Cosmological hydrodynamical simulations: Illustris and IllustrisTNG data release platform

www.tng-project.org/data

Dylan Nelson (ITA, Heidelberg University)
IVOA Meeting, May 2021

Emmy Noether
DFG

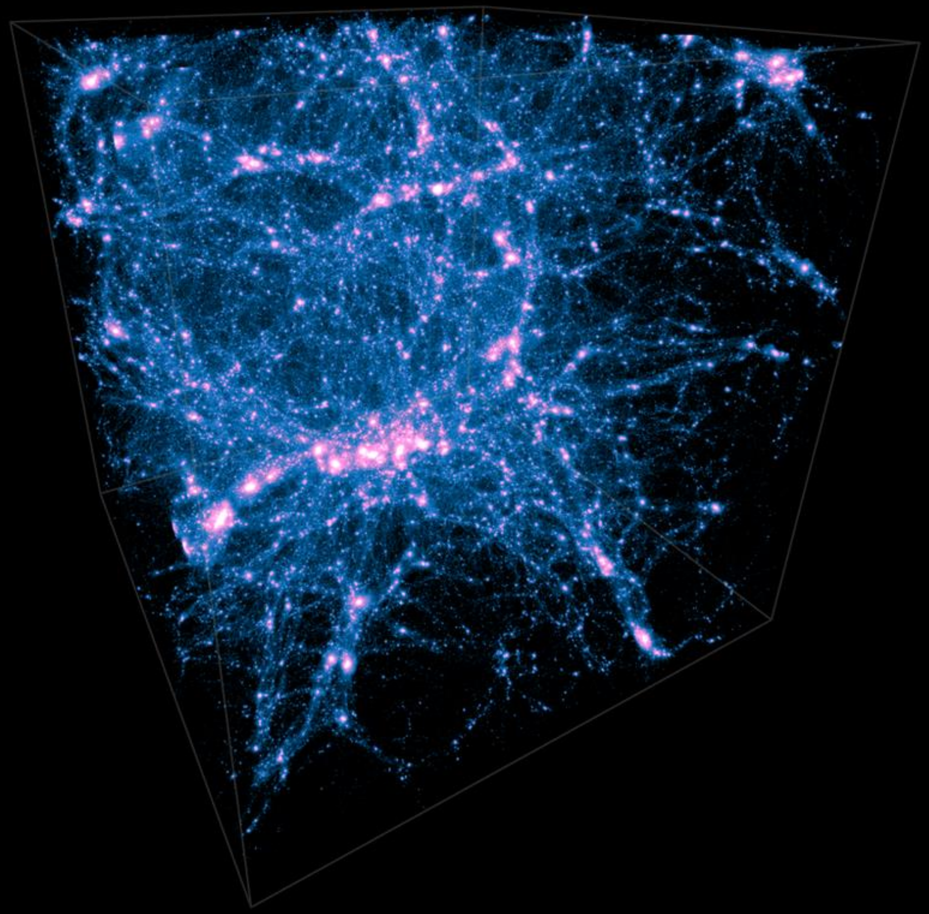
UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Cosmological hydro simulations:

[Illustris, Eagle, TNG, Horizon-AGN, Magneticum, SIMBA, ...]

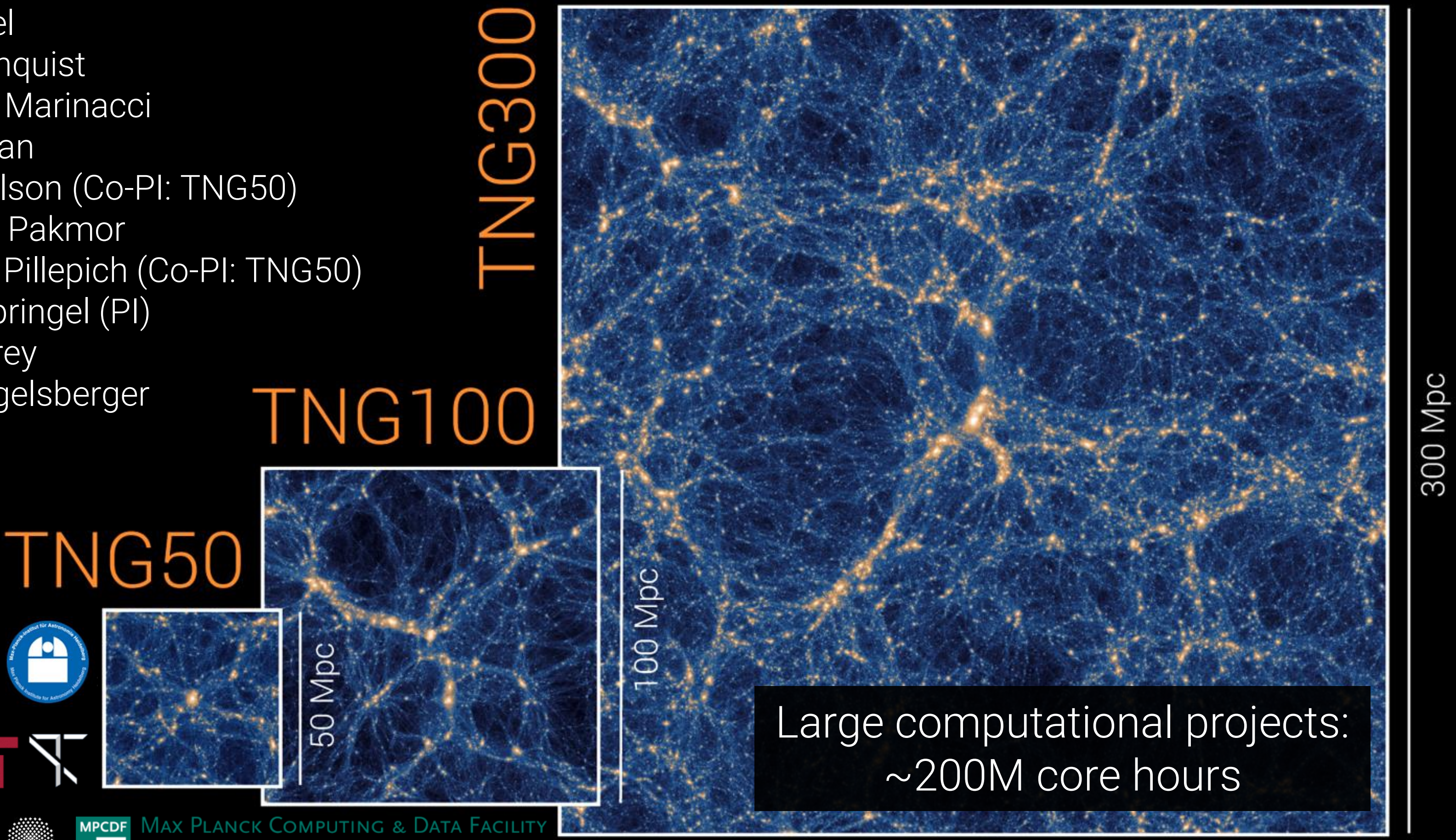
- "Universe in a Box" – periodic cube
- Snapshots: particle-level data
  - Gas
  - Dark matter
  - Stars
  - Black holes
- Catalogs: identified structures
  - Dark matter halos
  - Galaxies
- (x, y, z, property0, property1, ...)
- ~100 full time save points, few TB each

- Heterogeneous supplementary data products
  - Sub-volume snapshots (thousands)
  - Merger/assembly trees
  - Multiple substructure/galaxy catalogs
  - Observational mocks, images, lightcones, ...
  - Value-added catalogs of user-computed values



The IllustrisTNG project: the current state-of-the-art for cosmological simulations.

- Shy Genel
- Lars Hernquist
- Federico Marinacci
- Jill Naiman
- Dylan Nelson (Co-PI: TNG50)
- Ruediger Pakmor
- Annalisa Pillepich (Co-PI: TNG50)
- Volker Springel (PI)
- Paul Torrey
- Mark Vogelsberger



Large computational projects:
~200M core hours

Max-Planck-Institut für Astrophysik

HITS CfA

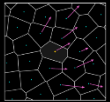
GCS Gauss Centre for Supercomputing

HLRS

MPCDF Max Planck Computing & Data Facility
Rechenzentrum Garching der Max-Planck-Gesellschaft

log10(M⋆/M⊙) = 9.69
r1/2, ⋆ = 2.31 kpc
5.9 kpc
log10(M⋆/M⊙) = 10.41
r1/2, ⋆ = 9.44 kpc
14.1 kpc
log10(M⋆/M⊙) = 11.69
r1/2, ⋆ = 8.72 kpc
13.1 kpc
log10(M⋆/M⊙) = 10.61
r1/2, ⋆ = 6.22 kpc
9.3 kpc
log10(M⋆/M⊙) = 11.06
r1/2, ⋆ = 9.47 kpc
14.2 kpc
log10(M⋆/M⊙) = 9.76
r1/2, ⋆ = 8.01 kpc
12.0 kpc
log10(M⋆/M⊙) = 12.06
r1/2, ⋆ = 13.06 kpc
19.6 kpc
log10(M⋆/M⊙) = 11.01
r1/2, ⋆ = 11.41 kpc
17.1 kpc
log10(M⋆/M⊙) = 11.29
r1/2, ⋆ = 7.69 kpc
11.5 kpc
log10(M⋆/M⊙) = 10.78
r1/2, ⋆ = 8.07 kpc
12.1 kpc
log10(M⋆/M⊙) = 10.17
r1/2, ⋆ = 5.42 kpc
8.1 kpc
log10(M⋆/M⊙) = 11.33
r1/2, ⋆ = 8.04 kpc
12.1 kpc
log10(M⋆/M⊙) = 10.01
r1/2, ⋆ = 9.38 kpc
14.1 kpc
log10(M⋆/M⊙) = 11.18
r1/2, ⋆ = 10.82 kpc
16.2 kpc
log10(M⋆/M⊙) = 10.70
r1/2, ⋆ = 5.24 kpc
7.8 kpc
log10(M⋆/M⊙) = 10.49
r1/2, ⋆ = 3.66 kpc
5.9 kpc
log10(M⋆/M⊙) = 10.85
r1/2, ⋆ = 11.44 kpc
17.2 kpc
log10(M⋆/M⊙) = 10.72
r1/2, ⋆ = 9.60 kpc
14.4 kpc
log10(M⋆/M⊙) = 11.27
r1/2, ⋆ = 5.95 kpc
8.9 kpc
log10(M⋆/M⊙) = 11.40
r1/2, ⋆ = 7.10 kpc
10.6 kpc
log10(M⋆/M⊙) = 10.21
r1/2, ⋆ = 5.42 kpc
8.1 kpc
log10(M⋆/M⊙) = 10.00
r1/2, ⋆ = 5.30 kpc
8.0 kpc
log10(M⋆/M⊙) = 10.03
r1/2, ⋆ = 6.65 kpc
10.0 kpc
log10(M⋆/M⊙) = 11.44
r1/2, ⋆ = 13.98 kpc
20.9 kpc
log10(M⋆/M⊙) = 10.89
r1/2, ⋆ = 6.79 kpc
10.2 kpc
log10(M⋆/M⊙) = 10.72
r1/2, ⋆ = 5.58 kpc
8.4 kpc
log10(M⋆/M⊙) = 11.62
r1/2, ⋆ = 9.93 kpc
14.9 kpc
log10(M⋆/M⊙) = 10.95
r1/2, ⋆ = 8.85 kpc
13.2 kpc

$ h5ls -r snap_099.hdf5



tracers

gas

stars

```
/                        Group
/Config                  Group
/Header                  Group
/Parameters              Group
/PartType0               Group
/PartType0/CenterOfMass  Dataset {14454722210, 3}
/PartType0/Coordinates   Dataset {14454722210, 3}
/PartType0/Density       Dataset {14454722210}
/PartType0/ElectronAbundance Dataset {14454722210}
/PartType0/EnergyDissipation Dataset {14454722210}
/PartType0/GFM_AGNRadiation Dataset {14454722210}
/PartType0/GFM_CoolingRate Dataset {14454722210}
/PartType0/GFM_Metallicity Dataset {14454722210}
/PartType0/GFM_Metals    Dataset {14454722210, 10}
/PartType0/GFM_MetalsTagged Dataset {14454722210, 6}
/PartType0/GFM_WindDMVelDisp Dataset {14454722210}
/PartType0/GFM_HostHaloMass Dataset {14454722210}
/PartType0/InternalEnergy Dataset {14454722210}
/PartType0/InternalEnergyOld Dataset {14454722210}
/PartType0/Machnumber    Dataset {14454722210}
/PartType0/MagneticField Dataset {14454722210, 3}
/PartType0/MagneticFieldDiv Dataset {14454722210}
/PartType0/Masses        Dataset {14454722210}
/PartType0/NeutralHydrogen Dataset {14454722210}
/PartType0/ParticleIDs   Dataset {14454722210}
/PartType0/Potential     Dataset {14454722210}
/PartType0/StarFormationRate Dataset {14454722210}
/PartType0/SubfindDMDensity Dataset {14454722210}
/PartType0/SubfindDensity Dataset {14454722210}
/PartType0/SubfindHsml   Dataset {14454722210}
/PartType0/SubfindVelDisp Dataset {14454722210}
/PartType0/Velocities    Dataset {14454722210, 3}
/PartType1               Group
/PartType1/Coordinates   Dataset {15625000000, 3}
/PartType1/ParticleIDs   Dataset {15625000000}
/PartType1/Potential     Dataset {15625000000}
/PartType1/SubfindDMDensity Dataset {15625000000}
/PartType1/SubfindDensity Dataset {15625000000}
/PartType1/SubfindHsml   Dataset {15625000000}
/PartType1/SubfindVelDisp Dataset {15625000000}
/PartType1/Velocities    Dataset {15625000000, 3}
```

dark matter

```
/PartType3               Group
/PartType3/ParentID      Dataset {15625000000}
/PartType3/TracerID      Dataset {15625000000}
/PartType4               Group
/PartType4/BirthPos      Dataset {711967677, 3}
/PartType4/BirthVel      Dataset {711967677, 3}
/PartType4/Coordinates   Dataset {711967677, 3}
/PartType4/GFM_InitialMass Dataset {711967677}
/PartType4/GFM_Metallicity Dataset {711967677}
/PartType4/GFM_Metals    Dataset {711967677, 10}
/PartType4/GFM_MetalsTagged Dataset {711967677, 6}
/PartType4/GFM_StarFormationTime Dataset {711967677}
/PartType4/GFM_StarPhotometrics Dataset {711967677, 8}
/PartType4/Masses        Dataset {711967677}
/PartType4/ParticleIDs   Dataset {711967677}
/PartType4/Potential     Dataset {711967677}
/PartType4/StellarHsml   Dataset {711967677}
/PartType4/SubfindDMDensity Dataset {711967677}
/PartType4/SubfindDensity Dataset {711967677}
/PartType4/SubfindHsml   Dataset {711967677}
/PartType4/SubfindVelDisp Dataset {711967677}
/PartType4/Velocities    Dataset {711967677, 3}
/PartType5               Group
/PartType5/BH_BPressure  Dataset {701122}
/PartType5/BH_CumEgyInjection_QM Dataset {701122}
/PartType5/BH_CumEgyInjection_RM Dataset {701122}
/PartType5/BH_CumMassGrowth_QM Dataset {701122}
/PartType5/BH_CumMassGrowth_RM Dataset {701122}
/PartType5/BH_Density    Dataset {701122}
/PartType5/BH_HostHaloMass Dataset {701122}
/PartType5/BH_Hsml       Dataset {701122}
/PartType5/BH_Mass       Dataset {701122}
/PartType5/BH_Mdot       Dataset {701122}
/PartType5/BH_MdotBondi  Dataset {701122}
/PartType5/BH_MdotEddington Dataset {701122}
/PartType5/BH_Pressure   Dataset {701122}
/PartType5/BH_Progs      Dataset {701122}
/PartType5/BH_U          Dataset {701122}
/PartType5/Coordinates   Dataset {701122, 3}
/PartType5/Masses        Dataset {701122}
/PartType5/ParticleIDs   Dataset {701122}
/PartType5/Potential     Dataset {701122}
/PartType5/SubfindDMDensity Dataset {701122}
/PartType5/SubfindDensity Dataset {701122}
/PartType5/SubfindHsml   Dataset {701122}
/PartType5/SubfindVelDisp Dataset {701122}
/PartType5/Velocities    Dataset {701122, 3}
```

black holes

$ h5ls -r groups_099.hdf5

dark & luminous subhalos (galaxies)

```
/                        Group
/Config                  Group
/Subhalo                 Group
/Subhalo/SubhaloBHMass   Dataset {14485709}
/Subhalo/SubhaloBHMdot   Dataset {14485709}
/Subhalo/SubhaloBfldDisk Dataset {14485709}
/Subhalo/SubhaloBfldHalo Dataset {14485709}
/Subhalo/SubhaloCM       Dataset {14485709, 3}
/Subhalo/SubhaloFlag     Dataset {14485709}
/Subhalo/SubhaloGasMetalFractions Dataset {14485709, 10}
/Subhalo/SubhaloGasMetalFracSfr Dataset {14485709, 10}
/Subhalo/SubhaloGasMetalFracSfrWeight Dataset {14485709, 10}
/Subhalo/SubhaloGasMetallicity Dataset {14485709}
/Subhalo/SubhaloGasMetallicityHalfRad Dataset {14485709}
/Subhalo/SubhaloGasMetallicitySfr Dataset {14485709}
/Subhalo/SubhaloGasMetallicitySfrWeight Dataset {14485709}
/Subhalo/SubhaloGrNr     Dataset {14485709}
/Subhalo/SubhaloHalfmassRad Dataset {14485709}
/Subhalo/SubhaloHalfmassRadType Dataset {14485709, 6}
/Subhalo/SubhaloIDMostbound Dataset {14485709}
/Subhalo/SubhaloLen      Dataset {14485709}
/Subhalo/SubhaloLenType  Dataset {14485709, 6}
/Subhalo/SubhaloMass     Dataset {14485709}
/Subhalo/SubhaloMassInHalfRad Dataset {14485709}
/Subhalo/SubhaloMassInHalfRadType Dataset {14485709, 6}
/Subhalo/SubhaloMassInMaxRad Dataset {14485709}
/Subhalo/SubhaloMassInMaxRadType Dataset {14485709, 6}
/Subhalo/SubhaloMassInRad Dataset {14485709}
/Subhalo/SubhaloMassInRadType Dataset {14485709, 6}
/Subhalo/SubhaloMassType Dataset {14485709, 6}
/Subhalo/SubhaloParent   Dataset {14485709}
/Subhalo/SubhaloPos      Dataset {14485709, 3}
/Subhalo/SubhaloSFR      Dataset {14485709}
/Subhalo/SubhaloSFRinHalfRad Dataset {14485709}
/Subhalo/SubhaloSFRinRad Dataset {14485709}
/Subhalo/SubhaloSpin     Dataset {14485709, 3}
/Subhalo/SubhaloStarMetalFractions Dataset {14485709, 10}
/Subhalo/SubhaloStarMetalFractionsHalfRad Dataset {14485709, 10}
/Subhalo/SubhaloStarMetallicity Dataset {14485709}
/Subhalo/SubhaloStarMetallicityHalfRad Dataset {14485709}
/Subhalo/SubhaloStellarPhotometrics Dataset {14485709, 8}
/Subhalo/SubhaloStellarPhotometricsMassInRad Dataset {14485709}
/Subhalo/SubhaloStellarPhotometricsRad Dataset {14485709}
/Subhalo/SubhaloVel      Dataset {14485709, 3}
/Subhalo/SubhaloVelDisp  Dataset {14485709}
/Subhalo/SubhaloVmax     Dataset {14485709}
/Subhalo/SubhaloVmaxRad  Dataset {14485709}
/Subhalo/SubhaloWindMass Dataset {14485709}
```

```
/Group                   Group
/Group/GroupBHMass       Dataset {17625892}
/Group/GroupBHMdot       Dataset {17625892}
/Group/GroupCM           Dataset {17625892, 3}
/Group/GroupFirstSub     Dataset {17625892}
/Group/GroupGasMetalFrac Dataset {17625892, 10}
/Group/GroupGasMetallicity Dataset {17625892}
/Group/GroupLen          Dataset {17625892}
/Group/GroupLenType      Dataset {17625892, 6}
/Group/GroupMass         Dataset {17625892}
/Group/GroupMassType     Dataset {17625892, 6}
/Group/GroupNsubs        Dataset {17625892}
/Group/GroupPos          Dataset {17625892, 3}
/Group/GroupSFR          Dataset {17625892}
/Group/GroupStarMetalFrac Dataset {17625892, 10}
/Group/GroupStarMetallicity Dataset {17625892}
/Group/GroupVel          Dataset {17625892, 3}
/Group/GroupWindMass     Dataset {17625892}
/Group/Group_M_Crit200   Dataset {17625892}
/Group/Group_M_Crit500   Dataset {17625892}
/Group/Group_M_Mean200   Dataset {17625892}
/Group/Group_M_TopHat200 Dataset {17625892}
/Group/Group_R_Crit200   Dataset {17625892}
/Group/Group_R_Crit500   Dataset {17625892}
/Group/Group_R_Mean200   Dataset {17625892}
/Group/Group_R_TopHat200 Dataset {17625892}
/Header                  Group
/Parameters              Group
```

dark matter halos

Download Simulation Data

Simulation Name	$L_{box}[Mpc]$	N_{DM}	$m_{DM}\ [M_\odot]$	$m_{gas}\ [M_\odot]$	N_{snap}	$N_{Subfind}(z=0)$	Snaps	FoF	Subfind	SubLink	LHaloTree
TNG100-1	110.7	1820^3	7.5×10^6	1.4×10^6	100	4371211	✔	✔	✔	✔	✔
TNG100-1-Dark	110.7	1820^3	8.9×10^6	0	100	5012155	✔	✔	✔	✔	✔
TNG100-2	110.7	910^3	6.0×10^7	1.1×10^7	100	698336	✔	✔	✔	✔	✔
TNG100-2-Dark	110.7	910^3	7.1×10^7	0	100	757923	✔	✔	✔	✔	✔
TNG100-3	110.7	455^3	4.8×10^8	8.9×10^7	100	118820	✔	✔	✔	✔	✔
TNG100-3-Dark	110.7	455^3	5.7×10^8	0	100	116020	✔	✔	✔	✔	✔
TNG300-1	302.6	2500^3	5.9×10^7	1.1×10^7	100	14485709	✔	✔	✔	✔	✔
TNG300-1-Dark	302.6	2500^3	7.0×10^7	0	100	15724587	✔	✔	✔	✔	✔
TNG300-2	302.6	1250^3	4.7×10^8	8.8×10^7	100	2471022	✔	✔	✔	✔	✔
TNG300-2-Dark	302.6	1250^3	5.6×10^8	0	100	2411570	✔	✔	✔	✔	✔
TNG300-3	302.6	625^3	3.8×10^9	7.0×10^8	100	391144	✔	✔	✔	✔	✔
TNG300-3-Dark	302.6	625^3	4.5×10^9	0	100	372177	✔	✔	✔	✔	✔
TNG50-1	51.7	2160^3	4.5×10^5	8.5×10^4	100	5688113	✔	✔	✔	✔	✔
TNG50-1-Dark	51.7	2160^3	5.4×10^5	0	100	6616159	✔	✔	✔	✔	✔
TNG50-2	51.7	1080^3	3.6×10^6	6.8×10^5	100	859077	✔	✔	✔	✔	✔
TNG50-2-Dark	51.7	1080^3	4.3×10^6	0	100	991928	✔	✔	✔	✔	✔
TNG50-3	51.7	540^3	2.9×10^7	5.4×10^6	100	134779	✔	✔	✔	✔	✔
TNG50-3-Dark	51.7	540^3	3.4×10^7	0	100	150309	✔	✔	✔	✔	✔
TNG50-4	51.7	270^3	2.3×10^8	4.3×10^7	100	22869	✔	✔	✔	✔	✔
TNG50-4-Dark	51.7	270^3	2.8×10^8	0	100	22963	✔	✔	✔	✔	✔
Illustris-1	106.5	1820^3	6.3×10^6	1.3×10^6	134	4366546	✔	✔	✔	✔	✖
Illustris-1-Dark	106.5	1820^3	7.5×10^6	0	136	4872374	✔	✔	✔	✔	✔
Illustris-2	106.5	910^3	5.0×10^7	1.0×10^7	136	689785	✔	✔	✔	✔	✔
Illustris-2-Dark	106.5	910^3	6.0×10^7	0	136	735751	✔	✔	✔	✔	✖
Illustris-3	106.5	455^3	4.0×10^8	8.1×10^7	136	121209	✔	✔	✔	✔	✔
Illustris-3-Dark	106.5	455^3	4.8×10^8	0	136	111992	✔	✔	✔	✔	✖

In total have **30** runs with **3004** total snapshots and **954.0 TB** total data volume, including **9,450,798,626** FoF groups, **8,970,441,171** Subfind groups, and **15,181,158,395,371** particles.

Simulation Name	Parent Sim	$L_{box}[Mpc]$	$N_{DM}(z=0)$	$m_{DM}\ [M_\odot]$	N_{snap}	Snaps	FoF	Subfind
Illustris-1-Subbox0	Illustris-1	10.7	33629705	6.3×10^6	3970	✔	✖	✖
Illustris-1-Subbox1	Illustris-1	11.4	4362068	6.3×10^6	3969	✔	✖	✖
Illustris-1-Subbox2	Illustris-1	7.1	1873205	6.3×10^6	3972	✔	✖	✖
Illustris-1-Subbox3	Illustris-1	7.1	1651059	6.3×10^6	3970	✔	✖	✖
Illustris-2-Subbox0	Illustris-2	10.7	4205674	5.0×10^7	2265	✔	✖	✖
Illustris-2-Subbox1	Illustris-2	11.4	544949	5.0×10^7	2265	✔	✖	✖
Illustris-2-Subbox2	Illustris-2	7.1	233917	5.0×10^7	2264	✔	✖	✖
Illustris-2-Subbox3	Illustris-2	7.1	206250	5.0×10^7	2265	✔	✖	✖
Illustris-3-Subbox0	Illustris-3	10.7	525682	4.0×10^8	1426	✔	✖	✖
Illustris-3-Subbox1	Illustris-3	11.4	68264	4.0×10^8	1426	✔	✖	✖
Illustris-3-Subbox2	Illustris-3	7.1	29095	4.0×10^8	1426	✔	✖	✖
Illustris-3-Subbox3	Illustris-3	7.1	25749	4.0×10^8	1426	✔	✖	✖
TNG100-1-Subbox0	TNG100-1	11.1	34833161	7.5×10^6	7908	✔	✖	✖
TNG100-1-Subbox1	TNG100-1	11.1	5915987	7.5×10^6	7908	✔	✖	✖
TNG100-2-Subbox0	TNG100-2	11.1	4353391	6.0×10^7	4380	✔	✖	✖
TNG100-2-Subbox1	TNG100-2	11.1	737680	6.0×10^7	4380	✔	✖	✖
TNG100-3-Subbox0	TNG100-3	11.1	543988	4.8×10^8	2431	✔	✖	✖
TNG100-3-Subbox1	TNG100-3	11.1	92277	4.8×10^8	2431	✔	✖	✖
TNG300-1-Subbox0	TNG300-1	22.1	89804035	5.9×10^7	2449	✔	✖	✖
TNG300-1-Subbox1	TNG300-1	22.1	26406751	5.9×10^7	2449	✔	✖	✖
TNG300-1-Subbox2	TNG300-1	14.8	1113796	5.9×10^7	2449	✔	✖	✖
TNG300-2-Subbox0	TNG300-2	22.1	11237893	4.7×10^8	3045	✔	✖	✖
TNG300-2-Subbox1	TNG300-2	22.1	3298609	4.7×10^8	3045	✔	✖	✖
TNG300-2-Subbox2	TNG300-2	14.8	139024	4.7×10^8	3045	✔	✖	✖
TNG300-3-Subbox0	TNG300-3	22.1	1403193	3.8×10^9	2050	✔	✖	✖
TNG300-3-Subbox1	TNG300-3	22.1	413272	3.8×10^9	2050	✔	✖	✖

Subboxes: in total have **39** runs with **118216** total snapshots and **409.2 TB** total data volume, including **0** FoF groups, **0** Subfind groups, and **3,449,350,532,922** particles.

total data volume ~1.4 PB

~20T total 'particles'

What physical properties exist, or can be computed?

Almost everything...

- Gas: Temperature
- Density
- Metallicity
- Star formation rate
- Ionization state, n_e, Hα, OIII, ...
- Magnetic field strength |B|, direction
- Metal abundances (H, He, C, N, O, Ne, Mg, Si, Fe)
- Entropy, pressure
- Neutral hydrogen
- Atomic hydrogen
- Molecular hydrogen, CO, CII
- Cooling rate, cooling time
- All metal ions (e.g. CLOUDY)
- X-ray emission (e.g. APEC)
- MgII, CIV, OVI, NeVIII absorption
- Absorption spectra (e.g. TRIDENT)
- Emission maps, Lyman-alpha
- Velocity fields / IFU datacubes
- Velocity, kinematics, line-of-sight/3D
- Lagrangian dynamical/thermal history

- Stars: mass
- Stellar age
- Stellar metallicity
- Stellar [optical] light (e.g. FSPS, BPASS)
- Stellar radiation field
- Metagalactic background radiation field
- SNII rate, SNIa rate, NS-NS merger rate
- Dust distribution
- Dust temperature (e.g. SKIRT)
- Dust [infrared] emission (e.g. SKIRT)
- Dust polarization
- Faraday rotation measure
- Dispersion measure (FRBs)
- Synchrotron emission
- Sunyaev-Zeldovich signal
- Mock HI data cubes (e.g. MARTINI)
- Mock strong/weak lensing
- 21cm
- Shock locations
- Shock Mach number
- Shock energy dissipation, γ-rays

- Black hole: mass
- Accretion rate
- Luminosity
- Feedback energy rates
- FB energy history
- Dark matter density
- DM annihilation
- DM velocity dispersion
- Gravitational potential

All within a galaxy or halo:

- Stellar mass
- Gas mass/fraction
- Size (Gas, Stars, DM)
- Morphology, shape
- Bars, spirals, warps
- Halo mass, c, z_{form}, ...
- Merger/assembly history
- SFR, sSFR, SFH
- Galaxy color
- Spin (Gas, Stars, DM)
- Environment, satellites
- Cosmic web structures
- Complete evo. history

TNG Data Platform

There are three fundamental ways you can work with the simulation data.

Local data, local analysis

(i) you can download the raw data files and example scripts in order to work completely on your local system.

Python
MATLAB
IDL

- If you have your own cluster & storage
- Good for complex & slow analysis
- Power users

Remote data, local analysis

(ii) you can use the online API to retrieve specific galaxies/halos of interest without needing to download any full datafiles.

Python
MATLAB
IDL

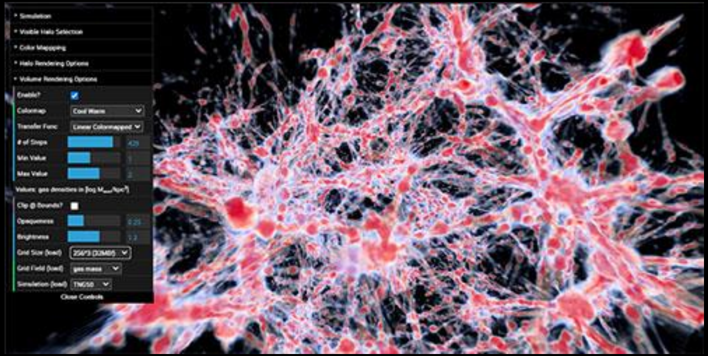
- REST
- Search and query online
- Download small pieces of data and catalogs
- Plotting on your laptop

Remote data, remote analysis

(iii) you can launch a web-based JupyterLab (or Jupyter notebook) session and explore the data, develop your analysis, run data- and compute-intensive tasks, and make final plots for publication.

Python	Julia
MATLAB	R
IDL	C/C++

- Quick exploration, prototyping
- Full particle-level data access without downloading
- Full paper reproducibility platform
- (JupyterHub-based containers, integrated auth, persistent storage)



Explorer3D

Interactive ray-traced volume rendering.

Explorer3D

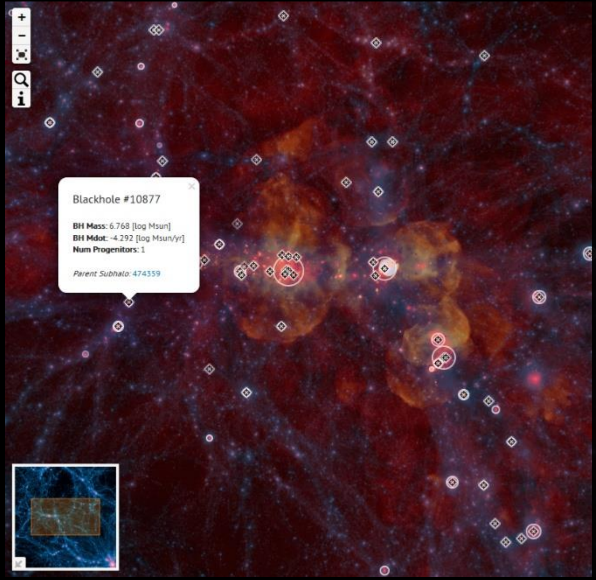
Interactive WebGL three.js, on-click API search.

online API

example applications

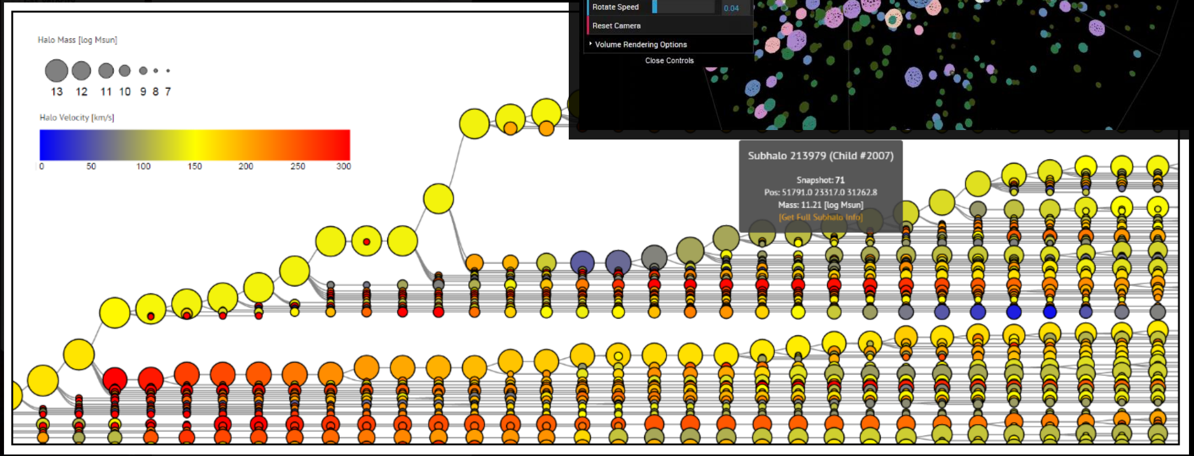
The Explorer

Leaflet tile-map renderer, spatial queries



Galaxy merger trees.

Interactive d3.js, each node links to its actual simu



time

Galaxy/Halo Search

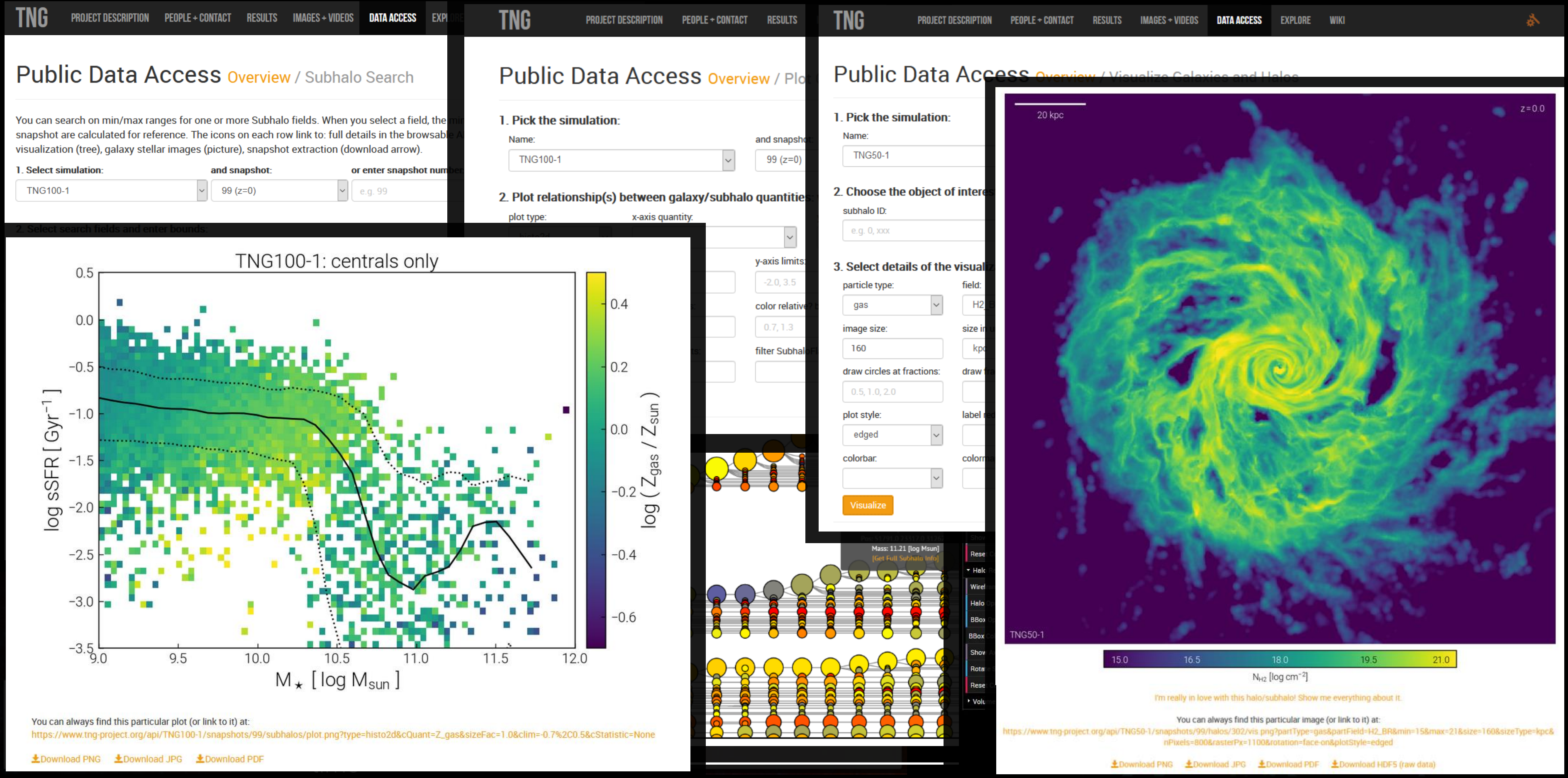
Convenient UI for quick catalog searches.

Plot Galaxy/Halo Catalog

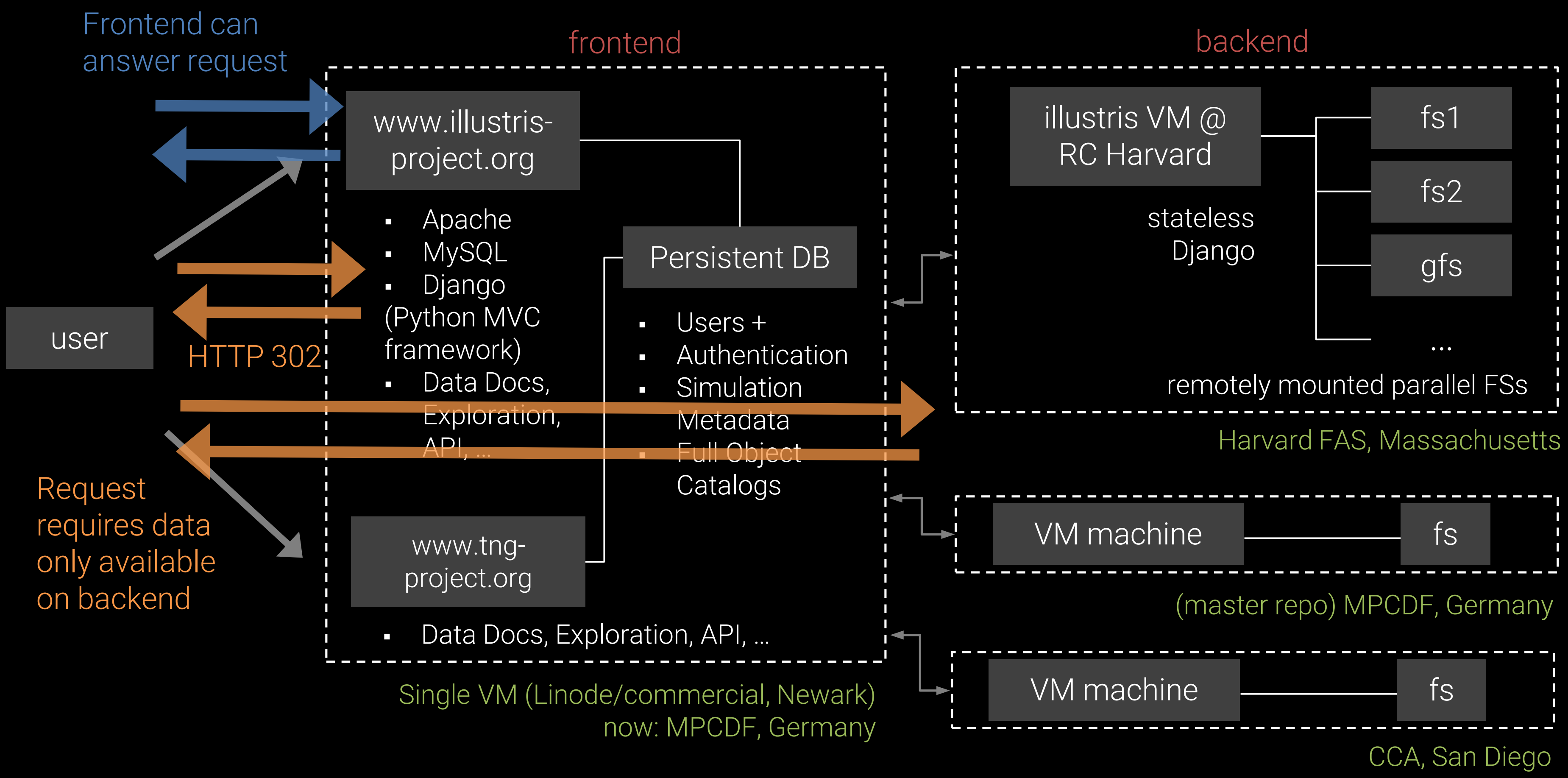
Study relationships and correlations between properties.

Visualize Galaxies and Halos

On-demand backend rendering.

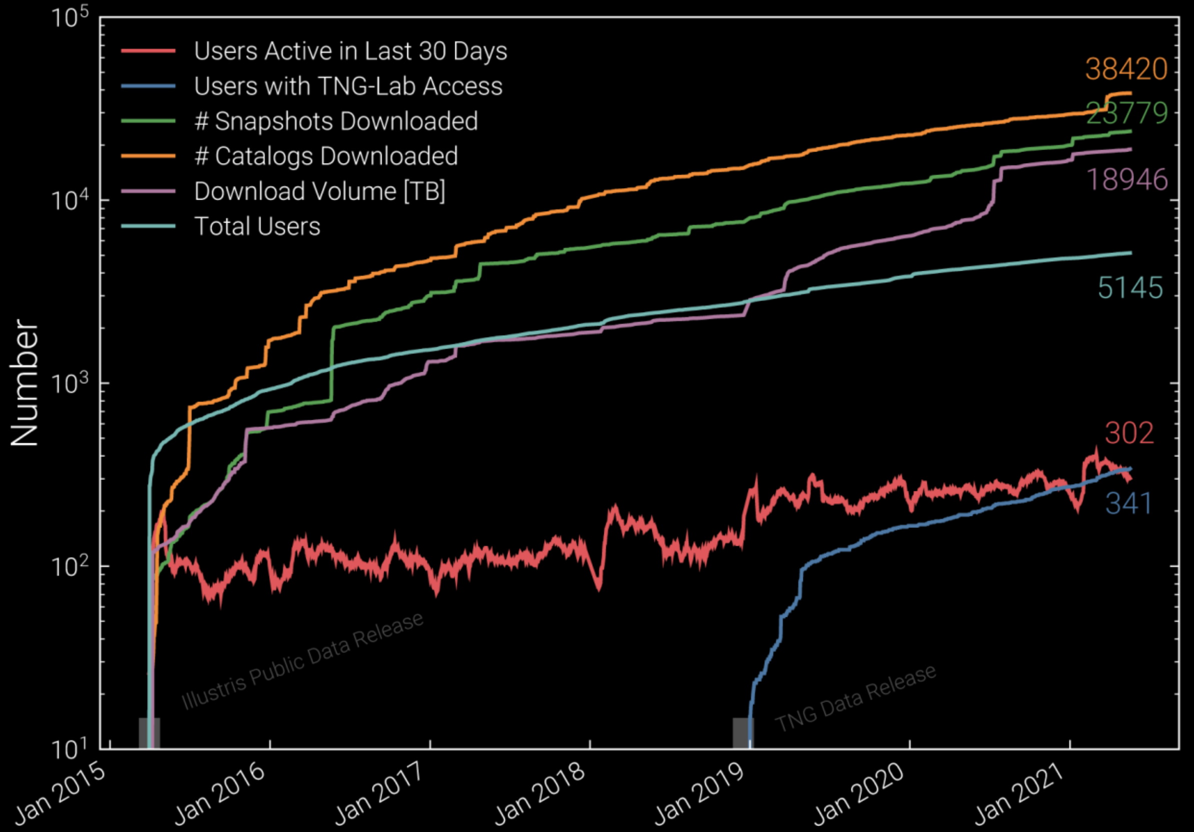


System architecture: (optional) division between data stores and user frontend.



Public data release platform for the Illustris & IllustrisTNG (50/100/300) simulations.

Field-leading example of a Open Data / Open Science initiative facilitates access to state-of-the-art theory by a global userbase of astrophysicists across a broad range of research topics.



Big data.

>10's k full snapshots & catalogs downloaded
Petabytes of data transferred

Significant community use.

More than 5,000 registered data users
450 million API requests

Usage is sustained and on-going.

Steady state level of ~200-300 active users

Web-based Lab service in demand.

341 approved users, 7,805 sessions to date

Simulation data = publications.

To date: **230** published papers with Illustris
278 published papers with TNG

MPCDF

Nelson+ (2015b)
Nelson+ (2019a)

Future directions & interoperability

- Current state of the world: every large simulation project handles their public data effort (not all participate).
- Goal: collect simulations under one (a few) public repositories, with at least some effort towards commonalities.
- Two levels of interop:
  (1) Simulation output “raw file” format, structure, metadata.
  (2) Web-accessible methods for accessing & analyzing simulations.

(1)

- Standards cannot be developed ahead of time or applied top-down.
- Always driven by computational/code considerations.
- Allow organic convergence of formats and approaches among the various groups.
- Machine readable metadata (G4, SWIFT).
- Vision 2020-2023: collect the main hydro simulations, re-write into *identical* common format at the file level, re-process with *uniform* pipelines, present with *identical* interface at the API level.

(2) www.tng-project.org/api/

www.tng-project.org/api/TNG100-1/
www.tng-project.org/api/TNG100-1/snapshots/68/
www.tng-project.org/api/TNG100-1/snapshots/99/subhalos/73664/
www.tng-project.org/api/TNG100-1/snapshots/99/subhalos/73664/stellar_mocks/broadband.fits
www.tng-project.org/api/TNG100-2/snapshots/80/halos/523312/cutout.hdf5?dm=Coordinates,Velocities&gas=Masses
www.tng-project.org/api/TNG100-2/snapshots/67/halos/100/vis.png?partType=gas&partField=temp
www.tng-project.org/api/TNG100-2/snapshots/50/subhalos/plot.pdf?xQuant=mstar2_log&yQuant=ssfr
www.tng-project.org/api/TNG300-1/snapshots/99/subhalos/?mass__gt=10.0&mass__lt=20.0
www.tng-project.org/api/TNG300-1/snapshots/68/subhalos/50000/sublink/full.hdf5
www.tng-project.org/api/EAGLE100/snapshots/28/subhalos/50000/sublink/mpb.json
www.tng-project.org/api/Millennium-1/files/groupcat-65.5.hdf5
www.tng-project.org/api/EAGLE100/files/snapshot-28.10.hdf5?dm=all

API metadata (json, ...)

On-disk files (HDF5, ...)

On-the-fly generated data (HDF5, PDF, ...)

All data is publicly available.

www.tng-project.org

/data

Most important aspect of the entire data release: [physical/model] documentation

PartType0 (gas)						
Field	**Full Snaps**	**Mini Snaps**	**Subbox Snaps**	**Dims**	**Units**	**Description**
CenterOfMass	✓	-	✓	N,3	$ckpc/h$	Spatial position of the center of mass, which in general differs from the geometrical center of the Voronoi cell (the offset should be small). Comoving coordinate.
Coordinates	✓	✓	✓	N,3	$ckpc/h$	Spatial position within the **periodic** simulation domain of BoxSize. Comoving coordinate.
Density	✓	✓	✓	N	$(10^{10} M_\odot/h)/(ckpc/h)^3$	Comoving mass density of cell (calculated as mass/volume).

PartType1 (dm)						
Field	**Full Snaps**	**Mini Snaps**	**Subbox Snaps**	**Dims**	**Units**	**Description**
Coordinates	✓	✓	✓	N,3	$ckpc/h$	Spatial position within the **periodic** simulation domain of BoxSize. Comoving coordinate.
ParticleIDs	✓	✓	✓	N	-	The unique ID (uint64) of this DM particle. Constant for the duration of the simulation.
Potential	✓	-	✓	N	$(km/s)^2/a$	Gravitational potential energy.

PartType4 (stars / wind particles)						
Field	**Full Snaps**	**Mini Snaps**	**Subbox Snaps**	**Dims**	**Units**	**Description**
BirthPos	✓	-	✓	N,3	$ckpc/h$	Spatial position within the periodic box where this star particle initially formed. Comoving coordinate.
BirthVel	✓	-	✓	N,3	$km\sqrt{a}/s$	Spatial velocity of the parent star-forming gas cell at the time of formation. The peculiar velocity is obtained by multiplying this value by $a^{1/2}$.
Coordinates	✓	✓	✓	N,3	$ckpc/h$	Spatial position within the **periodic** simulation domain of BoxSize. Comoving coordinate.
GFM_InitialMass	✓	✓	✓	N	$10^{10} M_\odot/h$	Mass of this star particle when it was formed (will subsequently decrease due to stellar evolution).
GFM_Metallicity	✓	✓	✓	N	-	See entry under PartType0. Inherited from the gas cell spawning/converted into this star, at the time of birth.

Use case #1: ‘example scripts’ (Python, Matlab, IDL, Julia) abstract away I/O complexities

Group Catalogs

First, make a base directory for the run and a subdirectory for the `z=0` group catalogs (snapshot 135), then download the catalog (~100 MB).

```
$ mkdir -p ~/Illustris-3/output/groups_135
$ cd ~/Illustris-3/output/groups_135/
$ wget -nd -nc -nv -e robots=off -l 1 -r -A hdf5 --content-disposition --header="API-Key: 10d143a0ef27c6461f94b50275d45d6f" "http://www.tng-project.org/api/Illustris-3/files/groupcat-135/?format=api"
```

Start up your interface of choice and import the example scripts:

```
$ python
>>> import illustris_python as il
>>>
```

Define the base path for the data (modify as needed), and load the total masses (SubhaloMass) and star formation rate within twice the stellar half mass radius (SubhaloSFRinRad) of all the Subfind subhalos.

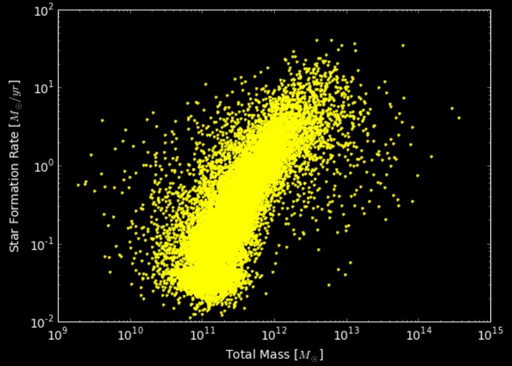
```
>>> basePath = './Illustris-3/output/'
>>> fields = ['SubhaloMass','SubhaloSFRinRad']
>>> subhalos = il.groupcat.loadSubhalos(basePath,135,fields=fields)
```

Inspecting the return, we see it is a dict/hash/struct with a key `count` indicating that there are **121209** total subhalos. Each requested field is returned as a numeric array with a key name equal to its field name in the group catalog.

```
>>> subhalos.keys()
['count', 'SubhaloSFRinRad', 'SubhaloMass']
>>> subhalos['SubhaloMass'].shape
(121209,)
```

Make a simple scatterplot of the relation (note the units).

```
>>> import matplotlib.pyplot as plt
>>> mass_msun = subhalos['SubhaloMass'] * 1e10 / 0.704
>>> plt.plot(mass_msun,subhalos['SubhaloSFRinRad'],'.')
>>> plt.xscale('log')
>>> plt.yscale('log')
>>> plt.xlabel('Total Mass [$M_\odot$]')
>>> plt.ylabel('Star Formation Rate [$M_\odot / yr$]')
```



Use case #2: enable the most important 'data volume reduction prior to transmission' tasks (search, extraction)

Task 2: for Illustris-1 at $z = 2$, search for all subhalos with total mass $10^{11.9} M_\odot < M < 10^{12.1} M_\odot$, print the number returned, and the Subfind IDs of the first five results (arbitrarily ordered, you may get different ids).

```
>>> # first convert log solar masses into group catalog units
>>> mass_min = 10**11.9 / 1e10 * 0.704
>>> mass_max = 10**12.1 / 1e10 * 0.704

>>> # form the search_query string by hand for once
>>> search_query = "?mass__gt=" + str(mass_min) + "&mass__lt=" + str(mass_max)
>>> search_query
'?mass__gt=55.9207077246&mass__lt=88.6283489903'

>>> # form the url and make the request
>>> url = "http://www.tng-project.org/api/Illustris-1/snapshots/z=2/subhalos/" + search_query
>>> subhalos = get(url)
>>> subhalos['count']
550

>>> ids = [ subhalos['results'][i]['id'] for i in range(5) ]
>>> ids
[109974, 110822, 123175, 107743, 95711]
```

Task 3: for Illustris-1 at $z = 2$, retrieve all fields for five specific Subfind IDs (from above: 109974, 110822, 123175, 107743, 95711), print the stellar mass and number of star particles in each.

```
>>> ids = [109974, 110822, 123175, 107743, 95711]
>>> for id in ids:
>>>     url = "http://www.tng-project.org/api/Illustris-1/snapshots/z=2/subhalos/" + str(id)
>>>     subhalo = get(url)
>>>     print id, subhalo['mass_stars'], subhalo['len_stars']

109974 0.283605 7270
110822 0.41813 5820
123175 0.529888 11362
107743 0.648827 10038
95711 0.623781 12722
```

Task 4: for Illustris-1 at $z = 2$, for five specific Subfind IDs (from above: 109974, 110822, 123175, 107743, 95711), extract and save full cutouts from the snapshot (HDF5 format).

```
>>> ids = [109974, 110822, 123175, 107743, 95711]
>>>
>>> for id in ids:
```

TNG-JupyterLab Interface