DCO

dco.uab.cat

# Data Compression Observatory (DCO)

## What it is

The DCO is a network of researchers, practitioners and users of data compression with the goal of improving the state of the art and facilitating its best usage.

Being part of the DCO costs nothing, demands no exclusivity, requires no minimum commitment and does not bind any member's institution.

## Who can join

Joining is free and open to anyone.

Whoever joins shares the purpose of the network, treats material shared within it on a non-public basis with discretion, and neither acts nor speaks in the name of the DCO without the consensus of its members.

English adherence · Adhesión en castellano

You can send your signed adherence forms to miguel.hernandez@uab.cat.

## Pillars

The DCO sets out to address four problems, and as many spaces to fill.

### 1. Curated tracking of the state of the art

The state of the art in data compression (what has been published and what is used in practice) is spread across journals of differing profiles and across repositories, private or public, belonging to institutions and agencies. Criteria on reproducibility, or on the use of datasets, vary from one venue to another, so that locating what is genuinely relevant within a given area, and knowing which part of it can be verified, takes considerable effort.

The DCO aims at a curated, accessible channel, organized by category (compression of images of a particular kind, for instance), gathering the relevant contributions, verified and with their experiments available, and open to anyone who wishes to subscribe. And at a qualified peer review supporting that selection, professionalized and dignified as the work it is.

### 2. Avoid repetition of negative results

Traveling a road others have already traveled without success is an avoidable waste of effort. For research to be efficient, it is necessary to know what has not worked.

The DCO aims at a curated repository of negative results, valued on a par with positive ones: the paths ruled out are one more element of the search space that research covers.

### 3. Standards

This pillar complements the first. Within each ecosystem, *de iure* and *de facto* standards coexist, competing or cooperating with one another, and a good part of them is published outside the academic circuit, which leaves them beyond that tracking. They are also, in many cases, hard to understand and to implement, and those who decide on compression are not always experts in compression: they design a compression stage within a larger system, in any field.

The DCO aims at bringing them into the tracking, by category of application, with homogeneous performance comparisons, with the available implementations of each and with the supporting documentation that helps to understand them. And at promoting their use, the CCSDS standards for instance, and at outreach and tools that make it possible to decide well without being an expert. All of it from the most impartial standpoint possible, including when the members of the network take part in the development of standards, on the academic or the applied side.

### 4. Common evaluation framework

This pillar runs across the other three. Without data prepared in common and without agreed protocols, results are neither comparable nor convincing beyond whoever publishes them.

The DCO aims at a repository dedicated to the verification, the evaluation and the comparison of compression techniques across all the categories covered, with datasets located, curated and prepared, published and kept available in the long term (that preparation is what makes results relevant), and with agreed mechanisms that make it possible to sustain that the *benchmarks* applied are genuinely useful, for academia as much as for industry. And at guidelines for producing and distributing the results of research projects (reproducible experiments, intermediate data, code, configurations) and, later on, an open repository of containers.

## Origin

The initiative comes from Miguel Hernández-Cabronero, of DEIC (Departament d'Enginyeria de la Informació i de les Comunicacions) at Universitat Autònoma de Barcelona.