

The Target-selection Pipeline for the Dark Energy Spectroscopic Instrument

Adam D. Myers¹, John Moustakas², Stephen Bailey³, Benjamin A. Weaver⁴, Andrew P. Cooper⁵, Jaime E. Forero-Romero⁶, Bela Abolfathi⁷, David M. Alexander⁸, David Brooks⁹, Edmond Chaussidon¹⁰, Jaime E. Forero-Romero⁶, Bela Abolfathi⁷, David M. Alexander⁸, David Brooks⁹, Edmond Chaussidon¹⁰, Chia-Hsun Chuang^{11,12}, Kyle Dawson¹¹, Arjun Dey⁴, Biprateep Dey¹³, Govinda Dhungana¹⁴, Peter Doel⁹, Kevin Fanning^{15,16}, Enrique Gaztañaga¹⁷, Satya Gontcho A Gontcho^{3,18}, Alma X. Gonzalez-Morales^{19,20}, ChangHoon Hahn^{3,21}, Hiram K. Herrera-Alcantar²⁰, Klaus Honscheid¹⁵, Mustapha Ishak²², Tanveer Karim²³, David Kirkby⁷, Theodore Kisner³, Sergey E. Koposov^{24,25}, Anthony Kremin³, Ting-Wen Lan^{26,27}, Martin Landriau³, Dustin Lang²⁸, Michael E. Levi³, Christophe Magneville¹⁰, Lucas Napolitano¹, Paul Martini^{16,29,30}, Aaron Meisner⁴, Jeffrey A. Newman¹³, Nathalie Palanque-Delabrouille^{3,10}, Will Percival^{28,31}, Claire Poppett³², Francisco Prada³³, Anand Raichoor³, Ashley J. Ross¹⁶, Edward F. Schlafly³⁴, David Schlegel³, Michael Schubnell³⁵, Ting Tan³⁶, Gregory Tarle³⁵, Michael J. Wilson^{3,37}, Christophe Yèche¹⁰, Rongpu Zhou³, Zhimin Zhou³⁸, and Hu Zou³⁸, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA ⁴ NSF's National Optical-Infrared Astronomy Research Laboratory, 950 N. Cherry Avenue, Tucson, AZ 85719, USA ⁵ Institute of Astronomy and Department of Physics, National Tsing Hua University, 101 Kuang-Fu Rd. Section 2, Hsinchu 30013, Taiwan ⁶ Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Bogotá, Colombia ¹ Department of Physics and Astronomy, University of California, Irvine, Frederick Reines Hall, Irvine, CA 92697, USA Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France ¹¹ Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA ¹² Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA ¹³ Department of Physics and Astronomy and PITT-PACC, University of Pittsburgh, 100 Allen Hall, 3941 O'Hara Street, Pittsburgh, PA 15260, USA Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75205, USA ¹⁵ Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA ¹⁶ Center for Cosmology and AstroParticle Physics, The Ohio State University, Columbus, OH 43210, USA ¹⁷ Institut de Ciencies de l'Espai, IEEC-CSIC, Campus UAB, Carrer de Can Magrans s/n, E-08913 Bellaterra, Barcelona, Spain ¹⁸ Department of Physics and Astronomy, University of Rochester, 500 Joseph C. Wilson Boulevard, Rochester, NY 14627, USA ¹⁹ Consejo Nacional de Ciencia y Tecnología, Av. Insurgentes Sur 1582. Colonia Credito Constructor, Del. Benito Jurez C.P. 03940, México D.F., Mexico Departamento de Física, Universidad de Guanajuato-DCI, C.P. 37150, Leon, Guanajuato, Mexico ²¹ Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA ²³ Department of Physics, The University of Texas at Danas, Richardson, TX 70000, Cont.
 ²³ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
 ²⁴ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK
 ²⁵ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK
 ²⁶ Department of Physics, National Taiwan University, Taipei 10617, Taiwan
 ²⁷ Texas and Static Astronomy, University of Lawan University, Taipei 10617, Taiwan ²⁷ Graduate Institute of Astrophysics, National Taiwan University, Taipei 10617, Taiwan ²⁸ Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA ³⁰ Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA 02138, USA ³¹ Department of Physics and Astronomy, University of Waterloo, 200 University Avenue W, Waterloo, ON, N2L 3G1, Canada ³² Space Sciences Laboratory at University of California, 7 Gauss Way, Berkeley, CA 94720, USA ³³ Institute de Astrofísica de Andalucía, Glorieta de la Astronomía s/n, E-18008 Granada, Snain Instituto de Astrofisica de Andalucía, Glorieta de la Astronomía, s/n, E-18008 Granada, Spain ³⁴ Lawrence Livermore National Laboratory, P.O. Box 808 L-211, Livermore, CA 94551, USA ³⁵ Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, MI 48109, USA ³⁶ Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), F-75005 Paris, France Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK

³⁸ National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing, 100012, People's Republic of China Received 2022 August 22; revised 2022 November 3; accepted 2022 November 7; published 2023 January 11

Abstract

In 2021 May, the Dark Energy Spectroscopic Instrument (DESI) began a 5 yr survey of approximately 50 million total extragalactic and Galactic targets. The primary DESI dark-time targets are emission line galaxies, luminous red galaxies, and quasars. In bright time, DESI will focus on two surveys known as the Bright Galaxy Survey and the Milky Way Survey. DESI also observes a selection of "secondary" targets for bespoke science goals. This paper gives an overview of the publicly available pipeline (desitarget) used to process targets for DESI observations. Highlights include details of the different DESI survey targeting phases, the targeting ID (TARGETID) used to define unique targets, the bitmasks used to indicate a particular type of target, the data model and structure of DESI targeting files, and examples of how to access and use the desitarget code base. This

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. paper will also describe "supporting" DESI target classes, such as standard stars, sky locations, and random catalogs that mimic the angular selection function of DESI targets. The DESI target-selection pipeline is complex and sizable; this paper attempts to summarize the most salient information required to understand and work with DESI targeting data.

Unified Astronomy Thesaurus concepts: Catalogs (205); Sky surveys (1464); Cosmology (343); Observational cosmology (1146); Galaxy distances (590); Galaxies (573); Redshift surveys (1378); Quasars (1319); Milky Way Galaxy (1054); Astronomy data analysis (1858); Open source software (1866); Astronomy software (1855)

1. Introduction

As sky surveys have expanded in size and sophistication, the accompanying software has also become increasingly complex. Improvements in software pipelines are typically driven by, and proceed hand-in-hand with, augmentations in instrumentation and more ambitious scientific goals. Some examples of these algorithmic leaps include the development of techniques to perform spectroscopic reductions in near real-time to improve the completeness of redshift surveys (e.g., Tonry & Davis 1979), and the automated cataloging of galaxies using scanning machines in order to better sample and characterize large-scale structure (e.g., Loveday et al. 1992).

With the advent of multiobject spectrographs (e.g., Bagnuolo et al. 1990; Lewis et al. 2002; Smee et al. 2013), spectroscopic surveys have greatly expanded in scope. This has facilitated larger campaigns with objectives that require greater accuracy and precision. Surveys with lower shot noise are more prone to being affected by subtly heterogeneous source catalogs, and survey pipelines have had to became increasingly elaborate to mitigate inhomogeneity. Targeting for the 2dF Surveys, for example, incorporated color equations to account for different emulsions on glass plates versus films, and the survey pipeline generated masks to mitigate density fluctuations caused by bright stars, varying flux limits, satellite tracks, and plate defects (e.g., Colless et al. 2001, 2003; Croom et al. 2004; Smith et al. 2005).

The Sloan Digital Sky Survey (SDSS; York et al. 2000), in particular, greatly advanced the frontiers of software development for large sky surveys. Specialist pipelines were developed for imaging reductions (Lupton et al. 2001; Blanton et al. 2011), spectroscopic reductions, source classifications and redshift fitting (SubbaRao et al. 2002; Bolton et al. 2012), and tiling and fiber assignment (Blanton et al. 2003), among other goals. Crucially, the source code for these pipelines was made public, and the associated algorithms and data model were extensively documented (e.g., Stoughton et al. 2002).

Targeting pipelines also advanced with later iterations of the SDSS, such as the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) and the SDSS-IV extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016). The BOSS and eBOSS quasar targeting pipelines (Ross et al. 2012; Myers et al. 2015), for instance, not only incorporated long-established approaches such as color cuts and cross-matches to multiwavelength catalogs, but also implemented targeting techniques based on variability in multiepoch data (Palanque-Delabrouille et al. 2011, 2016), forced photometry (Lang et al. 2016), machine-learning methods (Yèche et al. 2010), and rigorous Bayesian approaches (Bovy et al. 2011; Kirkpatrick et al. 2011; Bovy et al. 2012).

The Dark Energy Spectroscopic Instrument (DESI) is a robotic, fiber-fed, highly multiplexed spectroscopic surveyor that operates on the Mayall 4 meter telescope at Kitt Peak National Observatory (DESI Collaboration et al. 2016a, 2022).

DESI, which can obtain simultaneous spectra of almost 5000 objects over a $\sim 3^{\circ}$ field (DESI Collaboration et al. 2016a; Silber et al. 2022; T. Miller et al. 2023, in preparation), is currently conducting a 5 yr survey of about a third of the sky. This campaign will obtain spectra for approximately 40 million galaxies and quasars (Levi et al. 2013; DESI Collaboration et al. 2016b), which will produce about an order-of-magnitude more extragalactic redshifts than measured by BOSS and eBOSS combined (see Ahumada et al. 2020, for BOSS and eBOSS source statistics). DESI spectra span wavelengths of \sim 3600–9800 Å with a blue-end resolution of \sim 2000, growing to \sim 5000 at the red end. Effective exposure times are \sim 1000 s for dark-time targets and ~ 180 s in bright time (see Section 4 of Guy et al. 2022). Quasar targets that DESI measures to be at redshifts of $z \ge 2.1$ are observed multiple times to improve signal-to-noise in the Ly α forest, and can ultimately amass \sim 4000 s of exposure time (see Section 4 of E. Schlafly et al. 2023, in preparation).

The sheer scale of the DESI experiment necessitates multiple supporting software pipelines and products. These include significant imaging from the DESI Legacy Imaging Surveys (Zou et al. 2017; Dey et al. 2019; D. Schlegel et al. 2023, in preparation), an extensive spectroscopic reduction pipeline (Guy et al. 2022), a template-fitting pipeline to derive classifications and redshifts for each targeted source (Redrock; S. Bailey et al. 2023, in preparation), a pipeline to assign fibers to targets (A. Raichoor et al. 2023, in preparation), a pipeline to tile the survey and to plan and optimize observations as the campaign progresses (E. Schlaffy et al. 2023, in preparation), and a pipeline to select targets for spectroscopic follow-up (desitarget; this paper).

Target-selection approaches for DESI are, themselves, varied and extensive. Other publications that accompany this work include a paper describing the DESI Survey Validation (SV) phase (DESI Collaboration et al. 2023, in preparation), two papers describing how visual inspection of spectra of targets acquired during SV produced truth tables to inform target selection for the DESI Main Survey (Alexander et al. 2022; Lan et al. 2022), and a series of five papers detailing the selection of DESI bright-time and dark-time science targets (Chaussidon et al. 2022; Zhou et al. 2022; see also Section 4.1.1).

In this paper, we detail the DESI target-selection pipeline, which we refer to throughout as desitarget. In Section 2 we give an overview of how DESI uses bitmasks to record which targets were selected by each classification algorithm. In Section 3 we detail the unique identification number (TARGE-TID) that DESI adopts to track targets. In Section 4 we give a technical overview of the main DESI target classes. In Section 5 we highlight some important known issues and caveats that should be considered when working with DESI targets. Finally, we present some closing thoughts in Section 6. Throughout this paper, code examples are given in the Python programming language and HEALPixel³⁹ nside numbers (Górski et al. 2005; Zonca et al. 2019) are always expressed in the NESTED scheme. The desitarget code base, as well as the entire history of its development, is publicly accessible.⁴⁰ Much of the data model for files produced by desitarget, including definitions of quantities in these files, is also available online.⁴¹

2. Bits and Bitmasks

The chief purpose of the desitarget pipeline is to determine which sources selected by a variety of algorithms will be targeted for follow-up spectroscopy by DESI. To record that a specific source has been selected by a particular targeting algorithm, desitarget uses a number of bitmasks. A tutorial to further help elucidate DESI bitmasks is available on GitHub.⁴²

The full set of targeting bitmasks assigned by desitarget will be tabulated in forthcoming DESI Data Release papers (e.g., DESI Collaboration et al. 2023, in preparation). Here, though, we will introduce (and refer to) some of the principal DESI targeting bitmasks to help illustrate the functionality of the desitarget pipeline.

2.1. An Overview of DESI Target Classes

The primary dark-time target classes observed by DESI (DESI Collaboration et al. 2016b) are emission line galaxies (ELGs; Raichoor et al. 2020, 2022), luminous red galaxies (LRGs; Zhou et al. 2020, 2022) and quasi-stellar objects (QSOs, also known as quasars; Yèche et al. 2020; Chaussidon et al. 2022a). During survey bright-time, DESI also targets galaxies as part of a dedicated Bright Galaxy Survey (BGS; Ruiz-Macias et al. 2020; Hahn et al. 2022), and a variety of Galactic sources as part of a Milky Way Survey (MWS; Allende Prieto et al. 2020; Cooper et al. 2022). Finally, to calibrate the output from its spectroscopic pipeline, DESI also targets calibration sources such as standard stars and blank-sky locations. These target classes will be described in more detail in Section 4.

The DESI survey also incorporates a plethora of programs that are not driven by the main cosmological goals of the primary campaign. These are referred to as "secondary" programs, and include, for example, studies of high-propermotion stars in our Galaxy, a campaign to target M31 (Dey et al. 2022), studies of peculiar velocities in local galaxies, broad censuses of extragalactic sources, follow-up of gravitational lenses, and identification of quasars at redshifts $z \gtrsim 5$ (J. Yang et al. 2023, in preparation). These target classes will be detailed in forthcoming DESI Data Release papers (e.g., DESI Collaboration et al. 2023, in preparation)—but we include some technical details about DESI secondary targets in Section 3.2 and Section 4.6.

All DESI primary targets, and the vast majority of secondary classes, are selected using properties from imaging. Based on algorithms applied to these photometric quantities, a bitmask (see Section 2.2) is constructed to designate a source as being a member of a particular target class or classes. Note that as the

bits corresponding to "QSO," "ELG," "LRG," etc., are assigned using only imaging quantities, subsequent DESI spectroscopy will reveal that some sources are inconsistent with their target class or classes, i.e., not all QSO targets will have quasar-like spectra.

2.2. A Brief Introduction to Bitmasks

As an example of how a bitmask is constructed, consider two types of source that DESI will target: LRGs and ELGs. Let us assume that LRGs are designated by bit 0 and ELGs by bit 1, and that the name of the variable used to store the targeting bitmask is DESI_TARGET. Then, DESI_TAR-GET = 1 ($\equiv 2^0$) would signify a source selected by the LRG targeting algorithm, DESI_TARGET = 2 ($\equiv 2^1$) would signify a source selected by the ELG targeting algorithm, and DESI_TARGET = 3 ($\equiv 2^1 + 2^0$) would signify a source selected by both of the LRG and ELG targeting algorithms. The desitarget pipeline uses 64-bit signed integers to represent bitmasks. But the desitarget bit-values that are used to distinguish target classes are never negative numbers, so each of the bitmasks can register up to 63 distinct selections (bit 0 to bit 62).⁴³

A more realistic example of a DESI bitmask would be 1152921504606849827, which equals $2^{60} + 2^{11} + 2^9 + 2^8 + 2^5 + 2^1 + 2^0$. For the DESI Main Survey (see Section 2.3), a bitmask of 1152921504606849827 would signify that a target was part of the DESI Bright Galaxy Survey (2^{60}) as well as being selected by the LRG (2^0) and ELG (2^1) targeting algorithms—specifically, by the ELG "low-priority" targeting algorithm (2^5). The remaining bits ($2^{11} + 2^9 + 2^8$) are informational, and indicate that the method used to target ELGs and LRGs corresponded to the northern target-selection algorithm (see Section 4.1.3).

2.3. Commissioning, Survey Validation, and the Main Survey

The DESI survey has progressed through a number of targeting phases:

- 1. First, DESI underwent commissioning (e.g., Besuner et al. 2020; Meisner et al. 2020; Perruchot et al. 2020; Poppett et al. 2020; Shourt et al. 2020). During this phase, a number of target classes were defined to calibrate the instrument, such as dither stars to test pointing models, standard stars to test throughput, and "first light" science targets. The commissioning phase of DESI is typically abbreviated as CMXby the desitargetpipeline. Although a general reader is unlikely to encounter commissioning targets, the following should be noted:
 - (a) CMX observations were highly preliminary and, for expediency, were not carefully tracked and documented by desitarget. From a targeting perspective, CMX observations therefore should *not* be used for general scientific analyses. Nevertheless, broadly, commissioning bitmasks can be handled by replacing occurrences of SV1 or sv1 (see below) with cmx or CMX and occurrences of desi_mask with cmx_ mask.⁴⁴

³⁹ https://healpix.sourceforge.net/

⁴⁰ https://github.com/desihub/desitarget

⁴¹ https://desidatamodel.readthedocs.io

⁴² https://github.com/desihub/desitarget/blob/2.5.0/doc/nb/target-

selection-bits-and-bitmasks.ipynb, in Jupyter notebook format.

The sign was reserved in case especial cases arose; see, e.g., Section 3.3.2.
 Note that there were no specific MWS, BGS, or secondary targets for CMX, so there are no bitmasks for these program types.

- (b) The CMX data a general reader is most likely to encounter is a tile of observations of M33 with TILEID = 80615.⁴⁵
- 2. Prior to commencing its main 5 yr mission, DESI underwent a period of survey validation (see DESI Collaboration, 2023, in preparation). The major objectives of this phase were to validate the end-to-end DESI systems, to provide observations to refine the purity and completeness of the targeting algorithms that would be used for the main DESI survey, and to stress test the procedures that would be needed for day-to-day DESI operations. Survey validation is typically abbreviated as SV by the desitarget pipeline, and SV itself was divided into three distinct stages:

(a) SV1: the first iteration of SV ran, mainly,⁴⁶ from the night of 20201214 to the night of 20210402⁴⁷ and encompassed DESI tiles with TILEID numbers in the interval 80605–80975.⁴⁸ A total of 189 tiles were targeted in the context of SV1. These tiles were spread widely across the sky and were observed over multiple months in order to sample a wide range of observational conditions. The main goals of SV1, from a data-pipeline perspective, were to collect sufficient spectra of different target classes to train target-selection algorithms and to refine the spectroscopic pipeline (see also Guy et al. 2022). The targeting algorithms deployed during SV1 were deliberately loosened to allow targeting to be refined.

- (b) SV2: the second phase of SV (also referred to as "the 0.1% Survey") ran from the night of 20210324 to the night of 20210511, covering DESI tiles in the interval 81000–81099. During SV2, 39 tiles were observed. The purpose of SV2 was to test critical end-to-end operational procedures, such as establishing the processing of "Merged Target Ledgers" (MTLs; see E. Schlafly et al. 2023, in preparation), in which spectroscopic classifications and redshifts are used to decide whether a target requires additional observations.
- (c) SV3: the final stage of SV (also known as "the One-Percent Survey," but typically referred to as SV3 in this paper) ran from the night of 20210405 to the night of 20210610, although it was mostly completed by the end of the night of 20210513. SV3 covered DESI tiles with IDs in the interval 1–596. During SV3, 488 tiles were observed, 239 in dark time, 214 in bright time, and 35 as part of a "backup" program (see Sections 2.5 and 4.1.2). Observations were conducted in 20 separate "rosette" patterns consisting of 11 overlapping pointings arranged in a circle comprising a little more than 7 deg² of total area. A few additional passes were made in some rosettes to further increase completeness. The main goals of SV3 were to conduct DESI observations in a mode that mimicked Main

Myers et al.

Survey observations, and to sample Main Survey target classes at high completeness.

The DESI Collaboration typically refers to SV1 as "Survey Validation" and SV3 as "the One-Percent Survey" (again, see DESI Collaboration et al. 2023, in preparation). But, this paper will refer to the entire period encompassing SV1, SV2, and SV3 as "Survey Validation," as such terminology better matches the nomenclature embedded in the desitarget data model.

3. Finally, DESI embarked upon its 5 yr Main Survey, which is typically denoted main by the desitarget pipeline. First observations for the Main Survey began on 20210514. The target classes that had been refined using spectra from SV1 and observed at high completeness during SV3 then began to be observed in earnest. Observations were conducted using procedures that were introduced during SV2 and finalized near the end of SV3.

2.4. DESI Bitmask Denominations, Values, and Access

The DESI bit-names and bit-values are available on GitHub for SV1,⁴⁹ SV2,⁵⁰ SV3,⁵¹ and the Main Survey.⁵² Note that, although the vast majority of targets for SV1 used version 0.51.0 of the desitarget code, some SV1 secondary programs were updated to use version 0.52.0 of the code (see the Appendix of DESI Collaboration et al. 2023, in preparation, for more details). Further, for secondary targets from SV1, it is important to use the appropriate version of the desitarget code for the corresponding survey files (e.g., use version 0.51.0 of desitarget when working with bitmasks from files in the 0.51.0/targets/sv1 directory), as some SV1 secondary targeting bits were deprecated. For all primary targets (and throughout SV2, SV3, and the Main Survey), it is acceptable to use the corresponding version of the code or a later version as bitmasks were only added during software development. The bitmasks can be accessed directly in Python using the targetmask module in the desitarget code. For example, if a user is interested in the primary dark-time target classes (ELGs, LRGs, QSOs, etc.) the relevant bitmask can be accessed via the command from desitarget. targetmask import desi_mask. The specific commands needed to retrieve each bitmask for different combinations of survey phase and target type are listed in Table 1. The bits that correspond to the different bitmasks are stored in different columns depending on the type of target. For instance, the primary dark-time classes are stored in DESI data files in a column called DESI_TARGET for the Main Survey and SVX_DESI_TARGET for SV for $X \in 1, 2, 3$. The specific column names used to store each target type for each survey phase are listed in Table 2.

It may be helpful to illustrate the use of DESI targeting bitmasks with a short example. Say we wish to know if a target has been selected by (at least) the quasar targeting algorithm in the Main Survey. Further, let us assume that the appropriate

⁴⁵ See, e.g., E. Schlafly et al. 2023 (in preparation) for a description of DESI TILEIDS.

⁴⁶ A few SV1 observations were conducted alongside SV2, and occasional SV1 tiles were observed even later.

⁴⁷ The night of a DESI observation is recorded in the form YYYYMMDD where Y = year, M = month, and D = day.

 $^{^{48}}$ In 2022 January, an additional SV1 tile was observed with TILEID=82633.

⁴⁹ https://github.com/desihub/desitarget/blob/0.51.0/py/desitarget/sv1/ data/sv1_targetmask.yaml

⁵⁰ https://github.com/desihub/desitarget/blob/0.53.0/py/desitarget/sv2/ data/sv2_targetmask.yaml

⁵¹ https://github.com/desihub/desitarget/blob/0.57.0/py/desitarget/sv3/ data/sv3_targetmask.yaml

⁵² https://github.com/desihub/desitarget/blob/1.1.1/py/desitarget/data/ targetmask.yaml

 Table 1

 Commands for Importing DESI Targeting Bitmasks

Phase	Import Command
SV1	from desitarget.sv1.sv1_targetmask import desi_mask, bgs_mask, mws_mask, scnd_mask ^a
SV2	<pre>from desitarget.sv2.sv2_targetmask import desi_mask, bgs_mask, mws_mask, scnd_mask</pre>
SV3	<pre>from desitarget.sv3.sv3_targetmask import desi_mask, bgs_mask, mws_mask, scnd_mask</pre>
main	from desitarget.targetmask import desi_mask, bgs_mask, mws_mask, scnd_mask

^a desi_mask, bgs_mask, mws_mask, and scnd_mask here correspond to the different target type columns listed in Table 2. Any of the individual masks can be omitted from the command.

data file has been read into a structure called targs, then the necessary commands to determine which indexes in targs correspond to quasar targets would be as follows:

from desitarget.targetmask import desi_
mask as Mx
is_QSO = (targs["DESI_TARGET"] & Mx["QSO"])
! = 0.

Much additional functionality exists for working with DESI targeting bitmasks. Two utility functions, in particular, are worth mentioning. First, any integer (i) can be converted to the corresponding bit-names using the mask.names(i) function. For example, desi_mask.names(7) will return the list [``LRG,'' `ELG,'' `QSO'']. Second, the function main_cmx_or_sv can be used to directly extract the relevant bitmasks, column names, and survey phase (detailed in Tables 1 and 2) from a DESI data file. For example:

from desitarget.targets import main_cmx_or_sv
column_names, masks, survey = main_cmx_
or_sv(targs).

2.5. Observing Conditions and Programs

Depending on the quality of the observing conditions at Kitt Peak, DESI pursues one of three different observing programs, corresponding to three different sets of targets (see, e.g., E. Schlafly et al. 2023, in preparation, for more details). During DARK time, DESI observes ELG, LRG, and QSO targets; the bits for such targets are stored in columns with names that resemble DESI_TARGET (see Table 2) and are described in the desi_mask mask (see Table 1). During BRIGHT time, DESI observes BGS targets, stored in columns like BGS_TARGET and recorded in the bgs_mask mask, and MWS targets, stored in columns like MWS_TARGET and recorded in the mws_mask mask. Finally, during more marginal observing conditions, DESI pursues a BACKUP program (see Section 4.1.2); the bits that indicate targets that are part of the backup program are also stored in columns like MWS_TARGET and are recorded in the mws mask mask.

Because the same target can be selected to be observed in both, for example, bright and dark time, it is useful to record the specific programs associated with each target. In DESI targeting files, such information is stored in a column called OBSCONDITIONS and the corresponding bit-names and bitvalues are accessible in the obsconditions mask,⁵³ which can be accessed using a similar command to those listed in Table 1: from desitarget.targetmask import
obsconditions.

Note that there are some placeholder bits included in the obsconditions mask (e.g., GRAY) that were deprecated for the DESI Main Survey or that were never adopted for DESI operations.

3. TARGETID: A Unique Identifier for DESI Targets

DESI utilizes a unique ID, similar to the objID used by the SDSS (e.g., Stoughton et al. 2002), to track targeted objects through the end-to-end pipeline. In DESI, this unique targeting ID is denoted TARGETID. It is critical for DESI operations that each TARGETID is associated with only one object. However, it is less crucial that each object is associated with only one TARGETID, particularly for targets that will not be used in the key DESI large-scale structure analyses. So, it is important to note that (for most cases) TARGETID does not have coordinate-based provenance—and targets at the same coordinates can therefore have a different TARGETID. Instead, TARGETID is constructed using information from the imaging survey used to select a target.

An illustrative example in DESI might be a target selected from the Legacy Imaging Surveys versus a target selected using only information from Gaia (e.g., Gaia Collaboration et al. 2018a). Although an object might share a location on the sky in Gaia and the Legacy Surveys, desitarget does not perform an exhaustive coordinate match across all possible input surveys. Therefore, such a target can have a TARGETID derived from Legacy Surveys information (see Section 3.1) and a TARGETID derived from Gaia information (see Section 3.1.5).

An exception to the rule that the DESI TARGETID is not coordinate based is negative values of TARGETID. Negative TARGETIDs, which are described in Section 3.3.2, are used to distinguish sky locations that need to be rapidly assigned during survey operations.

3.1. Primary Targets

Figure 1 depicts the bitmask that defines the DESI TARGETID.⁵⁴ In the rest of this subsection, we will detail the bits that comprise a "primary" TARGETID, by which we mean the unique identifier for a source that is part of one of the primary DESI subprograms, i.e., LRGs, ELGs, QSOs, the BGS or the MWS, rather than being part of a secondary program (see, e.g., Table 2).

⁵³ https://github.com/desihub/desitarget/blob/1.1.1/py/desitarget/data/ targetmask.yaml#L184-L187

⁵⁴ See also https://github.com/desihub/desitarget/blob/1.1.1/py/desitarget/ data/targetmask.yaml#L214-L221.

 Table 2

 Relevant Column Names for DESI Targeting Bitmasks

		Primary Targets		Secondary Targets
Survey Phase	Dark time ^a	BGS	MWS	
SV1	SV1_DESI_TARGET	SV1_BGS_TARGET	SV1_MWS_TARGET	SV1_SCND_TARGET
SV2	SV2_DESI_TARGET	SV2_BGS_TARGET	SV2_MWS_TARGET	SV2_SCND_TARGET
SV3	SV3_DESI_TARGET	SV3_BGS_TARGET	SV3_MWS_TARGET	SV3_SCND_TARGET
main	DESI_TARGET	BGS_TARGET	MWS_TARGET	SCND_TARGET

^a Dark-time targets include LRGs, ELGs, and QSOs.

3.1.1. The OBJID and BRICKID Bits

Imaging from the Legacy Surveys (Dey et al. 2019; D. Schlegel et al. 2023, in preparation), on which DESI targeting draws, is processed through the legacypipe pipeline⁵⁵ to produce catalogs of sources. The legacypipe code extracts these sources in individual regions of the sky denoted as bricks. The bricks cover an area of roughly 0.25×0.25 on the sky, and each brick is assigned a unique integer from 1 to 662,174. Each extracted source in each brick is assigned an integer from 0 to N - 1, where N is the total number of sources extracted by legacypipe in the brick. The unique brick number is called a BRICKID and the unique source number is called an OBJID.⁵⁶

Together, OBJID and BRICKID encode a unique legacypipe-extracted source. TARGETID inherits these two numbers as a method to track unique sources, and they occupy the two least significant sets of bits in TARGETID (see Figure 1).

3.1.2. The RELEASE Bits

OBJID and BRICKID encode a unique source but not the imaging reductions from which that source was extracted. To track this, starting with Data Release 4 (DR4), the Legacy Surveys included a column RELEASE that records image-processing information.⁵⁷ RELEASE is an integer in the thousands, with the first digit denoting the Data Release and subsequent digits representing the photometric system.

The Legacy Imaging Surveys commenced with Data Release 1, which means RELEASE numbers <1000 have no predefined meaning. As we will discuss more in subsequent sections, we take advantage of this extra bit-space by using integers with RELEASE <1000 to indicate a target that (potentially) was selected using information from beyond the Legacy Surveys.

3.1.3. The MOCK Bit and TARGETIDs for Random Catalogs

In preparation for going on-sky, the DESI collaboration produced multiple catalogs of ersatz targets to stress test the end-to-end software pipelines and to model survey outcomes. The MOCK bit is a single bit (again, see Figure 1) that encodes whether or not a source is derived from such "mock" data. In addition, the desitarget pipeline includes a randoms module that can be used to produce catalogs of points that mimic the selection function of sources in the Legacy Imaging Surveys (see Section 4.5.1). A value of "1" for the MOCK bit indicates that a target was generated as part of either a "mock" data run or a random catalog.

Information at locations created in desitarget random catalogs is derived at the pixel level using images from the Legacy Surveys (again, see Section 4.5.1). As such, RELEASE and BRICKID are inherited directly from the Legacy Surveys for points in random catalogs. Within a brick, desitarget assigns an OBJID integer for random points sequentially in order of R.A. (henceforth RA), such that the object with the smallest RA is assigned OBJID = 0. Ordering by RA in this manner guarantees that random points in northern and southern portions of the Legacy Surveys (see Section 4.1.3) have the same OBJID. Note that even though [RELEASE, BRICKID, OBJID] for random points can resemble similar combinations for real sources from the Legacy Surveys, random points will always have a distinct TARGETID because the MOCK bit will be set.

3.1.4. The SKY Bit and TARGETIDs for Blank Skies

The desitarget pipeline also generates blank-sky positions to facilitate spectral calibration (see Section 4.4.1). The SKY bit is a single bit (again, see Figure 1), which, if set, indicates that a target corresponds to a sky position.

The majority of blank-sky locations in the DESI footprint are assigned at the pixel level using blob maps from the Legacy Surveys imaging (again, see Section 4.4.1). For these sky locations, RELEASE and BRICKID are therefore inherited directly from the Legacy Surveys. Within each brick, the desitarget code assigns sequential integers for OBJID. This schema is guaranteed to produce unique values of TARGETID as setting the SKY bit distinguishes combinations of [RELEASE, BRICKID, OBJID] from any target derived from the Legacy Surveys source catalogs.

DESI also creates files of "supplemental" sky targets to cover areas outside of the Legacy Imaging Surveys (see Section 4.4.2), which are assigned by using a catalog-level match to avoid bright sources in Gaia. If these supplemental skies are near a bright object in Gaia, they are retained but are designated to be BAD_SKY (see Section 4.4.2 for more details). For BAD_SKY supplemental sky targets, RELEASE is set to 0 and the GAIADR indicative bits (see Section 3.1.5) are set to the Gaia Data Release integer used to assign supplemental skies (typically "2" to indicate Gaia DR2). For other supplemental skies, both the RELEASE and GAIADR bits are populated with the Gaia DR integer. BRICKID is set to the nside = 256 HEALPixel integer of each supplemental sky location, and OBJID is a sequential integer generated within each HEAL-Pixel. The SKY bit, of course, is also always set. Note that setting RELEASE <1000 guarantees that the TARGETID for

⁵⁵ https://github.com/legacysurvey/legacypipe

⁵⁶ See, e.g., https://www.legacysurvey.org/dr9/catalogs/.

⁵⁷ See, e.g., https://www.legacysurvey.org/release/.



Figure 1. The bits that comprise a positive DESI TARGETID. The least significant bit (i.e., 2^0) is depicted to the right of this diagram. The 58 least significant bits ($2^{2^{57}}$) record information from the Legacy Imaging Surveys. The next two most significant bits are Boolean flags that indicate whether a target is drawn from a mock or random catalog (2^{58}) and whether a target is a blank-sky location (2^{59}). The remaining two populated bits (GAIADR), which are used to indicate targets that are selected solely using Gaia information, store the Gaia Data Release number ($2^{60}-2^{61}$). Bit 62 is not used. Note that DESI also uses a differently structured *negative* TARGETID to distinguish some sky locations (see Section 3.3.2).

supplemental skies will always differ from the TARGETID for pixel-level skies, as there were no Legacy Surveys Data Releases prior to DR1.

TARGETID that is guaranteed to always be unique as well as distinct from the TARGETID of any primary target.

3.2.1. The General TARGETID for Secondary Targets

3.1.5. The GAIADR bits and Gaia-based TARGETIDS

A substantial set of DESI targets are derived solely using information from Gaia. These targets generally lie outside of the Legacy Imaging Surveys footprint and are intended for BACKUP observations (see Section 2.5). Objects that are targeted purely using Gaia have the GAIADR bits (depicted in Figure 1) set to the Gaia Data Release used to select the target (i.e "2" for Gaia DR2).⁵⁸ For targets that are selected solely using Gaia, BRICKID is populated with the integer of the nside = 32 HEALPixel that contains the target, and OBJID is set to a sequential integer generated within each HEALPixel. Note that this schema differs slightly from the one described for supplemental skies in Section 3.1.4 (for which nside = 256 is adopted when populating BRICKID).

3.2. Secondary Targets

In addition to "primary" target classes, DESI observes a number of secondary programs (see also Section 4.6) proposed by the wider DESI collaboration to address specific science goals. In many cases, the proposers specified that targets should not share the observing characteristics of a DESI primary target. For example, a supernova in a galaxy that is part of the BGS sample should not be considered to be the same target as the galaxy itself. Such a supernova might not be at exactly the same location as the galactic center, and may, for instance, need to be scheduled for dark-time observations rather than during bright time. The end-to-end DESI pipeline uses TARGETID to ensure correspondence between a target, an allocated fiber, and the resulting spectrum. Therefore, secondary targets that cannot be merged with a corresponding primary target (see Section 4.6) must, by definition, have a different TARGETID. In some instances, there is an unused primary TARGETID, because a secondary target that is also a source from the Legacy Surveys is not part of the primary DESI campaign. In those instances, a secondary target is simply assigned the primary TARGETID detailed in Section 3.1. Still, there are cases where a secondary target does not have a corresponding source in the Legacy Surveys or a secondary target is prevented from being merged with a primary target that is already using the primary TARGETID. In those cases, the desitarget pipeline constructs a "secondary"

The secondary TARGETID is structured similarly to the schema displayed in Figure 1. BRICKID has the same definition for secondary and primary targets—it is the number of the brick from the Legacy Surveys on which the secondary target lies—but OBJID and RELEASE have different meanings. OBJID, for secondary targets, records the row that a target was listed in a file of secondary targets prepared by a proposer to pass to the desitarget pipeline, after first partitioning by BRICKID. RELEASE, for secondary targets, corresponds to

$$100[X - 1] + \log_2(S), \qquad (1)$$

where X is the iteration of SVX during which the target was observed (see, e.g., Table 2), with [X - 1] = 5 used to denote the DESI Main Survey.⁵⁹ Here, S is the bit-value from the scnd_mask (see, e.g., Table 1). For example, pseudo-RELEASE = 245 would denote secondary targeting bit 45 from the SV3 campaign and pseudo-RELEASE = 534 would denote secondary targeting bit 34 from the main observing campaign. Note that secondary targets that have not adopted the primary TARGETID can always be identified by virtue of having RELEASE <1000. TARGETID bits other than OBJID, BRICKID, and RELEASE are set to zero for secondary targets.

3.2.2. The TARGETID for Targets of Opportunity

When merited, the DESI survey can occasionally be suspended to pursue rapidly designed, dedicated tiles that place fibers on targets of opportunity (ToOs). An example science case might be to follow up potential host galaxies of gravitational waves or neutrino events (e.g., Palmese et al. 2021). DESI targeting also facilitates an observing mode where some ToOs are added to the existing pool of science targets. As with any DESI target class, ToOs must have a unique TARGETID so that they can be tracked by the end-to-end pipeline. ToOs have a TARGETID that follows the model of Figure 1, and BRICKID has the same meaning as for general primary and secondary targets. The list of all DESI ToOs is maintained in a single ledger, to which new targets can be added but not removed,⁶⁰ and the OBJID for ToOs is simply

⁵⁸ Gaia DR1 was never used for DESI targeting. GAIADR==1 instead denotes a "first light" target observed during commissioning.

 $[\]frac{59}{\text{New or updated sets}}$ of secondary targets assigned during the Main Survey will progressively use higher (integer) values of *X*.

 $^{^{60}}$ See E. Schlafly et al. 2023 (in preparation) for a full description of the ToO ledger.

the row number of the corresponding target in that ledger. For ToOs, RELEASE is always set to 9999, which is distinct from any other RELEASE value used to encode a TARGETID. The other possible TARGETID bits are set to zero for ToOs.

3.2.3. Tertiary Targets

Occasionally, special tiles need to be rapidly scheduled for a scientific purpose that does not comport with the desitarget scheme for ToOs. To facilitate prompt observation of such tiles, desitarget reserves RELEASE bits equivalent to 8888 to denote targets that were "directly" assigned fibers. Targets that have a TARGETID with a RELEASE of 8888 were not processed by the desitarget pipeline and are instead maintained and handled by the DESI fiberassign code (A. Raichoor et al. 2023, in preparation). Such programs are sometimes referred to as "tertiary" programs.

3.3. Encoding and Decoding TARGETID

The desitarget code base includes utility functions for extracting the bits depicted in Figure 1 from a TARGETID and for constructing a TARGETID from its constituent bits. These functions will work for all positively valued TARGETID integers.

3.3.1. desitarget Functionality

The specific functions in desitarget for encoding and decoding TARGETID can be called as follows:

from desitarget.targets import decode_targetid, encode_targetid objid, brickid, release, mock, sky, gaiadr = decode_targetid(targetid) targetid = encode_targetid(objid, brickid, release, mock, sky, gaiadr).

Both of the encode_targetid and decode_targetid functions are written with a high degree of flexibility. For example, targetid can be passed to decode_targetid as a single integer or as an array. Similarly, encode_targetid can be passed a mix of arrays and integers. Additionally, none of the inputs must be passed. This makes it straightforward to encode a TARGETID that is composed of, say, arrays of BRICKIDs and OBJIDs but only a single RELEASE number.

3.3.2. Negative TARGETIDS

Some TARGETID integers in DESI files have negative values. Typically, these correspond to stuck fibers that could not be used by DESI to observe a science target so were assigned to sky locations on-the-fly (see, e.g., E. Schlafly et al. 2023, in preparation). Assigning stuck fibers to sky locations frees up fibers to be assigned to science targets elsewhere in the focal plane. The least significant 29 bits encoded in a negative TARGETID correspond to decl., partitioning in equal bins in a northerly direction with the decl. = -90° bin corresponding to bit 0. The next most significant 30 bits correspond to RA, partitioned in equal bins with the R.A. = 0° bin corresponding to the least significant bit. The remaining four bits allow flexibility to record up to 16 different groups of distinct negative TARGETIDs.

The resulting schema allows unique TARGETIDs to be assigned by location at a resolution of ~ 1.2 milliarcseconds. Negative TARGETIDs can be created and decoded using desitarget routines as follows:

```
from desitarget.targets import decode_ne-
gative_targetid, encode_negative_targetid
ra, dec, group = decode_negative_targetid
(targetid)
targetid = encode_negative_targetid(ra,
dec, group = group).
```

Here, decode_negative_targetid(targetid) will return the lowest-valued bin edge in RA and decl. (i.e., the "left-hand" and "bottom" edges of the bin). A full description of how negative TARGETIDs are handled by the DESI fiberassignment code is provided in A. Raichoor et al. (2023, in preparation).

4. DESI Target Classes

In this section, we detail each of the main DESI target classes and the location and structure of the files in which they are stored. We will frequently refer to columns in these files being populated by certain "target selection bits," which are detailed in Section 2.

4.1. Primary Science Targets

The selection and optimization of DESI science targets for SV3 and the Main Survey are described in a series of papers, listed in Table 3.

4.1.1. Bright-time and Dark-time Targets

Bright-time and dark-time science targets are publicly available online for SV1,⁶¹ SV2,⁶² SV3,⁶³ and the Main Survey.⁶⁴ Each url links to bright and dark subdirectories which contain bright-time (MWS, BGS) and dark-time (ELG, LRG, QSO) targets, respectively. Note that operational procedures were in flux during SV1, meaning that a (very small) subset of SV1 targets can only be found in earlier versions of the target files than 0.51.0. Similarly, a few hundred targets from the DESI Main Survey are only available in the 1.0.0 version of the target files (see Section 5.1).

The bright/dark subdirectories host files containing targets that are only selected by bright-time/dark-time target-selection algorithms. However, every target in each subdirectory has all of the appropriate targeting bits populated. Consider, for instance, a QSO target that is *also* targeted by the MWS survey. Such a target will appear in the appropriate file in *both* of the bright/dark subdirectories with the same bits set in the DESI_TARGET and MWS_TARGET columns. However, a QSO target that is *not* also targeted by a bright-time survey will *only* appear in the appropriate file in the dark subdirectory. Similarly, a bright-time target that is not also

⁶¹ https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/0.51.0/targets/

sv1/resolve/

⁶² https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/0.53.0/targets/ sv2/resolve/

⁶³ https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/0.57.0/targets/ sv3/resolve/

⁶⁴ https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/1.1.1/targets/ main/resolve/

Target-selection Class	Main Survey Bit-names ^a	Brief Description	Selection Detailed In
Dark-time targets			
Luminous red galaxies	LRG	LRG target	Zhou et al. (2022)
Emission line galaxies	ELG	ELG target	Raichoor et al. (2022)
	ELG_LOP	ELG at standard priority	
	ELG_HIP	ELG observed at the	
		priority of an LRG ^b	
	ELG_VLO	Very-low-priority "filler" ELG	
Quasars	QSO	Quasar target	Chaussidon et al. (2022a)
Bright-time targets			
Bright Galaxy Survey	BGS_ANY	Any BGS bit is set	Hahn et al. (2022)
	BGS_FAINT ^C	Faint BGS target	
	BGS_BRIGHT	Bright BGS target	
	BGS_WISE	AGN-like BGS target	
	BGS_FAINT_HIP	Faint BGS target prioritized	
		like a bright BGS target ^b	
Milky Way Survey	MWS_ANY	Any MWS bit is set	Cooper et al. (2022)
	MWS_BROAD ^d	Magnitude-limited bulk sample	
	MWS_WD	White dwarf	
	MWS_NEARBY	Volume-limited ~100 pc sample	
	MWS_BHB	Blue horizontal branch target	
	MWS_MAIN_BLUE	Magnitude-limited blue sample	
	MWS_MAIN_RED	Magnitude-limited red sample	
Backup targets			
Part of the	BACKUP_GIANT_LOP ^e	Candidate giant star	MWS et al. 2023 (in preparation)
Milky Way Survey		observed at lower priority	
	BACKUP_GIANT	Candidate giant star	
	BACKUP_BRIGHT	Brighter backup target	
	BACKUP_FAINT	Fainter backup target	
	BACKUP_VERY_FAINT	Even fainter backup target	

 Table 3

 Publications That Describe Specific DESI Science Target Classes

^a Stored as bit-values in the DESI_TARGET, BGS_TARGET, and MWS_TARGET columns. Bit-names can be converted to bit-values using the desi_mask, bqs_mask, and mws_mask bitmasks (see Section 2.4).

^b Some targets with low observational priority are observed at higher priority to help characterize the survey selection function.

 $^{\rm c}$ BGS bits other than <code>BGS_ANY</code> are stored in the <code>BGS_TARGET</code> column and <code>bgs_mask</code> bitmask.

 d MWS bits other than MWS_ANY are stored in the MWS_TARGET column and <code>mws_mask</code> bitmask.

^e BACKUP bits are stored in the MWS_TARGET column and mws_mask bitmask.

targeted as part of the dark-time survey will only appear in the appropriate file in the bright subdirectory.

Files in the bright/dark subdirectories have names of the form AAAtargets-OBSCON-hp-HPX.fits. Here, AAA is sv1, sv2, or sv3 for SV files and is omitted for Main Survey files, OBSCON is either dark or bright, and HPX refers to the nside = 8 integer of the HEALPixel that contains the targets.⁶⁵

The contents of each file described in this section are similar. Information is stored in FITS format (Wells et al. 1981) and data are only included in extension 1 of the FITS file. Most of the data columns are directly derived from the DR9 "sweep" and "light curve sweep" catalogs⁶⁶ from which DESI Main Survey targets were selected. Additional columns that are derived by desitarget itself are detailed in Table 4.

In Tables 5 and 6 we show the densities of the principal DESI Main Survey bright-time and dark-time target classes. We also show how the densities of these target classes overlap (i.e., whether a target is included in both of two target classes). For example, Table 6 demonstrates that $\sim 0.4\%$ (7.6/1910.5) of

MWS targets are also quasar targets. This overlap is driven by the MWS_MAIN_BLUE sample described in Section 4.1 of Cooper et al. (2022), which contains blue point sources that resemble quasars in imaging. Densities in these tables are derived from the target files discussed in this section, with areal weights derived using the desitarget pixelized weight maps detailed in Section 4.5.2. The resulting matrices of densities can also be seen on the front page of the quality assurance (QA) web pages described in Section 4.5.3.

4.1.2. Backup Targets

In addition to other campaigns, DESI pursues a program of "backup" targets when observing conditions are of insufficient quality to collect signal on bright-time or dark-time science targets (see also Section 2.5). The principal backup targets are listed in the final part of Table 3 and will be detailed in a forthcoming paper (see MWS et al. 2023, in preparation). Backup targets differ from other science targets in that they are selected solely using Gaia information rather than imaging from the Legacy Surveys. As such, the directory structure for the files that contain backup targets differs slightly from that of other science targets, having gaiadr2 in the directory

 $^{^{65}}$ This nside is stored in the file header as FILENSID.

⁶⁶ https://www.legacysurvey.org/dr9/files/#sweep-catalogs-region-sweep

 Table 4

 Columns Added to Target Files by the desitarget Pipeline

Column Name	Python Format	Description	
BRICK_OBJID	i4	OBJID from the Legacy Surveys sweep catalogs.	
MORPHTYPE	U4	TYPE from the Legacy Surveys sweep catalogs.	
PHOTSYS	U1	"N" for MzLS/BASS photometric system, "S" for DECaLS. ^a	
TARGETID	i8	See Section 3.	
DESI TARGET 18 Column name		Column name differs for SV; see Section 2.4.	
BGS_TARGET	ARGET i8 Column name differs for SV		
MWS_TARGET	i8	Column name differs for SV; see Section 2.4.	
SCND_TARGET	i8	Column name differs for SV; see Section 2.4.	
SUBPRIORITY	f8	Random number (0-1) used to break fiber-assignment ties.	
OBSCONDITIONS	i8	See Section 2.5.	
PRIORITY_INIT	i8	See E. Schlafly et al. 2023 (in preparation).	
NUMOBS_INIT	i8	See E. Schlafly et al. 2023 (in preparation).	
HPXPIXEL	i8	The nside = 64 HEALPixel of the target. ^b	

^a BASS, MzLS, and DECaLS are the individual surveys that comprise the DESI Legacy Imaging Surveys (e.g., Dey et al. 2019). The photometric system (bok for BASS, mosaic for MzLS, and decam for DECaLS) can be derived from the RELEASE number according to https://www.legacysurvey.org/release/.

construction instead of dr9.⁶⁷ For versions of desitarget starting with 0.52.0, backup targets are stored in a backup subdirectory, but prior to 0.52.0 this subdirectory was instead called supp (for "supplemental"). Backup targets used for the DESI Main Survey are derived from a slightly later release of the desitarget code than for other target classes.⁶⁸

4.1.3. Overlapping Imaging: The North/South "Resolve"

The majority of DESI targets are selected from the Legacy Surveys, which comprises three individual imaging campaigns called BASS, MzLS, and DECaLS. Because these surveys overlap (see, e.g., Figure 1 of Dey et al. 2019), it is crucial to decide which imaging to use in each area of the sky. To help decide this, the desitarget pipeline adds a PHOTSYS column to targeting files (see Table 4), which is set to "N" for targets derived using imaging from the "northern" Legacy Surveys (BASS and MzLS) and "S" for targets derived using DECaLS imaging. A function called targets.resolve in the desitarget pipeline then uses PHOTSYS to determine (or "resolve") which imaging to adopt in a given area of the DESI footprint.

The desitarget resolve function first defines the "northern" sky to comprise areas that are both north of the Galactic plane (in Galactic coordinates) and north of decl. $\ge 32^\circ.375$ (in equatorial coordinates). Any area that does not meet both of these criteria is considered to be "southern" area (again, Figure 1 of Dey et al. 2019 may be helpful to visualize these areas). The desitarget resolve function then ensures unique imaging by only retaining targets that are in the northern sky area and that have PHOTSYS== `N' set, or that are in the southern sky area and have PHOTSYS== `S' set. Note that no resolve is necessary for targets—such as backup targets—that are selected purely from Gaia data.

The urls from which to retrieve targets that are provided in Section 4.1.1 all terminate with a resolve directory, indicating that they have been resolved by desitarget into unique targets in areas where the individual Legacy Surveys overlap. It is possible to use the desitarget pipeline to generate all targets in overlapping imaging areas rather than resolving them, and some (very early) versions of DESI targeting files were processed with this option. Such targets end up in a directory structure that includes noresolve instead of resolve.

4.1.4. Gaia DR2 or Gaia EDR3?

Version 0.58.0 of the desitarget code introduced the option of partially using Gaia Early Data Release 3 (EDR3; Gaia Collaboration et al. 2021) for target selection, with a view to optimizing astrometry-based selection of MWS targets without changing Gaia imaging information for selections used by other target classes. Target files processed with this option have a GAIASUB value in the FITS header that is set to True. EDR3 values are obtained by coordinate matching the Legacy Surveys sweep catalogs used for target selection (see Section 4.1.1) and EDR3 at 0.2, after first accounting for EDR3 proper motions. The EDR3 columns that are substituted for Gaia in such target files are listed in Table 7.

So, broadly, SV targets were selected using Gaia DR2 and Main Survey targets were selected using imaging quantities from Gaia DR2 and astrometric quantities from Gaia EDR3. An important caveat to this statement is that, as of the time of writing—corresponding to version 2.5.0 of desitarget backup targets (Section 4.1.2) still always use Gaia DR2 for all quantities. This is due to a bug where the Gaia EDR3 quantities are not substituted for backup targets even though the GAIASUB keyword is set to True.

4.2. Standard Stars

The DESI survey requires flux standards for spectrophotometric calibration. To achieve this goal, desitarget defines three principal standard star classes to be used inside of the footprint of the Legacy Surveys imaging, with bit-names

⁶⁷ For instance, SV3 backup targets are at https://data.desi.lbl.gov/public/ ets/target/catalogs/gaiadr2/0.57.0/targets/sv3/resolve/.

⁶⁸ https://data.desi.lbl.gov/public/ets/target/catalogs/gaiadr2/2.2.0/ targets/main/resolve/

 Table 5

 Densities (deg⁻²) of Principal Main Survey Target Classes in Dark Time

Bit-name	BGS_ANY	ELG	LRG	MWS_ANY	QSO	STD_BRIGHT ^a	STD_FAINT ^a
BGS_ANY	79.8 ^b	1.0 ^c	78.4	0.5	0.9	0.0	0.0
ELG		2394.3	2.3	0.0	108.5	0.0	0.0
LRG			625.7	0.4	4.0	0.0	0.0
MWS_ANY				119.0	7.6	51.9	104.2
QSO					319.8	0.0	0.0
STD_BRIGHT						54.9	54.9
STD_FAINT							107.2

^a STD_BRIGHT and STD_FAINT are the main standard star classes intended for use in bright time and dark time, respectively (see Section 4.2).

^b On-diagonal elements represent the density for a target class, with areas derived using the desitarget weight maps (see Section 4.5.2).

^c Off-diagonal elements represent the overlap in density between two target classes.

STD_BRIGHT, STD_FAINT (see also Table 5 and Table 6), and STD_WD,⁶⁹ and three classes for use outside of the Legacy Surveys imaging footprint, GAIA_STD_BRIGHT, GAIA_STD_FAINT, and GAIA_STD_WD. The Gaia standard star selections are designed to approximate the principal standard star classes but using imaging from Gaia in place of Legacy Surveys quantities. The STD_WD and GAIA_STD_WD classes are identical to each other—as well as to the MWS_WD class described in Cooper et al. (2022)—so will not be further detailed here.

The STD_BRIGHT and STD_FAINT target classes are designed to select main-sequence F-type stars in a similar fashion to the BOSS selection of spectrophotometric standards.⁷⁰ The color constraints differ somewhat from those applied in BOSS, though, both because DR9 of the Legacy Surveys does not include deep *u*- and *i*-band imaging (unlike the SDSS) and to ensure sufficient flux standards across the DESI focal plane. The overall approach introduces a small fraction of A-type stars, while retaining sufficiently precise flux calibration to measure signal in the Ly α forest. Broadly, the STD_BRIGHT target class is intended to be used for calibration during bright-time programs and the STD_FAINT target class is intended for use during dark time.

The STD_BRIGHT flux standards are selected as detailed in Table 8. The only difference between STD_BRIGHT and STD_FAINT targets is that STD_FAINT targets extend over the magnitude range $16 \le G < 19$ instead of $16 \le G < 18$ (in the Gaia *G* band).⁷¹ The selection of the GAIA_STD_BRIGHT and GAIA_STD_FAINT flux standards is detailed in Table 9. Again, GAIA_STD_BRIGHT and GAIA_STD_FAINT targets only differ because GAIA_STD_FAINT standards extend over the magnitude range $16 \le G < 19$. The proper-motion cuts listed in Tables 8 and 9 are included to help avoid stars in the thick disk of the Milky Way in favor of a more metal-poor halo turn-off population. Thick disk stars are more numerous at distances of a few kiloparsecs from the Sun, though the halo is hotter. Halo stars therefore typically have larger proper motions than thick disk stars at intermediate distances.

The algorithms used to select standard stars remained the same throughout DESI SV and the Main Survey. Gaia-only

standards (GAIA_STD_BRIGHT, etc.) were not introduced until version 0.48.0 of desitarget—and were not actually finalized until version 0.50.0—meaning that these standards do not appear in target files that predate SV1. As with other target classes (see Section 4.1.4) the option to substitute Gaia EDR3 astrometric data became available in version 0.58.0 of the desitarget code, and this option was used to produce targeting files for the DESI Main Survey.

Because the Gaia-only standards were designed for observations outside of the Legacy Surveys imaging footprint, the data model for these targets is slightly different to that of the principal flux standards. First, the Gaia-only bits (GAIA_STD_BRIGHT, etc.) can be accessed via the mws_mask bitmask and are stored in the MWS_TARGET column (see Tables 1 and 2), whereas the principal standards (STD_BRIGHT, etc.) can be accessed via the desi_mask bitmask and are stored in the DESI_TARGET column. Second, the Gaia-only standards, as is the case for all Gaia-only targets, reside in the backup files outlined in Section 4.1.2 rather than the bright-time and dark-time targeting files outlined in Section 4.1.1.

4.3. Guiding, Focusing, and Alignment targets

The DESI instrument requires samples of stars to use for guiding, focusing, and alignment. Collectively, we will refer to these as guiding, focusing and alignment targets, or just "GFAs." The desitarget pipeline includes a module called gfa specifically for assembling these targets. GFAs are publicly available online in directories accompanying those described in Section 4.1.1, where the directory names terminate with gfas instead of targets.⁷² Within a gfas directory, files have names of the form gfas-hp-HPX.fits. Here, HPX refers to the nside = 8 integer of the HEALPixel that contains the GFAs.⁷³

The selection of GFAs did not significantly change between iterations of SV and the DESI Main Survey. GFA targets essentially comprise all sources in Gaia limited to G < 21 in the Gaia G band. We include extra imaging information for Gaia sources that are also in the Legacy Surveys sweep catalogs⁷⁴ (i.e., that have REF_ID > 0) and that are *not* part of the 2020 Siena Galaxy Atlas (J. Moustakas et al. 2023, in preparation).

⁶⁹ "WD" denotes a white dwarf.

⁷⁰ https://www.sdss.org/dr12/algorithms/boss_std_ts/

⁷¹ Standards appropriate for bright time were also intended to be used in dark time, but not vice-versa, meaning that the STD_FAINT class is a "fainter" extension of the STD_BRIGHT selection. Hence the name "STD_FAINT" rather than, say, "STD_DARK."

⁷² For example, for the Main Survey the appropriate url is https://data.desi. lbl.gov/public/ets/target/catalogs/dr9/1.1.1/gfas/.

 $^{^{73}}$ As for science targets, this nside is stored in the file header as FILENSID.

⁷⁴ https://www.legacysurvey.org/dr9/files/#sweep-catalogs-region-sweep

Table 6Densities (deg^{-2}) of Principal Main Survey Target Classes in Bright Time

Bit-name	BGS_ANY	ELG	LRG	MWS_ANY	QSO	STD_BRIGHT ^a	STD_FAINT ^a
BGS_ANY	1389.7 ^b	1.0 ^c	78.4	0.6	0.9	0.0	0.0
ELG		1.0	0.4	0.0	0.0	0.0	0.0
LRG			78.8	0.4	0.3	0.0	0.0
MWS_ANY				1910.5	7.6	51.9	104.2
QSO					8.0	0.0	0.0
STD_BRIGHT						54.9	54.9
STD_FAINT							107.2

^a STD_BRIGHT (STD_FAINT) is the principal standard star class (see Section 4.2) intended for use in bright time (dark time).

^b On-diagonal elements represent the density for a target class, with areas derived using the desitarget weight maps (see Section 4.5.2).

^c Off-diagonal elements represent the overlap in density between two target classes.

Table 7
Columns That Use Gaia EDR3 Instead of DR2 when GAIASUB is Set in a File
Header

Note. Columns are as for the Legacy Surveys sweep files.

This ensures that high-precision fluxes and morphologies are retained for GFAs in the primary DESI targeting footprint to make quality cuts downstream of desitarget, if needed.

Between the SV and Main Survey phases of DESI the desitarget.gfa code was updated to include the option of substituting Gaia-based values from Gaia EDR3 in place of values from Gaia DR2 (see also Section 4.1.4). GFA targets for the DESI Main Survey included this substitution. In general, files of GFAs that were run with Gaia EDR3 have a GAIADR value in the FITS header that is set to "edr3." For GFAs that have a match in the Legacy Surveys sweep catalogs, only the columns in Table 7 are ever updated from Gaia DR2 to Gaia EDR3. For GFAs that do not match a Legacy Surveys source (i.e., typically objects outside of the Legacy Surveys imaging footprint) *all* Gaia-derived quantities are updated to EDR3.

Some sources that appear in Gaia—particularly in DR2—do not have measured proper motions. As astrometric information is essential for DESI guiding, the GFAs incorporate proper motions from the First U.S. Naval Observatory Robotic Astrometric Telescope Catalog (URAT; Zacharias et al. 2015) where Gaia proper motions are missing. Only proper motions for URAT sources that match a Gaia source within 0."5 are substituted for missing Gaia values. The desitarget code has the option of turning off this URAT-substitution mechanism. GFA targets that did include substitution of URAT proper motions in place of missing Gaia values can be identified by having NOURAT set to False in the FITS header of the GFA file. All of the columns in the desitarget GFA files are derived from the Legacy Surveys sweep catalogs⁷⁵ or are provided in Table 4, with the exception of URAT_ID, URAT_SEP, and GAIA_PHOT_G_N_OBS. The URAT_ID column records the recommended identifier for URAT1 sources (see http://cdsarc.u-strasbg.fr/ftp/1/329/ReadMe), URAT_SEP is the separation (in arcseconds) between a URAT source and a GFA target, and GAIA_PHOT_G_N_OBS records the number of observations in the *G* band from Gaia (which is useful for characterizing variable sources downstream of desitarget).

In addition, the MORPHTYPE column described in Table 4 has a different provenance for GFA targets that do not have a match in the Legacy Surveys. For such Gaia-only GFAs, a source is recorded as point-like if

$$[G \leq 19. \text{ AND AEN} < 10^{A}] \text{ OR}$$

 $[G \geq 19. \text{ AND AEN} < 10^{A+0.2(G-19.)}],$ (2)

where *G* denotes Gaia *G*-band magnitude (without any correction for Galactic extinction; $GAIA_PHOT_G_MEAN_$ MAG), AEN denotes $GAIA_ASTROMETRIC_EXCESS_$ NOISE, and *A* is 0.5 for Gaia DR2 or 0.3 for EDR3. Sources which do (do not) meet the criterion in Equation (2) are designated GPSF (GGAL) in the MORPHTYPE column.

4.4. Blank-sky Locations

The DESI fiber allocation software (A. Raichoor et al. 2023, in preparation) assigns fibers to empty or "blank" locations that the DESI spectroscopic pipeline (Guy et al. 2022) uses to perform sky subtraction. To assign these fibers, DESI operations utilizes precalculated lists of sky locations that are expected to include minimal flux from astronomical sources when integrated over a DESI fiber. The desitarget pipeline⁷⁶ assembles such lists of blank-sky locations using two different techniques, a pixel-level approach and a Gaia-avoidance approach, which we describe in this section.

In addition, as of the Main Survey, DESI operations software determines, on-the-fly, whether stuck fibers in the DESI focal plane would collect significant flux during survey observations. Where possible, these stuck fibers then become part of the sky budget, which releases other fibers that can then be used to

 ⁷⁵ https://www.legacysurvey.org/dr9/files/#sweep-catalogs-region-sweep
 ⁷⁶ Using a module called skyfibers.

Reason for Selection	How Selection Is Applied
Not masked	MASKBITS ^{a,b} is not BRIGHT or GALAXY
Point-like	TYPE=="PSF"
Isolated	$FRACFLUX_X^c < 0.01$
Measured flux	$FLUX_IVAR_X > 0$
Observed	$NOBS_X > 0$
Low fraction of masked pixels	FRACMASKED_X < 0.6
Color of halo turn-off or bluer ^d	(r-z) < 0.2 AND $0.0 < (g-r) < 0.35$
Reasonably bright in Gaia ^e	$16 \leqslant \text{PHOT}_G_\text{MEAN}_\text{MAG} < 18$
Bp – Rp color is measured	GAIA_PHOT_BP_MEAN_MAG - GAIA_PHOT_RP_MEAN_MAG is not NaN
No astrometry issues	GAIA_ASTROMETRIC_EXCESS_NOISE < 1 AND
	GAIA_ASTROMETRIC_PARAMS_SOLVED==31
Proper motions are finite	Neither PMRA or PMDEC is NaN
Parallax smaller than 1 mas	PARALLAX < 1
Proper motion larger than 2 mas yr^{-1}	$\sqrt{ extsf{PMRA}^2 + extsf{PMDEC}^2} > 2$
Unique source in Gaia	GAIA_DUPLICATED_SOURCE==0

^a Parameters that appear in this table are derived from the Legacy Surveys sweep files at https://www.legacysurvey.org/dr9/files/#sweep-catalogs-region-sweep.

^b The Legacy Surveys MASKBITS mask is described at https://www.legacysurvey.org/dr9/bitmasks/.

^c In this table, X denotes that the criterion is applied for *all* bands from the Legacy Surveys (g, r, and z).

 d g, r, z denote FLUX_G, FLUX_R, FLUX_Z converted to magnitudes and corrected for Galactic extinction.

 e° The only difference between the STD_BRIGHT and STD_FAINT target classes is that STD_FAINT extends across $16 \leq \text{PHOT}_G_{MEAN}_{MAG} < 19$.

observe science targets (see Section 3.3.2). Reassigning sky locations on-the-fly is detailed in A. Raichoor et al. (2023, in preparation) and E. Schlafly et al. (2023, in preparation).

4.4.1. Pixel-based Sky Locations

The legacypipe software used by the Legacy Surveys to extract sources from astronomical images operates on areas, called "blobs," in which pixels have signal that exceeds a certain detection threshold. The legacypipe code builds maps where pixels that lie in a particular blob are set to a positive integer that encodes that blob and pixels that are not in any blob are set to -1. These "blob maps" are created for each brick of the Legacy Surveys and are available in the metrics subdirectory of the DR9 Legacy Surveys release (D. Schlegel et al. 2023, in preparation), with names resembling metrics/ 000/blobs-0003m015.fits.gz.⁷⁷ The desitarget code adapts these blob maps by setting all pixels that do (do not) have a value of -1 to True (False), essentially building a pixel map that stores values of True where no sources are detected. The resulting source-detection maps share the Legacy Surveys native resolution for a brick of 3600×3600 pixels, at $0^{\prime\prime}_{...262}$ pixel⁻¹. So, each map contains 12.96M pixels spread across the approximately 0.0623 deg^2 unique area of a brick.

The desitarget pipeline produces catalogs of pixel-level blank locations at a user-specified density of sky positions. The density used for the Main Survey, for instance, was 18,000 sky locations per square degree, which is easily sufficient to fill the DESI focal plane.⁷⁸ The desitarget code then performs a binary erosion on the source-detection maps (in each unique brick) to achieve the required density. For example, if 18,000 deg⁻² sky locations are desired in a 0.0623 deg² brick, then a total of about 1120 locations are needed—meaning a brick needs to be eroded until it has dimensions of approximately 34×34 . In each step of the erosion, in a given grid cell, desitarget calculates the distance (in pixels) of the pixel that is farthest from a detected source; we will refer to this quantity as BLOBDIST. At the conclusion of the erosion procedure, a blank-sky target is placed in each grid cell at the location with the maximum value of BLOBDIST. If a grid cell completely overlaps pixels that include source detections, then BLOBDIST will be equal to 0, meaning that there are no completely blank-sky locations in that grid.

The desitarget pipeline then uses the photutils (Bradley et al. 2019) routines CircularAperture and aperture_photometry to extract flux at each calculated sky location. Fluxes are measured in the per-filter image and invvar stacks provided by the Legacy Surveys.⁷⁹ The measured fluxes, and their associated inverse variances, are then stored as quantities called FIBERFLUX_X and FIBER-FLUX_IVAR_X, where X refers to each of the Legacy Surveys optical bands (g, r, and z). The measured flux quantities can have multiple values if apertures of different radii are specified when running desitarget,⁸⁰ though processing for the DESI SV and Main Survey files exclusively adopted a 0."75 aperture, corresponding to the radius of a DESI fiber.

Given the high density of sky locations required by DESI, it is inevitable that apertures placed at certain positions will not be empty. The desitarget pipeline therefore labels two distinct types of sky locations: particularly good positions and more dubious ones. Sky locations that have any of FIBERFLUX_G==0 or FIBERFLUX_IVAR_G==0 or FIBERFLUX_R==0 or FIBERFLUX_IVAR_R==0 or BLOBDIST==0 are referred to as "bad" skies, and are assigned the bit-name BAD_SKY.⁸¹ Sky locations that do not meet these criterion are referred to just as

⁷⁷ For additional information regarding the Legacy Surveys maps, see, e.g., https://www.legacysurvey.org/dr9/files/#image-stacks-region-coadd.

⁷⁸ This quantity is stored as NPERSDEG in the header of files produced by desitarget.

⁷⁹ https://www.legacysurvey.org/dr9/files/#image-stacks-region-coadd

⁸⁰ The applied apertures are stored in a series of values called APO, AP1, AP2, etc., in the headers of the files produced by desitarget.

⁸¹ All sky-related bits appear in the desi_mask bitmask and are stored in the DESI_TARGET column (see Section 2).

 Table 9

 Target Selection for the GAIA_STD_BRIGHT and GAIA_STD_FAINT Target Classes

Reason for Selection	How Selection Is Applied
Point-like	GAIA_ASTROMETRIC_EXCESS_NOISE < 10 ^{0.5}
Isolated	No other Gaia source within 3."5
Color of halo turn-off or bluer ^a	0.2 < (Bp - Rp) < 0.9 AND
	(G - Bp) > -0.5(Bp - Rp) AND
	(G - Bp) < 0.3 - 0.5(Bp - Rp)
Reasonably bright in Gaia ^b	$16 \leqslant ext{PHOT_G_MEAN_MAG} < 18$
Bp – Rp color is measured	GAIA_PHOT_BP_MEAN_MAG - GAIA_PHOT_RP_MEAN_MAG is not NaN
No astrometry issues	$GAIA_ASTROMETRIC_EXCESS_NOISE < 1 AND$
	GAIA_ASTROMETRIC_PARAMS_SOLVED==31
Proper motions are finite	Neither PMRA or PMDEC is NaN
Parallax smaller than 1 mas	PARALLAX < 1
Proper motion larger than 2 mas yr^{-1}	$\sqrt{ extsf{PMRA}^2 + extsf{PMDEC}^2} > 2$
Unique source in Gaia	GAIA_DUPLICATED_SOURCE==0

^a In this table, Bp, *G*, Rp denote versions of GAIA_PHOT_BP_MEAN_MAG, GAIA_PHOT_G_MEAN_MAG, GAIA_PHOT_RP_MEAN_MAG that have been corrected for Galactic extinction using the prescription in Table 1 of Gaia Collaboration et al. (2018b).

^b The only difference between the GAIA_STD_BRIGHT and GAIA_STD_FAINT target classes is that GAIA_STD_FAINT extends across $16 \leq \text{PHOT}_G\text{_MEAN}\text{_MAG} < 19$.

"skies" and are assigned the bit-name SKY. The logic behind the criteria that determine whether a sky location is bad is that locations with zero flux or infinite flux variance in the *g* or *r* bands are typically missing entirely from the Legacy Surveys rather than genuinely having flux of exactly zero—and that sky locations with BLOBDIST==0 certainly intersect a source, as described above.

Pixel-level sky locations are derived within Legacy Surveys bricks. It is important to note that this generally contrasts with other files produced by desitarget, which are stored within HEALPixels. Therefore, before using the desitarget pipeline to make supplemental sky locations by avoiding bright stars in Gaia—which is discussed in the next section of this manuscript—it is important to recast the derived sky locations from a brick-based scheme to a HEALPixel-based scheme. The desitarget function that can be used to achieve this is called skyfibers.repartition_skies, and how this function is applied is detailed in the tutorial linked to in Section 4.7.

4.4.2. Supplemental Sky Locations

To account for the fact that certain areas of the sky covered by DESI might be outside of—or missing in—the Legacy Surveys, the desitarget pipeline generates an additional catalog of sky positions using Gaia. These "supplemental" sky locations are produced at random positions across the sky within HEALPixels at a user-specified density.⁸² For SV and Main Survey files, this density was chosen to be 18,000 deg⁻² to match the density at which pixel-based sky locations were generated. Supplemental sky locations are then removed if they are within a radius of 2" of any source in Gaia DR2. Finally, to limit the total density of sky locations that could be used by the DESI fiber-assignment software, supplemental skies are removed if they share a HEALPixel with a pixel-based sky location at nside = 4096, which corresponds to a resolution of approximately 18,000 deg⁻². Quantities generated for pixel-based sky locations are also "mocked up" for supplemental sky locations. The BLOBDIST value is set to be the (2") radius for avoiding Gaia sources, divided by the 0."262 pixel scale of the Legacy Surveys imaging. All FIBERFLUX_X and FIBERFLUX_IVAR_X quantities (see 4.4.1) are set to -1. Finally, supplemental skies are assigned the bit-name SUPP_SKY; as with other sky locations, this bit is set in the DESI_TARGET column and can be extracted using the desi_mask bitmask (see Section 2).

4.4.3. The Data Model for Sky Locations

Files of sky locations derived by desitarget at the pixel level (see Section 4.4.1) are available online in the dr9 directory in a similar way to the target files detailed in Section 4.1.1, with directory names that terminate with skies rather than targets.⁸³ Files of supplemental sky locations derived by avoiding bright sources in Gaia (see Section 4.4.2) are available in gaiadr2 directories in a similar manner to what is outlined in Section 4.1.2. Directory names for supplemental skies, however, terminate in skies-supp instead of in skies or targets.⁸⁴

Within the skies (skies-supp) directory, files have names of the form skies-hp-HPX.fits (skies-supp-hp-HPX.fits). Here, HPX refers to the nside = 8 integer of the HEALPixel that contains the sky locations. As is the case for other desitarget products, the adopted nside is stored in the file header as FILENSID. The skies directory also contains an unpartitioned subdirectory. This stores the original sky locations generated by desitarget before applying the repartitioning code discussed in Section 4.4.1, i.e., unpartitioned contains files assembled according to which *bricks* occupy each HEALPixel rather than by which sky locations occupy each HEALPixel. Columns in the skies and suppskies files are either inherited from the Legacy Surveys

 $[\]frac{82}{82}$ As with other sky locations, this quantity is stored as NPERSDEG in the header of files produced by desitarget.

⁸³ For example, the appropriate url for the Main Survey is https://data.desi. lbl.gov/public/ets/target/catalogs/dr9/1.1.1/skies.

⁸⁴ For example, Main Survey supplemental skies are at https://data.desi.lbl. gov/public/ets/target/catalogs/gaiadr2/1.1.1/skies-supp/.

sweep catalogs, are provided in Table 4, or were introduced in Section 4.4.1.

4.4.4. Masking and "Safe" Locations

Locations assigned near very bright stars are potentially unusable for sky subtraction by the DESI spectroscopic pipeline. To help characterize such locations for SV and the Main Survey, the desitarget code⁸⁵ was used to compile a "bright-star" catalog using all Gaia DR2 sources, supplemented with sources from the Tycho 2 (Høg et al. 2000) catalog.⁸⁶ Any matches between Gaia and Tycho 2 within a radius of 10" are first removed from the Tycho 2 catalog to prevent duplication of sources. Then, Gaia-Tycho coordinate matches are performed, after first shifting the location of Gaia sources to the Tycho epoch using Gaia astrometric parameters. Gaia sources with missing proper motions are assigned information from URAT using the same routine as detailed for GFAs (see Section 4.3). This catalog is then limited to only sources with a magnitude <12, where the magnitude that is used is Gaia G, Tycho V_T , Tycho H_P , and Tycho B_T , in that order of preference -because G is not measured for Tycho-only sources, V_T is not measured for every Tycho source, etc.

A "bright-star mask" is then constructed around all sources in this bright-star catalog by shifting sources to an epoch of 2023.0 to roughly match the expected central time of the DESI survey. Any sky locations within 5" of this (circa 2023) brightstar mask are assigned the bit IN BRIGHT OBJECT from the desi mask bitmask in the DESI TARGET column (see Section 2). Similarly, any sky locations within 10'' of a bright star are assigned the NEAR_BRIGHT_OBJECT bit.

Given the extensive sky coverage of the DESI survey, it is impossible to completely avoid placing fibers near bright stars. So, to assist with finding reasonable locations for sky subtraction in regions near bright stars, both the pixel-based and supplemental skies are augmented by additional "safe" locations. Any sky location that has IN_BRIGHT_OBJECT set is supplemented with six additional, symmetrically placed locations (see Figure 2) around the periphery of the circular mask for any individual star. This placement pattern corresponds to approximately one safe location per radius of the mask (with a margin of one extra location), which should be sufficient to ensure a safe location for the vast majority of fibers in the DESI focal plane. These safe locations all have the BAD_SKY bit set. They are also assigned the bits IN_BRIGHT_OBJECT and NEAR_BRIGHT_OBJECT using the bright-star mask (as for any other sky location).

Finally, any pixel-based or supplemental sky locations that have both IN_BRIGHT_OBJECT and BAD_SKY set (including "safe" locations) are removed completely from their respective desitarget output files. This step ensures that any freshly generated safe locations are not, themselves, inside the boundary of a bright star (again, see Figure 2). Sky locations with other combinations of the sky and bright object bits set are retained, but with all of the bits set in the DESI TARGET column to help downstream code identify potentially problematic positions.





Figure 2. Example blank-sky locations. Crosses depict supplemental sky locations (SUPP_SKY; see Section 4.4.2). Large blue circles represent masks around bright stars, and red diamonds depict "safe" locations generated when a SUPP_SKY falls in a mask (BAD_SKY; see Section 4.4.4). SUPP SKY locations in masks are not discarded, but have the IN_BRIGHT_OBJECT bit set (green circles). The masks in the top-right corner demonstrate how safe locations that have both BAD_SKY and IN_BRIGHT_OBJECT set are discarded: safe (BAD_SKY) locations were generated around the left-hand star mask based on a SUPP_SKY in that mask, but half fell in the right-hand star mask and so were discarded.

4.5. Monte Carlo Sampling of the Target-selection Footprint

To help facilitate large-scale structure analyses with DESI, the desitarget pipeline produces Monte Carlo samples that attempt to recreate the angular selection function of the Legacy Surveys imaging that was used to select DESI targets. These "random catalogs," and related products, are detailed in this section.

4.5.1. Random Catalogs

The main desitarget module used to create random catalogs is called desitarget.randoms and the main worker function is randoms.guantities_at_positions_in_a_brick(). Initially, a desired density of random points (deg^{-2}) is specified.⁸⁷ Then, for each brick in the Legacy Surveys, the desired density is multiplied by the $(\sim 0.0623 \text{ deg}^2)$ brick area, with the resulting number being modified by a (random) draw from the Poisson distribution to produce a final number of points to be sampled in a given brick. Locations (in RA and decl.) are then generated in the brick according to that final number of points.

The desitarget pipeline then queries the Legacy Surveys imaging at the pixel level, by directly sampling information from the CCDs used to produce the Legacy Surveys. In general, the files used as inputs to desitarget in this context are from the Legacy Surveys coadded stack files.⁸² Each random location is converted to a pixel using the World Coordinate System information for a given brick-based stack. The quantities produced for the DESI random catalogs are described in the documentation for the Legacy Surveys.

 $^{^{\}overline{87}}$ This density appears in the header of output random catalogs as DENSITY. ⁸⁸ The stacks are documented at https://www.legacysurvey.org/dr9/files/ #image-stacks-region-coadd.
 ⁸⁹ Specifically, under "random catalogs" at https://www.legacysurvey.org/

dr9/files/#random-catalogs-randoms.



Figure 3. An example use of the desitarget "allsky" random catalogs. Colors distinguish different numbers of g-band observations in DR9 of the Legacy Surveys. For instance, areas colored green (pink) were covered by two (\geq 5) observations. Areas that have zero observations because they are outside of the Legacy Surveys are immediately obvious, as is the lower coverage near the edge of the footprint. The division between BASS and MzLS at a decl. of ~32°.375 is also apparent (see also Section 4.1.3). It is trivial to use these random catalogs to characterize how much area in the Legacy Surveys match specified observational criteria.

Flux-related columns are derived by integrating over an aperture of a specified radius, as described in Section 4.4.1. The chosen aperture radius is recorded as APRAD in the header of an output random catalog, and is likely to always be 0."75, corresponding to the radius of a DESI fiber.

In addition to the random catalogs described above, the desitarget pipeline can produce randoms catalogs in bricks that lie *outside* of the Legacy Surveys footprint. Such catalogs only include a limited number of columns, as most of the informative quantities are not defined where imaging does not exist. The "inside" and "outside" catalogs can also be combined by desitarget to produce all-sky catalogs, which are particularly useful for determining, for example, the areal coverage of the Legacy Surveys (see Figure 3). Again, columns present in these additional catalogs are detailed in the Legacy Surveys documentation.

As for other files produced by desitarget, random catalogs are available online at the urls specified in Section 4.1.1. For random catalogs, the directory names terminate with randoms (instead of, e.g., targets) and the most recent desitarget version number used to run a substantial set of random catalogs is 0.49.0.90 Versions of the random catalogs that have and have not been resolved according to the criteria in Section 4.1.3 are available in the resolve and noresolve subdirectories. A randomsall subdirectory may also exist, which contains a larger file created by concatenating multiple smaller random catalogs to increase sampling density, as outlined in an accompanying README file.

Within the randoms/resolve directory, the inside-the-Legacy-Surveys random catalogs have the form randoms-ISEED-ISPLIT.fits, the outside-the-Legacy-Surveys random catalogs have the form randoms-outside-ISEED-ISPLIT.fits, and the combined random catalogs have the form randomsallsky-ISEED-ISPLIT.fits. Here, ISEED refers to the random seed used to produce the catalog and ISPLIT is an integer specifying a particular subset of randoms; the catalogs are split into smaller files from a larger parent set (which is itself randomized) during processing to make them more manageable. The specific integer seed (ISEED) used to produce a given file is stored in the file header as SEED.

4.5.2. Pixelized Weight Maps

The desitarget pipeline—specifically the desitarget.randoms.pixmap() function—can be used to combine files of DESI targets with random catalogs in a convenient HEALPixel-based format to produce maps of important quantities derived from the Legacy Surveys imaging. The resulting "pixweight" files have been made publicly available for both SV3⁹¹ and the Main Survey.⁹² Maps of dark-time and bright-time targets are created separately, and so each release directory contains dark and bright subdirectories. The HEALPixel resolution of each pixweight file is nside = 256, which is stored in the header of the file as HPXNSIDE. This resolution corresponds to a pixel area of ~0.0525 deg², which approximates the size of a Legacy Surveys brick.

The pixweight maps contain the quantities detailed in Table 10, which are all easily derived from the random catalogs discussed in Section 4.5.1 (typically from the file in the randomsall subdirectory) and the bright- and dark-time target files described in Section 4.1.1. One column that appears in Table 10 that we have not previously described is STARDENS, which contains a measure of the stellar density in a given HEALPixel. The STARDENS quantity is derived from a count of sources in Gaia DR2 that have a Gaia *G*-band magnitude in the range $12 \leq G < 17$ and that are point-like. Point-like, here, is defined using Gaia quantities via

$$AEN == 0. OR \log 10(AEN) < 0.3G-5.3,$$
 (3)

where AEN is the Gaia quantity ASTROMETRIC_EXCESS_ NOISE and *G* is the Gaia *G*-band magnitude (i.e., PHOT_G_MEAN_MAG).

4.5.3. Quality Assurance Web Pages

The "pixweight" files described in Section 4.5.2 can be used to derive useful plots of sky coverage for targets. In particular, the FRACAREA quantity allows target densities to be correctly weighted by the actual areal coverage of the Legacy Surveys imaging. In addition, as the "pixweight" files collate information about observational properties of the imaging surveys, such as depth and local stellar density, they can be used to determine how target densities correlate with systematic effects (see, e.g., Chaussidon et al. 2022b).

The desitarget. QA module can be used to combine the "pixweight" files from Section 4.5.2 with the target files detailed in Section 4.1.1 to create web pages of diagnostic plots for DESI targeting. The gateway to these web pages also includes a version of the target density matrices provided in

 $[\]frac{90}{90}$ Le., the catalogs are at https://data.desi.lbl.gov/public/ets/target/ catalogs/dr9/0.49.0/randoms/.

⁹¹ E.g., https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/0.57.0/pixweight/ sv3/resolve/bright/sv3pixweight-1-bright.fits.

⁹² E.g., https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/1.1.1/ pixweight/main/resolve/dark/pixweight-1-dark.fits.

THE ASTRONOMICAL JOURNAL, 165:50 (22pp), 2023 February



Figure 4. Example quality assurance diagnostic plots derived from running the desitarget code base on DR9 of the Legacy Surveys. Left panel: number density contours for DESI Main Survey LRG (dark-time) targets in an optical color (r - z) compared to the 3.4 and 4.6 μ m bands (W1 – W2) from the WISE survey. Right panel: density of all SV3 BGS targets in equatorial coordinates. The solid black line depicts the plane of our Galaxy and the red dotted line depicts the ecliptic.

 Table 10

 Columns in the Pixweight Files

Column Name	Python Format	Description
HPXPIXEL	i4	HEALPixel integers at an input nside (typically nside = 256).
FRACAREA	£4	Fraction of the HEALPixel with ≥ 1 observation in any band. ^a
STARDENS	£4	Density (deg^{-2}) of stars in the HEALPixel from Gaia DR2 (see Section 4.5.2).
EBV	£4	Galactic dust extinction $(E[B-V])^{b}$
PSFDEPTH_X	£4	$\frac{5}{\sqrt{\text{PSFDEPTH}_X}}$ gives the 5 σ flux-detection-limit in nanomaggies for a point source. ^{b,c}
GALDEPTH_X	£4	As for PSFDEPTH_X but for a round, exponential galaxy of effective radius 0."45. ^{b,c}
PSFSIZE_X	£4	Weighted average point-spread function FWHM, in arcseconds. ^{b,d}
FRACAREA_X	£4	Fraction of HEALPixel with ≥ 1 observation in any band with MASKBITS = X. ^e
target name	£4	Target density in the HEALPixel for each target class bit-name (see Table 3).

Notes.

^a Derived from the NOBS_G, NOBS_R, NOBS_Z quantities in the random catalog.

^b This quantity is the median in the HEALPixel from the passed random catalog.

^c One each for optical and Wide-field Infrared Survey Explorer (WISE) bands (i.e., X = G, R, Z, W1, W2), where the WISE values are AB, not Vega.

^d One for each optical band (i.e., X = G, R, Z).

^e This is a bitwise OR, so, e.g., if X = 7 then it is the fraction for $2^0 | 2^1 | 2^2$. MASKBITS is described in the Legacy Surveys documentation at https://www.legacysurvey.org/dr9/bitmasks/.

Tables 5 and 6, which are trivial to derive from the "pixweight" files.

These QA web pages have been made publicly available by the DESI collaboration for SV3⁹³ and the Main Survey.⁹⁴ At each url, there are individual subdirectories for bright-time and dark-time targets. Clicking on a subdirectory launches a page that includes target density matrices, plots of observational systematics, and a list of targeting bit-names (see, e.g., Table 3). Clicking on any bit-name produces a page with a number of diagnostic plots for that target class, including sky maps, color–color plots, and trends in target density with observational systematics. Some examples of these diagnostic plots are included in Figure 4.

4.6. Secondary Target Classes

In addition to the dark-time targets, bright-time targets, backup targets, calibration targets, and random catalogs

discussed elsewhere in this section, the DESI survey incorporates additional targets to pursue science efforts beyond the primary DESI goals. These targets, which we refer to as "secondary" targets, are either conducted as dedicated campaigns or as "filler" programs to mop up spare fibers where no primary DESI targets are available. The dedicated campaigns include tertiary programs (see Section 3.2.3) and some large ToO campaigns (see Section 3.2.2 and, e.g., Palmese et al. 2021).

4.6.1. An Overview of Secondary Targets

In Section 2 we discussed how secondary targets can be retrieved from the DESI target files (Section 4.1.1) by applying the scnd_mask bit-names to bit-values in columns that resemble SCND_TARGET. In Section 3.2 we detailed the specific, unique TARGETID that desitarget assigns to secondary targets. However, a full description of all of the scientific goals that underlie the DESI secondary target programs is beyond the scope of this paper and will be reserved for forthcoming DESI Data Release papers (e.g., DESI

 ⁹³ https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/0.57.0/QA/sv3/
 ⁹⁴ https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/1.1.1/QA/main/

Collaboration et al. 2023, in preparation). We will, however, provide an overview of the process for selecting secondary targets here.

Every secondary program provides lists of targets which are preserved for posterity in a specific directory.⁹⁵ This "secondary directory" contains subdirectories for SV1, SV3, the Main Survey, and some bespoke programs (sv1, sv3, main, bespoke; no secondary targets were assigned during SV2). Secondary targets added after the Main Survey may also be present in additional directories named main2, main3, etc. Each of these subdirectories includes a file of notes about each secondary target class; a docs directory of text files or Jupyter notebooks documenting a particular secondary program; an indata directory of the targets supplied by a secondary program to be processed by desitarget; and an outdata directory containing files for bright-time (bright) and darktime (dark) secondary targets that include the TARGETID, DESI_TARGET, and SCND_TARGET assigned by desitarget (again, see Sections 2 and 3.2). The outdata directory contains one or more subdirectories, which correspond to the different versions of the desitarget pipeline that were run to produce secondary targets. Each secondary targeting file is named \$bit-name.fits, \$bit-name.txt, or \$bitname.ipynb, where \$bit-name is the specific bit-name assigned to a secondary program in the scnd mask (see Section 2).

Files of targets provided by a secondary program were required to include the columns RA, DEC, PMRA, PMDEC, REF EPOCH, and OVERRIDE. The first five of these columns have the same meaning as for the Legacy Surveys sweep catalogs.⁹⁶ The OVERRIDE column, which was required to be set to True or False for each secondary target for each program, is used by desitarget to determine whether a secondary target should be merged with a primary target. When a secondary target is merged with a primary target by desitarget (corresponding to OVERRIDE==False), it inherits all of the properties of that primary target (such as the target's coordinates, observational priority, whether it will be observed in bright time or dark time, the primary TARGETID, etc.). If, on the other hand, OVERRIDE is set to True for a target, then the target is free to be observed as defined by the secondary program but a new TARGETID is generated for the secondary target (see Section 3.2). Secondary and primary targets are merged by performing a coordinate match with a 1''separation.97

One benefit to a secondary program of specifying OVER-RIDE==False for a secondary target that matches a primary target is that such targets are easier to track, because the same TARGETID will be assigned for the secondary and primary target throughout the DESI program. This choice also guarantees standard processing for a secondary target (such as being automatically coadded by the spectroscopic pipeline in the same manner as for primary targets). But, a major drawback of providing targets with OVERRIDE==False is that the target exactly resembles any matching DESI primary target, meaning that bespoke observational strategies—such as observing a target 20 times, or observing a match to a bright-time primary target in dark time—cannot be adopted. Secondary targets that specified OVERRIDE==False are present in the standard dark-time and bright-time files detailed in Section 4.1.1. Whether a target is both a secondary and a primary target can be determined by checking if the SCND_ANY bit from the desi_mask is set in the DESI_-TARGET column (see Section 2). Secondary targets that specified OVERRIDE==True⁹⁸ are written to their own monolithic files. These files are in the same root directory as detailed in Section 4.1.1 for primary targets, except that instead of being stored in the resolve directory they are in the secondary directory. Bright-time standalone secondary targets are compiled in the secondary/bright/targets-bright-secondary.fits file and dark-time standalone secondary targets are in the secondary/dark/ targets-dark-secondary.fits file.

Secondary targets that have OVERRIDE set to False are not necessarily as carefully vetted as primary targets. So, the desitarget code takes several precautions to ensure that DESI fibers are not placed on very bright OVERRIDE targets that could contaminate the spectra of primary targets. First, any OVERRIDE targets that lie inside of the bright-star mask discussed in Section 4.4.4 are removed completely from the output files of secondary targets. Second, when secondary targets are matched to primary targets (see above), they are also matched to the Legacy Surveys sweep catalogs.⁹⁹ During this match, information from Gaia and the Legacy Surveys is added to the secondary targeting file, including flux and magnitude quantities (FLUX_G, FLUX_R, FLUX_Z, GAIA_PHOT_G_ MEAN MAG, GAIA PHOT BP MEAN MAG, and GAIA PHOT_RP_MEAN_MAG). If an OVERRIDE target is brighter than 16th magnitude in any of these Gaia or Legacy Surveys bands, it is removed completely from all of the output files of secondary targets.

4.6.2. Is This Target Primary?

The introduction of secondary programs and ToOs somewhat complicate the provenance of a given target—particularly as secondary programs were allowed to be merged with primary programs if OVERRIDE was set to False. So, it is useful to provide a simple recipe for applying the information in Sections 2 and 3 to determine whether a target is part of the primary DESI campaign or is more exotic. One possible recipe is as follows:

- 1. Check the DESI_TARGET column for a target, or the equivalent SVX_DESI_TARGETcolumn (see Section 2).
- 2. If DESI_TARGET *only* has the bit SCND_ANY set (i.e., if it is equal to 2^{62}), then the target is either a secondary or tertiary target, or a ToO.
 - (a) Otherwise the target is primary (consult Section 4).
- 3. Use the decode_targetid() utility (see Section 3.3) to determine release for the target.
 - (a) If release < 1000 then the target is a general secondary target (see Equation (1) for the different meanings of release and DESI Collaboration et al. 2023, in preparation, for a description of the secondary programs).

⁹⁵ Publicly available at https://data.desi.lbl.gov/public/ets/target/secondary.

⁹⁶ https://www.legacysurvey.org/dr9/files/#sweep-catalogs-region-sweep

⁹⁷ This separation is recorded in output desitarget secondary files as SEP.

⁹⁸ These targets are sometimes referred to as "standalone" secondaries in DESI parlance.

⁹⁹ https://www.legacysurvey.org/dr9/files/#sweep-catalogs-region-sweep

- (b) If release = 8888 then the target is a tertiary target (consult Section 3.2.3 and A. Raichoor et al. 2023, in preparation).
- (c) If release==9999 then the target is a ToO (consult Section 3.2.2 and E. Schlafly et al. 2023, in preparation).

4.7. How to Run the desitarget Pipeline

Researchers may wish to fully recreate a suite of desitarget output files in order to conduct large-scale structure studies in the context of DESI or to inform target selection for future sky surveys. To help facilitate such endeavors, we have provided a tutorial on GitHub, in the form of a Jupyter notebook, that includes a detailed overview of how each of the data products outlined in this section were produced for the DESI Main Survey using the desitarget pipeline.¹⁰⁰

5. Discussion: Important Known Issues

As with any evolving software, desitarget underwent several bug fixes during catalog creation. In this section, we try to detail some of the most prominent issues and discrepancies that users of the DESI targeting files might encounter.

5.1. Target Reproducibility

Perhaps the most important known issue is that output from the desitarget code was not wholly reproducible throughout the Main Survey. This was due to the fact that although primary targets typically supersede secondary targets (see Section 4.6) the desitarget code was written to allow MWS targets to bump each other, regardless of whether they were primary or secondary. Unfortunately, because some MWS targets have similar colors to primary targets-for instance, white dwarfs resemble some QSOs-and because MWS white dwarf targets had a relatively high observational priority, the desitarget pipeline fostered occasional glitches where an MWS target could bump a primary target intended for largescale structure analyses. Notably, in situations where multiple MWS targets matched a primary target during primarysecondary merging (Section 4.6), which MWS target was chosen as the match was not ordered. So, on repeated execution of the desitarget code a different MWS target could end up being merged with a particular primary target. The upshot of this bug was that, prior to version 1.1.1 of desitarget, a small subset of bright-time targets, and very occasional darktime targets, were selected differently each time desitarget was run.

This MWS prioritizing issue was noticed because a new, high-density target class (called MWS_FAINT) was designed to be introduced for the first iteration of Main Survey targets, which were processed with version 1.0.0 of desitarget. Due to an oversight, these MWS_FAINT targets were not included in the 1.0.0 target catalogs. When this oversight was corrected and the target catalogs were updated, it became clear that a slightly different subset of targets had been selected, even when the MWS_FAINT targets were redacted. Any reproducibility bugs were subsequently fixed and new targets were processed with the MWS_FAINT selection once more omitted—resulting in the version 1.1.1 targets discussed throughout this paper.

While this bug fix was being addressed, 157 Main Survey tiles were observed with version 1.0.0 targets before updating to version 1.1.1 of desitarget and producing the bright-time and dark-time targets that will be used throughout the remainder of the DESI survey. Fortunately, the scope of the differences between which 1.0.0 targets were scheduled to be observed and which 1.1.1 targets would have been scheduled to be observed is minor. Across the 157 initial tiles, a total of 270 bright-time targets differed in a manner that could have produced discrepancies in fiber placement in the DESI focal plane.¹⁰¹ Only four dark-time targets differed in such a way. Given that there were \sim 785,000 fibers that could have been assigned fibers across 157 tiles, any reproducibility issues moving from version 1.0.0 to 1.1.1 targets near the start of the DESI Main Survey should have a negligible effect on large-scale structure analyses.

5.2. SUBPRIORITY Reproducibility

Assigning fibers to targets during early DESI Main Survey observations revealed an additional issue that could affect the overall reproducibility of the DESI survey. As noted in Table 4, desitarget adds a SUBPRIORITY column to target files that is used to break ties in the event that the DESI fiberassignment code has to choose between two targets that are otherwise equivalent in observational priority. A similar SUBPRIORITY column is added by desitarget to the pixel-based and supplemental sky locations discussed in Section 4.4. As of version 1.0.0 of the desitarget pipeline, SUBPRIORITY values were added as part of the function that writes files to disk. This meant that if an external routine read a target (or sky) file produced by desitarget and then wrote it back out using the desitarget writing utilities then each SUBPRIORITY would be overwritten with a new value.

Rewriting SUBPRIORITY using the approach of version 1.0.0 of the desitarget pipeline was not a bug, per se, but it proved problematic downstream of desitarget. In particular, the software that assigns fibers to DESI targets (A. Raichoor et al. 2023, in preparation) uses the desitarget writing routines to produce its own output data files. To guard against accidentally overwriting SUBPRIORITY values throughout the Main Survey, the desitarget pipeline was updated to only overwrite zero-valued SUBPRIORITY entries -meaning that no values are accidentally overwritten when the DESI fiber-assignment pipeline (or any other software) uses the desitarget writing functions. Although this issue with SUBPRIORITY was fixed before the Main Survey proceeded with version 1.1.1 targets, differences would persist in how fibers were assigned on the 157 Main Survey tiles that used targets compiled with version 1.0.0 of the desitarget pipeline. To ameliorate any discrepancies, files were constructed that contained the mapping between SUBPRIORITY and TARGETID for the SUBPRIORITY values that were initially assigned to version 1.0.0 Main Survey targets.¹⁰²

¹⁰⁰ See https://github.com/desihub/desitarget/blob/2.5.0/doc/nb/how-torun-target-selection-main-survey.ipynb.

¹⁰¹ Note that the crucial point is not that 270 different targets *were* observed, merely that a maximum of 270 different targets *could have* been observed, potentially impacting the survey selection function for large-scale structure studies.

studies. ¹⁰² These files are publicly available in the https://data.desi.lbl.gov/public/ ets/target/catalogs/subpriority/fba-version-4.0.0/ directory.

Versions of the desitarget pipeline subsequent to 1.1.0 incorporate a command-line option that can use these mappings to enforce the values of SUBPRIORITY used on the 157 Main Survey tiles that proceeded with version 1.0.0 targets.¹⁰³

5.3. Targets with RELEASE of 9010 in SV1

Early survey validation (i.e., SV1) took input information from the DR9 Legacy Surveys sweep catalogs before they were patched due to a data-processing issue¹⁰⁴ that altered the imaging properties of some sources. The desitarget pipeline switched to using the "patched" DR9 sweeps as of version 0.49.0 of the desitarget code. Targets that might have different source properties because of this patching bug can be identified because they will have a RELEASE of 9010 in files produced by versions of desitarget prior to 0.49.0but a RELEASE of 9012 in later target files. Such targets will have similar coordinates but a different TARGETID, both because the RELEASE number changed for all such sources in the patched sweep catalogs *and* because the OBJID changed for some such sources (see, e.g., Figure 1).

6. Summary

In this paper, we have provided a detailed overview of the DESI target-selection pipeline. The exact algorithms used to classify specific subsets of DESI bright-time and dark-time targets are detailed in other papers (Chaussidon et al. 2022a; Hahn et al. 2022; Raichoor et al. 2022; Zhou et al. 2022; Cooper et al. 2022) but aspects of how these targeting algorithms are *implemented* are crucial to understanding how to work with DESI data. This paper also details the approaches used to classify targets to help calibrate DESI spectroscopy, such as GFA targets and blank-sky locations.

The DESI target-selection pipeline, which is known as desitarget, is publicly available on GitHub. As with any software product, some documentation of the desitarget code and the resulting products will be described online and in other papers related to the DESI end-to-end pipeline (E. Schlafly et al. 2023, in preparation; A. Raichoor et al. 2023, in preparation; Guy et al. 2022; S. Bailey et al. 2023, in preparation). As such, here we have tried to focus on an overview of the most important concepts related to the technical aspects of DESI target selection.

The desitarget pipeline sets bit-values in output data files in a number of different columns, which indicate whether a particular target meets specific selection criteria. The names of these columns resemble DESI_TARGET (for dark-time DESI targets), BGS_TARGET (for bright-time galaxies), MWS_TARGET (for Milky Way targets), and SCND_TARGET (for targets beyond the primary DESI campaigns). Values in these columns can be interpreted using bitmasks with names like desi_mask, bqs_mask, mws_mask, and scnd_mask.

Another important output of the desitarget pipeline is a unique TARGETID that allows each target to be tracked throughout the DESI survey. This TARGETID encodes additional, useful information about each DESI target, such as whether a target is selected from Legacy Surveys imaging or from Gaia, whether a target is a sky location or part of a random catalog, and whether a target is part of a secondary program or is a ToO.

It is important to note that different versions of the desitarget pipeline were used for different phases of the DESI survey, such as the Survey Validation phase—including the DESI One-Percent Survey—and the Main Survey phase. The evolving nature of desitarget spawned multiple complexities in target selection that we have attempted to document in this paper. Examples of such technical details include (but are not limited to):

- 1. That sometimes Gaia EDR3 data, instead of Gaia DR2 data, is incorporated by desitarget when processing certain target-selection classes.
- 2. That the URAT survey is used to help select GFA targets for sources that are missing astrometric information in Gaia DR2.
- 3. That a bright-star mask based on Tycho and Gaia is applied to all sky locations and secondary targets, and that secondary targets are further removed from any DESI program if they are too bright in Gaia or Legacy Surveys imaging.
- 4. That targets processed using version 1.0.0 of the desitarget pipeline were observed for the earliest part of the DESI Main Survey but that, for reproducibility reasons, DESI rapidly updated to observing targets processed with version 1.1.1 of the desitarget pipeline.

The desitarget code base is now mostly stable as regards producing the static files used for DESI target selection. Further modifications to desitarget might be expected in modules related to producing random catalogs for large-scale structure analyses, and to converting static files to the dynamic files used to monitor the state of a DESI target as it is observed (see also E. Schlafly et al. 2023, in preparation). Any updates will be documented in later works as the DESI survey progresses.

A.D.M. and J.M. were supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Numbers DE-SC0019022 and DE-SC0020086. A.P.C. is supported by a Taiwan Ministry of Education Yushan Fellowship and Taiwan Ministry of Science and Technology grant No. 109-2112-M-007-011-MY3.

This research is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DEAC0205CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; additional support for DESI is provided by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF's National Optical-Infrared Astronomy Research Laboratory; the Science and Technologies Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Science and Technology of Mexico (CONACYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: https://www.desi.lbl.gov/collaboratinginstitutions.

¹⁰³ As detailed at https://github.com/desihub/desitarget/blob/2.5.0/doc/ nb/how-to-run-target-selection-main-survey.ipynb in a Jupyter notebook tutorial.

¹⁰⁴ See https://www.legacysurvey.org/dr9/issues/#bricks-that-were-processedusing-the-burst-buffer-at-nersc.

The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing-Arizona Sky Survey (BASS), and the Mayall z-band Legacy Survey (MzLS). DECaLS, BASS, and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory. Legacy Surveys also uses data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys was supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the U.S. National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences, and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy. The complete acknowledgments can be found at https:// www.legacysurvey.org/acknowledgment/.

The authors are honored to be permitted to conduct scientific research on Iolkam Duag (Kitt Peak), a mountain with particular significance to the Tohono Oodham Nation.

This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/ gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC; https://www.cosmos.esa.int/web/gaia/ dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

This research made use of Photutils, an Astropy package for detection and photometry of astronomical sources (Bradley et al. 2019). Some of the results in this paper have been derived using the healpy and HEALPix packages.

Facilities: Blanco (DECam), Bok (90Prime), Gaia, Mayall (DESI, Mosaic-3), NEOWISE, WISE

Software: astropy (Astropy Collaboration et al. 2013, 2018), fitsio (https://github.com/esheldon/fitsio), healpy (Zonca et al. 2019), matplotlib (Caswell et al. 2021), numpy (Harris et al. 2020), photutils (Bradley et al. 2019), pyyaml (https://pyyaml.org/), scipy (Gommers et al. 2021).

Data Availability

Data generated or analyzed as part this work are publicly available at URLs included in the text of the paper. Data points and code to produce the figures from this paper are at https://zenodo.org/record/7015786.

ORCID iDs

John Moustakas thttps://orcid.org/0000-0002-2733-4559 Stephen Bailey thttps://orcid.org/0000-0003-4162-6619 Andrew P. Cooper thttps://orcid.org/0000-0001-8274-158X Jaime E. Forero-Romero https://orcid.org/0000-0002-2890-3725

Bela Abolfathi b https://orcid.org/0000-0003-1820-8486 David M. Alexander b https://orcid.org/0000-0002-5896-6313

David Brooks b https://orcid.org/0000-0002-8458-5047 Edmond Chaussidon b https://orcid.org/0000-0001-8996-4874

Chia-Hsun Chuang https://orcid.org/0000-0002-3882-078X Kyle Dawson https://orcid.org/0000-0002-0553-3805 Arjun Dey https://orcid.org/0000-0002-4928-4003 Biprateep Dey https://orcid.org/0000-0002-5665-7912 Govinda Dhungana https://orcid.org/0000-0002-5402-1216 Enrique Gaztañaga https://orcid.org/0000-0001-9632-0815 Alma X. Gonzalez-Morales https://orcid.org/0000-0003-4089-6924

ChangHoon Hahn https://orcid.org/0000-0003-1197-0902 Hiram K. Herrera-Alcantar https://orcid.org/0000-0002-9136-9609

Mustapha Ishak https://orcid.org/0000-0002-6024-466X Tanveer Karim https://orcid.org/0000-0002-5652-8870 David Kirkby () https://orcid.org/0000-0002-8828-5463 Theodore Kisner (1) https://orcid.org/0000-0003-3510-7134 Sergey E. Koposov https://orcid.org/0000-0003-2644-135X Anthony Kremin https://orcid.org/0000-0001-6356-7424 Ting-Wen Lan ⁽ⁱ⁾ https://orcid.org/0000-0001-8857-7020 Martin Landriau https://orcid.org/0000-0003-1838-8528 Dustin Lang https://orcid.org/0000-0002-1172-0754 Michael E. Levi https://orcid.org/0000-0003-1887-1018 Lucas Napolitano https://orcid.org/0000-0002-5166-8671 Paul Martini ⁽¹⁾ https://orcid.org/0000-0002-4279-4182 Aaron Meisner https://orcid.org/0000-0002-1125-7384 Jeffrey A. Newman https://orcid.org/0000-0001-8684-2222 Nathalie Palanque-Delabrouille https://orcid.org/0000-0003-3188-784X

Will Percival **b** https://orcid.org/0000-0002-0644-5727 Francisco Prada **b** https://orcid.org/0000-0001-7145-8674 Anand Raichoor **b** https://orcid.org/0000-0001-5999-7923 Edward F. Schlafly **b** https://orcid.org/0000-0002-3569-7421 David Schlegel **b** https://orcid.org/0000-0002-5042-5088 Ting Tan **b** https://orcid.org/0000-0001-8289-1481 Gregory Tarle **b** https://orcid.org/0000-0003-1704-0781 Christophe Yèche **b** https://orcid.org/0000-0001-5146-8533 Rongpu Zhou **b** https://orcid.org/0000-0001-5381-4372 Zhimin Zhou **b** https://orcid.org/0000-0002-4135-0977 Hu Zou **b** https://orcid.org/0000-0002-6684-3997

References

Ahumada, R., Allende Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3 Alexander, D. M., Davis, T. M., Chaussidon, E., et al. 2022, arXiv:2208.08517

Allende Prieto, C., Cooper, A. P., Dey, A., et al. 2020, RNAAS, 4, 188

- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
- Bagnuolo, W. G. J., Furenlid, I. K., Gies, D. R., et al. 1990, PASP, 102, 604
- Besuner, R., Allen, L., Baltay, C., et al. 2020, Proc. SPIE, 11447, 1144710
- Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, AJ, 142, 31
- Blanton, M. R., Lin, H., Lupton, R. H., et al. 2003, AJ, 125, 2276
- Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, AJ, 144, 144
- Bovy, J., Hennawi, J. F., Hogg, D. W., et al. 2011, ApJ, 729, 141
- Bovy, J., Myers, A. D., Hennawi, J. F., et al. 2012, ApJ, 749, 41

Bradley, L., Sipőcz, B., Robitaille, T., et al. 2019, astropy/photutils: v0.7.2, Zenodo, doi:10.5281/zenodo.3568287

- Caswell, T. A., Droettboom, M., Lee, A., et al. 2021, matplotlib/matplotlib: REL: v3.5.1, Zenodo, doi:10.5281/zenodo.5773480
- Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2022a, arXiv:2208.08511
- Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2022b, MNRAS, 509, 3904
- Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039
- Colless, M., Peterson, B. A., Jackson, C., et al. 2003, arXiv:astro-ph/0306581
- Cooper, A. P., Koposov, S. E., Allende Prieto, C., et al. 2022, arXiv:2208. 08514
- Croom, S. M., Smith, R. J., Boyle, B. J., et al. 2004, MNRAS, 349, 1397
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016a, arXiv:1611. 00037
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016b, arXiv:1611. 00036
- DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, AJ, 164, 207
- Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
- Dey, A., Najita, J. R., Koposov, S. E., et al. 2022, arXiv:2208.11683
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018a, A&A, 616, A1
- Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018b, A&A, 616, A10
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1
- Gommers, R., Virtanen, P., Burovski, E., et al. 2021, scipy/scipy: SciPy 1.7.3, Zenodo, doi:10.5281/zenodo.5725464
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
- Guy, J., Bailey, S., Kremin, A., et al. 2022, arXiv:2209.14482
- Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2022, arXiv:2208.08512
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
- Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27
- Kirkpatrick, J. A., Schlegel, D. J., Ross, N. P., et al. 2011, ApJ, 743, 125
- Lan, T.-W., Tojeiro, R., Armengaud, E., et al. 2022, arXiv:2208.08516
- Lang, D., Hogg, D. W., & Schlegel, D. J. 2016, AJ, 151, 36
- Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv:1308.0847
- Lewis, I. J., Cannon, R. D., Taylor, K., et al. 2002, MNRAS, 333, 279

- Loveday, J., Peterson, B. A., Efstathiou, G., & Maddox, S. J. 1992, ApJ, 390, 338
- Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., & Kent, S. 2001, in ASP Conf. Ser. 238, Astronomical Data Analysis Software and Systems X, ed. J. Harnden, F. R. F. A. Primini, & H. E. Payne (San Francisco, CA: ASP), 269
- Meisner, A. M., Abareshi, B., Dey, A., et al. 2020, Proc. SPIE, 11447, 1144794
- Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, ApJS, 221, 27
- Palanque-Delabrouille, N., Yeche, C., Myers, A. D., et al. 2011, A&A, 530, A122
- Palanque-Delabrouille, N., Magneville, C., Yèche, C., et al. 2016, A&A, 587, A41
- Palmese, A., BenZvi, S., Bailey, S., et al. 2021, GCN, 30923, 1
- Perruchot, S., Blanc, P.-É., Guy, J., et al. 2020, Proc. SPIE, 11447, 1144786
- Poppett, C., Jelinsky, P., Guy, J., et al. 2020, Proc. SPIE, 11447, 1144711
- Raichoor, A., Eisenstein, D. J., Karim, T., et al. 2020, RNAAS, 4, 180
- Raichoor, A., Moustakas, J., Newman, J. A., et al. 2022, arXiv:2208.08513
- Ross, N. P., Myers, A. D., Sheldon, E. S., et al. 2012, ApJS, 199, 3
- Ruiz-Macias, O., Zarrouk, P., Cole, S., et al. 2020, RNAAS, 4, 187
- Shourt, W., Besuner, R., Silber, J., et al. 2020, Proc. SPIE, 11445, 114456J
- Silber, J. H., Fagrelius, P., Fanning, K., et al. 2023, AJ, 165, 9
- Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
- Smith, R. J., Croom, S. M., Boyle, B. J., et al. 2005, MNRAS, 359, 57
- Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485
- SubbaRao, M., Frieman, J., Bernardi, M., et al. 2002, Proc. SPIE, 4847, 452 Tonry, J., & Davis, M. 1979, AJ, 84, 1511
- Wells, D. C., Greisen, E. W., & Harten, R. H. 1981, A&AS, 44, 363
- Yèche, C., Palanque-Delabrouille, N., Claveau, C.-A., et al. 2020, RNAAS, 4, 179
- Yèche, C., Petitjean, P., Rich, J., et al. 2010, A&A, 523, A14
- York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579
- Zacharias, N., Finch, C., Subasavage, J., et al. 2015, AJ, 150, 101
- Zhou, R., Dey, B., Newman, J. A., et al. 2022, arXiv:2208.08515
- Zhou, R., Newman, J. A., Dawson, K. S., et al. 2020, RNAAS, 4, 181
- Zonca, A., Singer, L., Lenz, D., et al. 2019, JOSS, 4, 1298
- Zou, H., Zhou, X., Fan, X., et al. 2017, PASP, 129, 064101