NANORIGO

DELIVERABLE No.	D1.7
DELIVERABLE TITLE	Proposal for Guidance on Data
RESPONSIBLE AUTHOR	Dmitri Ciornii, Vasile-Dan Hodoroaba - BAM

D1.7 Proposal for Guidance on Data

The aim of D1.7 is to offer guidance on the suitability of scientific-technical data for risk governance which supports the user with information on data/information access, use, and "transformation" for specific purposes. This document compiles the outcomes of T1.1 (data sources), T1.2 (existing guidelines), T1.3 (data quality criteria), T1.5 (fitness-for-purpose of data), T1.6 (interviews with stakeholders), and T1.8 (SEIN aspects).

Currently, the full exploitation of data presents a considerable challenge despite a large quantity of data sources and repositories due to problems such as incompleteness, varied data formatting, and a lack of metadata.[1-2] Therefore, much of the currently available data, information, and knowledge cannot be properly reused. Furthermore, data with low reliability may lead to distrust and, consequently, to documents of dubious trustworthiness and wrong decisions with detrimental impact on a large (societal) scale. For the production of certain documents, such as meta-analyses, reviews, comments, opinions, responses, statements on proposed decisions, safety reporting, risk mitigation, and deciding on innovation or investment, users need trustful, reliable data. Moreover, the final products listed above need to be checked for quality and compliance with regulations. D1.7 is an attempt to offer solutions to the actual problems described above by employing a data readiness approach.

Different stakeholder interests may influence scientific practices and shape policy and commercial decisions. Furthermore, different stakeholders may assess data quality differently. For exemple, when assessing data fitness for industrial purposes, industrial companies pay attention to "*the level of quality required … in compliance with regulatory or customer expectations*"(D1.6). Here, data quality is defined at levels that fulfil 'fitness-for-purpose'. A solution might be improvement in communication and dissemination, and moderated dialogue between stakeholders which accounts for their different needs (analyses of stakeholders' expectations and opinions, perception, and education by dissemination).

The major aim of this deliverable is to provide a way to transform "raw" data into reliable data, information (data in context), and knowledge (information in context) for users. However, reliable data has to be defined. In close collaboration with other WPs from NANORIGO and NMBP-13 projects WP1 has elaborated an approach which easily assesses data maturity and

integrates data on nanomaterial biological effects into the Nano Risk Governance Framework (NRGF) and the Nano Risk Governance Council (NRGC).

NANORIGO WP1 (particularly D1.7) has developed a method (DRL approach) to transform all existing data and information (extracted from available databases) into operational knowledge for nano-risk decision making. Data Readiness Levels (DRLs) present a method for understanding the maturity of data during its acquisition phase for nano risk governance (NRG). Furthermore, the DRL approach allows the NRGC to assess the maturity (readiness) of particular data and compare data sets. It is a classification system for data of different complexities and has four inputs: raw data, tools, users, and the NRGC. The outputs are technical data readiness (**DRL1-3**), data readiness for specific/individual stakeholders (**DRL4-6**) and data readiness for risk governance (**DRL7-9**).
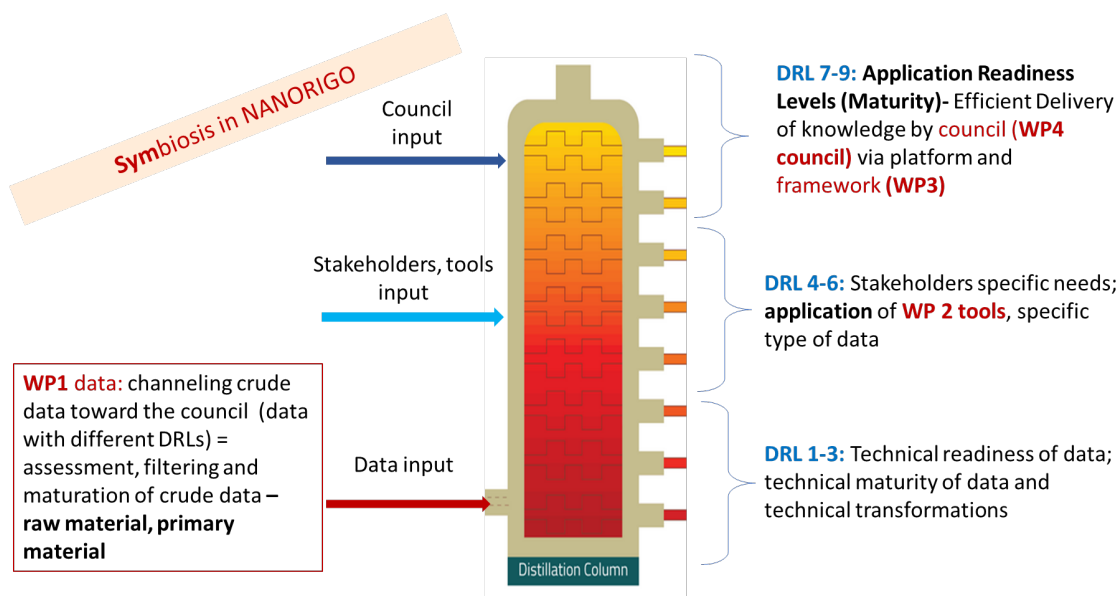
Figure 1 Schematic representation of the DRL "column" (in analogy with a refinery column). "Raw" data come into the DRL "column" at the beginning of the process and are step-by-step "refined" and transformed to obtain final products: technical, actionable data, stakeholder-specific types of data, actionable documents, and a "knowledge delivery system."

The DRLs are designed to support decision making and risk governance. Furthermore, DRLs support the users who need a service which will deliver actionable documents (refined data-related products). For this reason, the DRL approach will lead to different forms of refined data (in analogy with chemical risk assessment output):

1. Clear stand on the strength of evidence (certificate, statement on data readiness level)
2. Characteritics of the utility of data (fitness-for-purpose)
3. Guidance on how to reach higher readiness levels

4. Contribution to actionable documents (e.g., meta-analyses, reviews, weight-of-evidence approachs, and systemic reviews, achieved together with data users in a "co-creation" process).

Even if there are many good and reliable databases, and all of them provide useful information for different stakeholders, we conclude that at least in case of the user or "general public," an expert needs to be placed between the user and the data source/tool.

The novelties in the DRL approach are: i) DRL approach includes not only data, tools and NRGC input, but also users input (co-creation and co-refinement), all stakeholders are actively engaged in the risk-governance process; ii) Outputs: addressing stakeholders' specific needs regarding nanomaterial safety and risk, knowledge transfer for risk governance will be a significant novelty; iii) the way to operate: a symbiosis, iv) peer review and quality check by the NRGC of the actionable documents prepared by consultancy companies; v) another significant novelty issued by NANORIGO is the introduction also of a social dimension (perception, SEIN principles). By developing the DRLs, WP1 (in particular D1.7) offers something new, which is not present yet but highly demanded on the EU market.

References

1   Chao, T. C. Enhancing metadata for research methods in data curation. *Proceedings of the American Society for Information Science and Technology* **51**, 1-4, doi:https://doi.org/10.1002/meet.2014.14505101103 (2014).

2   Papadiamantis, A. G. *et al.* Metadata Stewardship in Nanosafety Research: Community-Driven Organisation of Metadata Schemas to Support FAIR Nanoscience Data. *Nanomaterials* **10**, 2033 (2020).

3   Willis, C., Greenberg, J. & White, H. Analysis and synthesis of metadata goals for scientific data. *Journal of the American Society for Information Science and Technology* **63**, 1505-1520, doi:https://doi.org/10.1002/asi.22683 (2012).