati sviluppati sistemi basati su immagini raccolte da droni per analizzare gli edifici da vicino, ricostruire modelli 3D e identificare automaticamente potenziali criticità strutturali - crepe, instabilità o altri segnali di rischio - attraverso tecniche di machine learning nate originariamente per l'analisi astronomica. Un altro esempio interessante riguarda un progetto sviluppato con il settore bancario, in collaborazione con Intesa Sanpaolo. In astronomia utilizziamo algoritmi sofisticati per identificare fenomeni estremamente rari e transienti - ad esempio intense emissioni luminose improvvise e brevissime provenienti dallo spazio profondo, come gamma ray burst o fast radio burst - che possono essere confusi e non venire individuati a causa del rumore osservativo. Gli stessi approcci sono stati utilizzati per individuare operazioni bancarie sospette o fraudolente all'interno di enormi volumi di transazioni normali. Anche qui il parallelismo è forte: trovare un evento raro e significativo nascosto dentro milioni di eventi ordinari.
Accanto agli innovation grant ci sono stati poi anche i bandi a cascata. In quel caso non si trattava tanto di trasferimento tecnologico verso l'industria, quanto piuttosto di coinvolgere aziende capaci di

risolvere problemi tecnici emersi durante le attività dello Spoke. Per esempio, nel campo della visualizzazione abbiamo collaborato con aziende che si sono occupate di reingegnerizzare i software sviluppati dai ricercatori, rendendoli più robusti, portabili e manutenibili. Il ricercatore sviluppa spesso codici molto efficaci dal punto di vista scientifico, ma non necessariamente progettati secondo tutti gli standard dell'ingegneria software. In questo senso il contributo industriale è stato strategico.

All'interno del Centro Nazionale c'è stato anche un dialogo tra i diversi Spoke? In particolare, avete collaborato con altri gruppi che lavorano su temi a voi vicini?
Gheller: Sì, anche se con modalità diverse a seconda dei progetti. La collaborazione più naturale è stata sicuramente quella con lo Spoke 2, dedicato ai problemi computazionali della ricerca fondamentale e alle loro applicazioni nella Space Economy. In un certo senso, lo Spoke 2 e lo Spoke 3 erano quasi "gemelli": nello Spoke 3 il ruolo di responsabile scientifico del progetto era affidato a INAF, con INFN come co-responsabile, mentre nello Spoke 2 avveniva l'opposto. Entrambi lavoriamo su tematiche legate alla fisica fondamentale e quindi c'erano molte sovrapposizioni naturali tra problemi di fisica delle alte energie e problematiche astrofisiche. C'è stata poi una collaborazione rilevante con lo Spoke 1, dedicato alle tecnologie HPC innovative, e con lo Spoke 10 sul quantum computing. In questi casi il dialogo era quasi inevitabile, perché molte delle problematiche affrontate erano comuni: architetture di calcolo, modelli di programmazione, accelerazione, ottimizzazione dei codici. Con altri Spoke il rapporto è stato un po' meno diretto sul piano scientifico, perché lavorano su temi più distanti dalla fisica e dall'astrofisica. Però ci sono comunque stati punti di contatto, soprattutto nel contesto degli innovation grant. Per esempio, con lo Spoke 5 abbiamo collaborato all'interno del progetto HaMMon, legato ai temi del territorio e dei rischi naturali.
Le collaborazioni quindi ci sono state, anche se spesso più su progettualità specifiche e applicative che non attraverso una cooperazione scientifica continua tra tutti gli Spoke. Ritengo comunque che il tentativo del Centro Nazionale di costruire connessioni tra competenze e comunità che normalmente lavorano in settori separati tra loro sia stato importante.

Guardando al futuro, quali sono oggi le priorità per dare continuità al lavoro avviato nello Spoke 3 e alle competenze costruite attraverso il Centro Nazionale?
Gheller: Credo che ormai una cosa sia abbastanza chiara: le tecnologie legate al supercalcolo, all'intelligenza artificiale e più in generale al calcolo avanzato sono diventate assolutamente imprescindibili per la ricerca scientifica. Questo messaggio sta arrivando progressivamente agli enti di ricerca e alle università, anche se non è ancora del tutto interiorizzato. Tradizionalmente, un dottorando, un assegnista o un ricercatore vengono formati soprattutto per affrontare uno specifico problema scientifico. Però oggi, per lavorare davvero su queste sfide, è necessario avere anche competenze specialistiche legate all'HPC, all'intelligenza artificiale, al machine learning e, in prospettiva, al quantum computing. Non significa che ogni astronomo debba diventare un esperto di programmazione, ma si tratta di formare persone capaci di muoversi sia nel linguaggio scientifico sia in quello delle tecnologie computazionali avanzate. Queste figure sono fondamentali perché fanno da collegamento tra mondi diversi: devono essere in grado di parlare con gli scienziati, capire i problemi scientifici reali e allo stesso tempo padroneggiare strumenti tecnologici avanzati. Senza questo tipo di professionalità sarà difficile continuare a progredire scientificamente.
Per queste ragioni stiamo cercando di dare continuità a quanto costruito nello Spoke 3. L'idea è trasformare questa esperienza in qualcosa di più stabile e strutturato, come un vero laboratorio permanente, con una prospettiva di lungo periodo e non limitata alla singola progettualità del PNRR. Un laboratorio che possa sia supportare gli scienziati nelle loro attività di ricerca, sia continuare a

sviluppare azioni di trasferimento tecnologico verso l'industria. Un altro punto fondamentale riguarda il personale formato in questi anni. Abbiamo investito molto nella costruzione di competenze e sarebbe un errore disperdere questo patrimonio. Il rischio, altrimenti, è perdere in poco tempo le professionalità che sono state costruite con grande fatica e che oggi sono sempre più necessarie per fare ricerca avanzata. La speranza è quindi riuscire a consolidare queste figure professionali nell'ambito delle future evoluzioni del Centro Nazionale.

Dal punto di vista scientifico, invece, quali trasformazioni si aspetta nei prossimi anni grazie all'uso di queste tecnologie?
Gheller: Dal punto di vista scientifico le prospettive sono estremamente promettenti. Le nuove infrastrutture osservative che stanno entrando in funzione oggi - e che potranno essere sfruttate soltanto grazie a queste tecnologie di calcolo - produrranno una crescita esponenziale delle conoscenze. Basti pensare agli esopianeti: in pochi anni siamo passati da poche decine di pianeti conosciuti a migliaia. E questo è avvenuto non solo perché gli strumenti osservativi sono diventati più potenti, ma anche perché siamo diventati capaci di elaborare i dati in modo sempre più sofisticato.
Oggi capita spesso che dei dati raccolti venti o trent'anni fa vengano nuovamente analizzati con tecniche moderne e producano nuove scoperte. Questo significa che il valore scientifico dei dati cresce nel tempo insieme alla nostra capacità di interpretarli. Il futuro è promettente, ma è importante capire che ormai il calcolatore è diventato parte integrante dello strumento scientifico stesso. Il telescopio non è più separabile dall'infrastruttura computazionale che permette di analizzare i dati. E quel calcolatore non è più il semplice PC del ricercatore: richiede competenze tecniche specialistiche, infrastrutture avanzate e comunità scientifiche capaci di far evolvere continuamente questi strumenti.