Population Variation and Equity
Volume 1 · Number 1 · 2026

A Static Probe Appears Displaced When the Camera Pose Is Late

Wesley Scott[1] · Trevor Morgan[2] · James Harrison[3]

[1] Germany · [2] United Kingdom · [3] France

Original publication: November 3, 2026 | Article 615

Abstract

Apparent motion can arise from a mismatched camera pose even when the tested world point is fixed. We placed a constructed point three meters ahead of each selected recorded camera pose and reprojected it using poses delayed by 10 or 50 milliseconds. Across 52 TUM RGB-D files, equal-file mean normalized-coordinate error was 0.003270 at the smaller delay and 0.012774 at the larger one. The paired increase was 0.009504. These calculations isolate a timing sensitivity under supplied geometry; they neither detect a moving object nor evaluate a learned motion-separation system.

Recorded camera trajectories with constructed geometric probes. The imposed delays are experimental inputs, not measured synchronization errors in the public dataset.

A Static Probe Appears Displaced When the Camera Pose Is Late

> Recorded camera trajectories with constructed geometric probes. The imposed delays are experimental inputs, not measured synchronization errors in the public dataset.

Abstract

Apparent motion can arise from a mismatched camera pose even when the tested world point is fixed. We placed a constructed point three meters ahead of each selected recorded camera pose and reprojected it using poses delayed by 10 or 50 milliseconds. Across 52 TUM RGB-D files, equal-file mean normalized-coordinate error was 0.003270 at the smaller delay and 0.012774 at the larger one. The paired increase was 0.009504. These calculations isolate a timing sensitivity under supplied geometry; they neither detect a moving object nor evaluate a learned motion-separation system.

Timing belongs in a motion interpretation

An apparent displacement is often interpreted by considering whether the camera moved, whether the subject moved or whether both did. There is another possibility in a computation that uses externally supplied camera geometry: the pose can describe the camera at the wrong time. Even an otherwise accurate position and orientation then belong to a different projection configuration from the one whose image is being interpreted.

We examine how much normalized image-coordinate error can arise solely from deliberately using an earlier camera pose on real recorded trajectories. The test point is fixed in world coordinates during each comparison. We evaluate delays of ten and fifty milliseconds on a common cohort, keeping the point construction and projection equation identical. The purpose is a sensitivity calculation with known inputs rather than a diagnosis of a dataset's clocks.

Meng, 2026 addresses camera-subject motion disentanglement through geometry-semantics orthogonal attention. We implement none of its attention operations. The relevant issue is what can produce a geometric residual before that residual is interpreted as subject movement. Our experiment supplies a timing-only explanation in a constructed case, without claiming that the target system makes this attribution or would fail under the imposed delays.

The controlled comparison also differs from a real false-alarm study. No object detector or decision threshold labels a subject as moving. We calculate the size of a projection error and retain the information that produced it. A later system-level evaluation would need observed images, actual subject motion labels and a stated decision rule before counting erroneous motion detections.

Recorded poses and a documented coordinate convention

The camera paths come from the public TUM RGB-D reference-trajectory text subset (Sturm et al., 2012). Its official format records the color camera's optical-center position and orientation in a motion-capture world frame (TUM Computer Vision Group, 2026). The quaternion order is explicitly given as its vector components followed by the scalar component. We retain that convention in the numerical rotations.

The acquired subset has 53 linked files, of which one contains no numerical poses. We use the other 52 after a documented unique-time curation: all 149 rows at 74 timestamps with conflicting pose values are

removed, with original source files preserved. The resulting archive contains 615,775 poses. No RGB images, depth maps or estimated SLAM trajectories are required or obtained for this calculation.

These files cover different motion and calibration roles and can share equipment or environment. They are not fifty-two independently sampled camera systems. Their role here is to provide concrete position and rotation histories for a geometric perturbation. The collection's public availability and its large pose count do not supply a population sampling design.

Other resources separate acquisition and reference information differently. EuRoC supplies another real visual-inertial setting (Burri et al., 2016), while CARLA and Virtual KITTI 2 offer synthetic environments with their own annotations (Dosovitskiy et al., 2017; Yohann et al., 2020). We do not merge their records into this study. Their contrast helps locate the present evidence: recorded real camera poses combined with hypothetical, explicitly constructed points.

Constructing a point that has no motion to explain

Every hundredth retained pose is a candidate anchor, beginning with the first record of each file. At an anchor, we take a camera-coordinate point [0, 0, 3] meters and transform it into world coordinates using that anchor's recorded pose. Its correct normalized projection at the anchor is the image-plane origin, up to numerical precision.

That world point is then held fixed while the earlier camera poses are used. Each anchor creates its own probe point; the full set is not a persistent physical object tracked across the recording. The construction also does not establish that a real surface lies three meters ahead or that it would be visible in an image. It is an analytic geometry probe whose depth and direction are chosen by design.

Multi-view geometry supplies the camera projection relationship (Hartley and Zisserman, 2004). Rigid-state estimation and transformation conventions clarify how a local point is mapped to the world and back (Barfoot, 2017; Blanco-Claraco, 2021). We calculate that mapping directly from the supplied pose. There is no depth network, point-cloud reconstruction or fitted object trajectory behind the probe location.

The fixed three-meter distance is consequential. Translation-induced projection changes depend on depth, while rotations affect bearing in another way. We do not generalize the measured error to every scene depth or estimate separate causal contributions from the scalar norm. A different distance would define another controlled geometry and require its own reported results.

Looking up the deliberately stale poses

For each candidate anchor, we request camera poses at ten and fifty milliseconds before its timestamp. Both earlier lookups must be supported for the anchor to enter either condition. This common-cohort requirement prevents the larger delay from being evaluated on a different set of camera motions simply because its history is harder to obtain.

If a lookup coincides with a recorded timestamp within two microseconds, we use that pose. Otherwise, the two neighboring recorded times must bracket the lookup and be at most fifty milliseconds apart. Translation is interpolated linearly; orientation uses spherical linear interpolation between the provided rotations. No extrapolation outside a file or interpolation across a wider observation gap is permitted.

Quaternion interpolation has a geometric foundation (Shoemake, 1985), and the local SciPy implementation documents its operation (SciPy developers, 2026; Virtanen et al., 2020). Provided quaternions are normalized by the rotation constructor. The largest deviation of their source norms from

one is about 0.0000941. We record that small representational adjustment rather than describe the rounded text as exact physical orientation truth.

The two-microsecond matching tolerance is a numerical convention for timestamp lookup, not a claim about sensor synchronization accuracy. The source values are represented as floating-point numbers, and time is shifted to a local origin for interpolation. Numerical analysis motivates stating the convention used near timestamp boundaries (Higham, 2002; Goldberg, 1991), especially when requested delays are small relative to the epoch timestamp itself.

There are 6,180 candidate anchors. Seventy lack the required common bounded history, leaving 6,110 probes. All 52 files contribute. We also require the fixed world point to have positive depth under both stale cameras; none of the otherwise eligible probes fails that check. The eligibility screen does not inspect which delay produces a more favorable error.

The measured error and its interpretation

For each delayed pose, we map the fixed world point into that camera's coordinates and divide the first two coordinates by depth. The Euclidean norm of the resulting two-dimensional vector is the error relative to the correct anchor projection. It is dimensionless in normalized image coordinates. No focal length is applied, so it is not a pixel distance.

The correctly timed projection has a maximum residual norm of approximately 2.92e-16. This validates the implemented point construction and coordinate reversal at the executed precision. It does not test the real reference system's absolute accuracy. The induced delayed errors result from using a different recorded camera configuration, with no world-point movement introduced in the calculation.

The zero-delay control is particularly limited because the world point is constructed from the same anchor pose used to project it back. Numerical cancellation is expected. It cannot independently establish that the anchor orientation or location is physically exact, nor can it estimate an unknown clock offset. The imposed delays remain sensitivity inputs throughout the study, rather than parameters inferred from a zero-residual criterion.

Trajectory-evaluation guidance distinguishes temporal correspondence from spatial alignment (Zhang and Scaramuzza, 2018). Our perturbation deliberately violates the former while retaining a consistent world frame. A spatial alignment operation would not in general make two poses from different times represent the same camera configuration. The experiment therefore addresses a different problem from estimating a constant transformation between trajectory coordinate systems.

Equal-file results

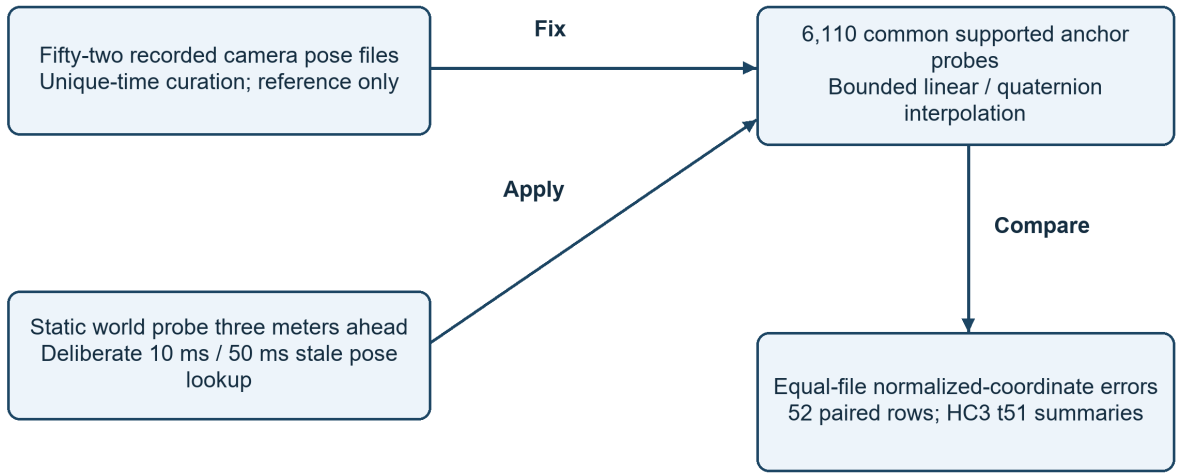


Figure 1. *An anchor pose defines a static three-meter probe. Two bounded earlier-pose lookups reproject that same point; common probes are averaged within file before the two descriptive regressions are fitted.*

We first average each delay's probe errors within each file. Model 1 estimates the equal-file mean error at ten milliseconds. Model 2 estimates the within-file mean increase from ten to fifty milliseconds. Both models contain an intercept and 52 file summaries, with HC3 standard errors and t51 confidence intervals. They use no additional covariates or file effects.

Robust covariance states a working variance assumption for these file summaries (White, 1980). Shared acquisitions and devices may create dependence outside that assumption, and clustering would need an appropriate independent grouping (Cameron and Miller, 2015). The intervals condition on the source files, curation, probe geometry and delay choices; they do not propagate motion-capture or interpolation uncertainty.

Table 1. *Normalized-coordinate error summaries, N=52 files, intercept-only OLS, HC3 covariance and t51 intervals. M1 reports ten-millisecond error; M2 reports the paired fifty-minus-ten-millisecond increase. Intercept-only R-squared is numerically zero.*

Model	Term	Estimate (SE)	95% CI	N
M1	Mean outcome/contrast	0.003270 (0.000236)	[0.002797, 0.003744]	52
M1: fit	R-squared = 0.0000	-	-	52
M2	Mean outcome/contrast	0.009504 (0.001019)	[0.007458, 0.011550]	52
M2: fit	R-squared = 0.0000	-	-	52

Model labels: M1 = Mean normalized-bearing error under 10 ms pose delay; M2 = Increase in normalized-bearing error from 10 ms to 50 ms.

The equal-file mean error at ten milliseconds is 0.00327028, with standard error 0.00023573 and interval [0.00279704, 0.00374353]. At fifty milliseconds, the descriptive mean is 0.01277448. The paired increase is 0.00950419, standard error 0.00101921 and interval [0.00745804, 0.01155034]. The common source points make this a paired timing contrast rather than a comparison of unrelated scenes.

Pooling probes instead gives mean errors of approximately 0.003038 and 0.010832. That calculation gives greater weight to files with more retained anchors. The primary equal-file summary gives a short eligible file the same importance as a longer one. Both are defined quantities, but their different weights should remain visible rather than be treated as independent replications.

The larger average error does not imply a monotonic increase for every point under every possible delay. Camera rotations can reverse, and translation and rotation can partly cancel in projection. The observed mean contrast belongs to the two requested time offsets and this cohort. We do not infer a universal linear relationship between milliseconds and apparent displacement.

Recorded pose differences can also contain reference-measurement variation in addition to smooth physical camera motion. No independent decomposition of those sources is available here. Interpolation supplies a value at a requested supported time but does not remove that uncertainty. Consequently, the result concerns delayed projections along the recorded reference process; it is not a clean estimate of the contribution of physical motion alone or a calibrated noise floor for another sensor.

Why the residual cannot identify subject motion

The probes remain stationary in world coordinates during every comparison. A nonzero projection residual is therefore generated entirely by the changed camera-pose time in this experiment. It should not be described as a measured subject velocity, a dense optical-flow field or evidence that an actual person or vehicle moved. Measurement-error reasoning helps preserve that distinction between the computed variable and an intended physical construct (Carroll et al., 2006).

The error's unit is also important (Stevens, 1946). A norm in normalized coordinates is neither an angle in radians nor a direct number of pixels. Converting it to another quantity would require the relevant geometric definition or intrinsics. We do not attach an unstated camera model to the values merely to make their practical magnitude sound more familiar.

Agreement methods encourage examining the difference being evaluated rather than relying on a broad correlation (Bland and Altman, 1986; Lin, 1989). Two pose series can follow closely related paths and still be mistimed at the moment needed for a projection. Their overall positional similarity would not validate the temporal association required by a motion interpretation.

Pose-estimation methods such as bundle adjustment, EPnP, ORB-SLAM2 and COLMAP address how camera geometry is recovered (Triggs et al., 2000; Lepetit et al., 2009; Mur-Artal and Tardos, 2016; Schonberger and Frahm, 2016). We do not run them, perturb their internal estimates or rank their synchronization robustness. All delayed poses come from the same curated reference series, so the controlled source of change is the lookup time.

The evidence needed for an operational motion system

A deployed system would need to know the timing relationship between its images and geometry, as well as the uncertainty in that relationship. Our ten- and fifty-millisecond delays are imposed conditions, not estimated dataset defects. The probe construction uses the correct anchor geometry deliberately, so the analysis is an offline sensitivity audit rather than an online method for recovering missing synchronization information.

Distribution-shift research cautions against promoting an internal perturbation result into a broad reliability guarantee (Ovadia et al., 2019). Validation should match the intended data and decision boundary (Roberts et al., 2017). A system-level follow-up would require actual image observations,

representative depths, subject-motion labels and a prespecified rule for interpreting residuals. The current study supplies none of those outcomes by implication.

Influential recordings should remain inspectable because a shared average can conceal very different camera dynamics (Cook, 1977). The saved probe table retains file identity, anchor index, lookup delays and computed depths. It also records the excluded-history count, preserving the fact that the result conditions on locally supported interpolation rather than complete temporal coverage of every recording.

Dataset documentation makes the constructed point and the source pose roles explicit (Gebru et al., 2021), while reusable provenance connects them to the original public text and curation policy (Wilkinson et al., 2016). The resulting evidence is a controlled numerical demonstration: on these recorded camera paths, a deliberately stale pose can produce an apparent static-point displacement substantial enough to require its own interpretation before any claim about subject motion.

References

1. Cameron, A. Colin, and Douglas L. Miller. “A Practitioner’s Guide to Cluster-Robust Inference.” *Journal of Human Resources*, vol. 50, no. 2, 2015, pp. 317-372. https://doi.org/10.3368/jhr.50.2.317.
2. Goldberg, David. (1991). What every computer scientist should know about floating-point arithmetic. ACM Computing Surveys. https://doi.org/10.1145/103162.103163.
3. Roberts, David R., et al. “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure.” *Ecography*, vol. 40, no. 8, 2017, pp. 913-929. https://doi.org/10.1111/ecog.02881.
4. Higham, Nicholas J.. (2002). Accuracy and Stability of Numerical Algorithms: Second Edition. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898718027.
5. White, Halbert. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, vol. 48, no. 4, 1980, pp. 817-838. https://doi.org/10.2307/1912934.
6. Barfoot, Timothy D.. (2017). State Estimation for Robotics. Cambridge University Press, first edition. https://doi.org/10.1017/9781316671528.
7. Burri, Michael; Nikolic, Janosch; Gohl, Pascal; Schneider, Thomas; Rehder, Joern; Omari, Sammy; Achtelik, Markus W; Siegwart, Roland. (2016). The EuRoC micro aerial vehicle datasets. The International Journal of Robotics Research. https://doi.org/10.1177/0278364915620033.
8. Lin, Lawrence I-Kuei. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics. https://doi.org/10.2307/2532051.
9. Gebru, Timnit, et al. "Datasheets for datasets" *Communications of the ACM*, 2021. https://doi.org/10.1145/3458723.
10. Ovadia, Yaniv, et al. “Can You Trust Your Model’s Uncertainty? Evaluating Predictive Uncertainty Under Dataset Shift.” *Advances in Neural Information Processing Systems*, vol. 32, 2019. https://arxiv.org/abs/1906.02530.
11. Mur-Artal, Raul; Tardos, Juan D.. (2016). ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. https://arxiv.org/abs/1610.06475.
12. Shoemake, Ken. (1985). Animating rotation with quaternion curves. Proceedings of the 12th annual conference on Computer graphics and interactive techniques - SIGGRAPH '85. https://doi.org/10.1145/325334.325242.
13. Meng, Zijie. (2026). OrthoMotion:Disentangling Camera and Subject Motion via Geometry Semantics Orthogonal Attention. arXiv:2606.22835. https://arxiv.org/abs/2606.22835.
14. Schonberger, Johannes L.; Frahm, Jan-Michael. (2016). Structure-from-Motion Revisited. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/cvpr.2016.445.
15. Cook, R. Dennis. (1977). Detection of Influential Observation in Linear Regression. Technometrics. https://doi.org/10.1080/00401706.1977.10489493.
16. Bland, J. Martin, and Douglas G. Altman. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet*, vol. 327, no. 8476, 1986, pp. 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8.
17. Sturm, Jürgen; Engelhard, Nikolas; Endres, Felix; Burgard, Wolfram; Cremers, Daniel. (2012). A benchmark for the evaluation of RGB-D SLAM systems. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. https://doi.org/10.1109/IROS.2012.6385773.

18. Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. “Measurement Error in Nonlinear Models.” 2006. https://doi.org/10.1201/9781420010138.

19. Triggs, Bill; McLauchlan, Philip F.; Hartley, Richard I.; Fitzgibbon, Andrew W.. (2000). Bundle Adjustment - A Modern Synthesis. Lecture Notes in Computer Science. https://doi.org/10.1007/3-540-44480-7_21.

20. Zhang, Zichao; Scaramuzza, Davide. (2018). A Tutorial on Quantitative Trajectory Evaluation for Visual(-Inertial) Odometry. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). https://doi.org/10.1109/IROS.2018.8593941.

21. Blanco-Claraco, José Luis. (2021). A tutorial on SE(3) transformation parameterizations and on-manifold optimization. arXiv:2103.15980; arXivversionyear, originaltechnicalreport2010. https://arxiv.org/abs/2103.15980.

22. TUM Computer Vision Group. (2026). RGB-D SLAM Dataset: File Formats [Official documentation; accessed8November2026, captureyearnotbenchmarkyear]. https://cvg.cit.tum.de/data/datasets/rgbd-dataset/file_formats.

23. Virtanen, Pauli; Gommers, Ralf; Oliphant, Travis E.; Haberland, Matt; Reddy, Tyler; Cournapeau, David; Burovski, Evgeni; Peterson, Pearu; Weckesser, Warren; Bright, Jonathan; van der Walt, Stéfan J.; Brett, Matthew; Wilson, Joshua; Millman, K. Jarrod; Mayorov, Nikolay; Nelson, Andrew R. J.; Jones, Eric; Kern, Robert; Larson, Eric; Carey, C J; Polat, İlhan; Feng, Yu; Moore, Eric W.; VanderPlas, Jake; Laxalde, Denis; Perktold, Josef; Cimrman, Robert; Henriksen, Ian; Quintero, E. A.; Harris, Charles R.; Archibald, Anne M.; Ribeiro, Antônio H.; Pedregosa, Fabian; van Mulbregt, Paul; SciPy 1.0 Contributors; Vijaykumar, Aditya; Bardelli, Alessandro Pietro; Rothberg, Alex; Hilboll, Andreas; Kloeckner, Andreas; Scopatz, Anthony; Lee, Antony; Rokem, Ariel; Woods, C. Nathan; Fulton, Chad; Masson, Charles; Häggström, Christian; Fitzgerald, Clark; Nicholson, David A.; Hagen, David R.; Pasechnik, Dmitrii V.; Olivetti, Emanuele; Martin, Eric; Wieser, Eric; Silva, Fabrice; Lenders, Felix; Wilhelm, Florian; Young, G.; Price, Gavin A.; Ingold, Gert-Ludwig; Allen, Gregory E.; Lee, Gregory R.; Audren, Hervé; Probst, Irvin; Dietrich, Jörg P.; Silterra, Jacob; Webber, James T; Slavič, Janko; Nothman, Joel; Buchner, Johannes; Kulick, Johannes; Schönberger, Johannes L.; de Miranda Cardoso, José Vinícius; Reimer, Joscha; Harrington, Joseph; Rodríguez, Juan Luis Cano; Nunez-Iglesias, Juan; Kuczynski, Justin; Tritz, Kevin; Thoma, Martin; Newville, Matthew; Kümmerer, Matthias; Bolingbroke, Maximilian; Tartre, Michael; Pak, Mikhail; Smith, Nathaniel J.; Nowaczyk, Nikolai; Shebanov, Nikolay; Pavlyk, Oleksandr; Brodtkorb, Per A.; Lee, Perry; McGibbon, Robert T.; Feldbauer, Roman; Lewis, Sam; Tygier, Sam; Sievert, Scott; Vigna, Sebastiano; Peterson, Stefan; More, Surhud; Pudlik, Tadeusz; Oshima, Takuya; Pingel, Thomas J.; Robitaille, Thomas P.; Spura, Thomas; Jones, Thouis R.; Cera, Tim; Leslie, Tim; Zito, Tiziano; Krauss, Tom; Upadhyay, Utkarsh; Halchenko, Yaroslav O.; Vázquez-Baeza, Yoshiki. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. https://doi.org/10.1038/s41592-019-0686-2.

24. Stevens, S. S.. (1946). On the Theory of Scales of Measurement. Science. https://doi.org/10.1126/science.103.2684.677.

25. Lepetit, Vincent; Moreno-Noguer, Francesc; Fua, Pascal. (2009). EPnP: An Accurate O(n) Solution to the PnP Problem. International Journal of Computer Vision. https://doi.org/10.1007/s11263-008-0152-6.

26. SciPy developers. (2026). scipy.spatial.transform.Slerp [API documentation; snapshot accessed 7 November 2026]. https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.transform.Slerp.html.

27. Yohann, Cabon,; Naila, Murray,; Martin, Humenberger,. (2020). Virtual KITTI 2. https://arxiv.org/abs/2001.10773.

28. Hartley, Richard; Zisserman, Andrew. (2004). Multiple View Geometry in Computer Vision. Cambridge University Press. https://doi.org/10.1017/cbo9780511811685.

29. Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; ’t Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://doi.org/10.1038/sdata.2016.18.

30. Dosovitskiy, Alexey; Ros, German; Codevilla, Felipe; Lopez, Antonio; Koltun, Vladlen. (2017). CARLA: An Open Urban Driving Simulator. arXiv:1711.03938. https://arxiv.org/abs/1711.03938.