Adopting FAIR principles for Rothamsted's long-term agricultural experiments

Richard Ostler

ROTHAMSTED RESEARCH

Biotechnology and Biological Sciences Research Council

Lawes Agricultural Trust

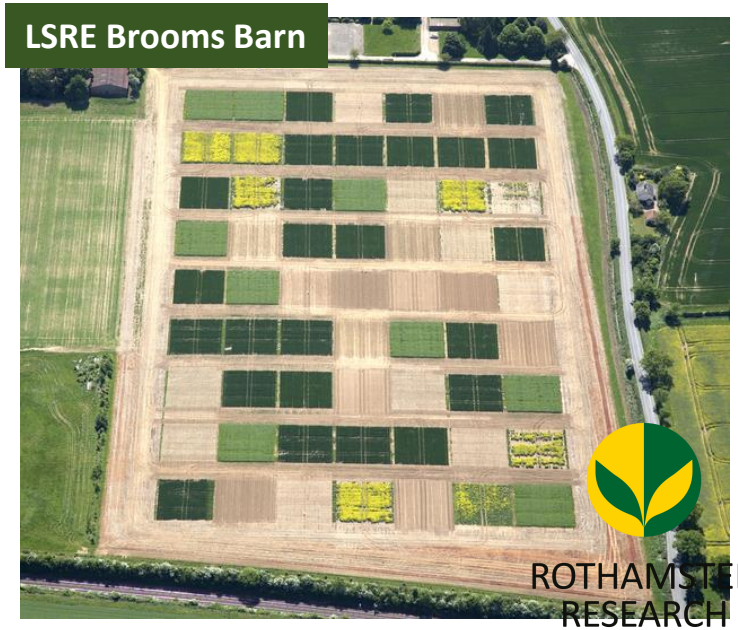
The Rothamsted Long-term Experiments

- 1843 - 1856 Lawes and Gilbert started the 9 “Classical Experiments” including Broadbalk and Park Grass
- > 40 LTEs additional LTEs, past and present, running for at least 20 years
- Examine multiple treatment factors (fertilization, cropping systems, crop protection, management…)
- 177 years of data have (yield traits, disease, botanical and invertebrate surveys, soil properties, management…)
- Funded as a BBSRC National Capability with a requirement to make data as accessible as possible
- LTE information and datasets available from e-RA website http://era.rothamsted.ac.uk, launched 2013

Broadbalk Wheat

Park Grass Hay

LSRE Brooms Barn



ROTHAMSTED RESEARCH

What does FAIR mean

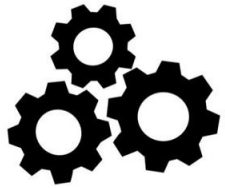
Findable

- Can people find your data?

Accessible

- If people can find your data do they know how it can be accessed
- FAIR ≠ Open Access

Interoperable



- If people can access your data can they integrate it with other related datasets

Reusable

- If people can find and access your data, is it sufficiently well described to allow them to re-use it?

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018 (2016). https://doi.org/10.1038/sdata.2016.18

https://www.go-fair.org/fair-principles/

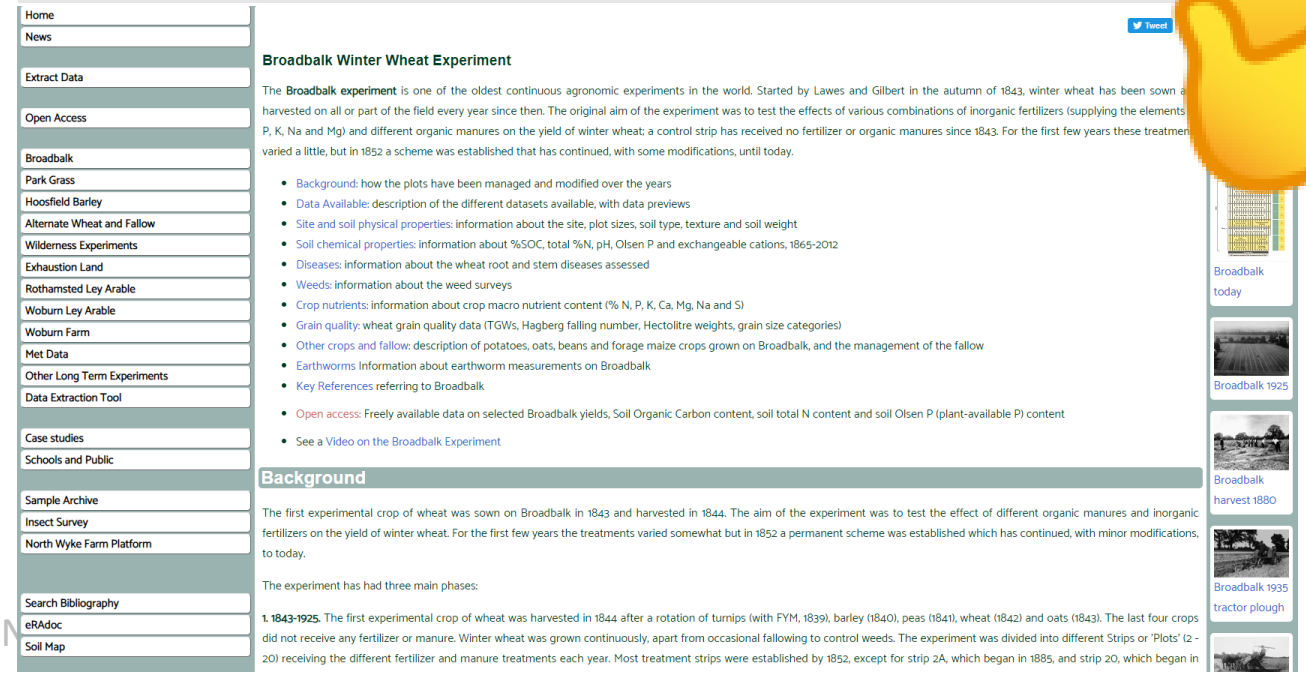
ROTHAMSTED RESEARCH

We do some of that – Does that mean we're FAIR?

Restricted access to annual plot data

plot	section	year	totalgrain85	totalstraw85	dmgrain	dmstraw	area	prev_crop	yr_of_wheat
			tonnes/hectare	tonnes/hectare	%	%	hectares		
10	0	2010	1.87		86.0		0.00320	F	6
11	0	2010	5.32		86.1		0.00320	F	6
12	0	2010	6.86		86.1		0.00320	F	6
13	0	2010	6.42		85.7		0.00320	F	6
14	0	2010	6.35		86.1		0.00320	F	6
15	0	2010	6.88		85.7		0.00320	F	6
16	0	2010	6.87		86.3		0.00320	F	6
17	0	2010	6.98		86.3		0.00320	F	6
18	0	2010	6.85		86.4		0.00320	F	6
19	0	2010	5.64		86.8		0.00320	F	6
21	0	2010	5.93		87.0		0.00320	F	6
22	0	2010	3.95		86.5		0.00320	F	6
3	0	2010	1.03		86.2		0.00320	F	6
5	0	2010	1.33		86.2		0.00320	F	6
6	0	2010	3.72		86.4		0.00320	F	6
7	0	2010	5.23		86.1		0.00320	F	6
8	0	2010	6.04		85.6		0.00320	F	6
9	0	2010	6.21		85.6		0.00320	F	6
10	0	2011	1.47		84.6		0.00320	F	7
11	0	2011	4.89		84.7		0.00320	F	7
12	0	2011	6.06		85.8		0.00320	F	7

Background information for Rothamsted Long-term experiments



Freely accessible aggregated datasets

	B	C	D	E	F	G	H	I	J	K	L
1	**Broadbalk Soil Total % Nitrogen (N) in topsoil (0-23cm) for "continuous" wheat sections**[1]										
2	Analysis of air-dried finely ground (0.354mm) soil. See notes below for analytical details.										
3	Data are the means of all holes (samples) in each strip (up to 1944) and of all sections in 1966; and then of										
4	specific sections from 1966 (see notes below).										
5	Data in red are estimated "starting" values (see notes below).										
6											
7	**% Soil Total N 0-23cm**										
8						Treatment					
9		Nil	PKMg	N3PKMg	N4PKMg	FYM	FYM1885	FYM1968			
10		Plot 3	Plot 5	Plot 8	Plot 9	Plot 2.2	Plot 2.1	Plot 1[5]			
11	1843	0.110	0.110	0.110	0.110	0.110	ns	ns			
12	1865	0.105	0.107	ns	ns	0.178	ns	ns			
13	1881	0.100	0.103	0.125	0.118	0.187	ns	ns			
14	1884	ns	ns	ns	ns	ns	0.100	ns			
15	1893	0.091	0.098	0.115	0.116	0.218	0.159	ns			
16	1914	0.094	0.104	0.130	0.114	0.256	0.202	ns			
17	1936	0.097	0.099	ns	ns	0.220	0.190	ns			
18	1944	0.102	0.103	0.120	0.111	0.229	0.198	ns			
19	1966 (all)	0.101	0.107	0.120	0.111	0.253	0.218	ns			
20	1966[2]	0.103	0.109	0.121	0.114	0.255	0.217	0.094			
21	1987[3]	0.102	0.104	0.126	0.122	0.270	0.228	ns			
22	1992[4]	0.091	0.101	0.118	0.122	0.271	0.271	0.203			
23	1997[4]	0.090	0.098	0.119	0.124	0.304	0.272	0.212			

FAIR Principles

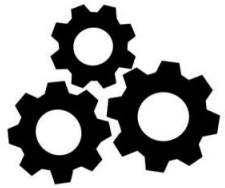
Findable

- (Meta)data assign unique persistent identifiers
- Described with rich metadata
- Metadata are indexed and searchable

Accessible

- Can be retrieved using a unique identifier (e.g. a DOI)

Interoperable



- (Meta)data use formal, shared languages for knowledge representation (i.e. data standards, controlled vocabularies, ontologies)
- (Meta)data has qualified references to related datasets

Reusable

- Metadata extended to include other relevant attributes (e.g. Descriptions of treatment factors)
- Published with data licences and access rights
- Dataset provenance included

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018 (2016). https://doi.org/10.1038/sdata.2016.18

https://www.go-fair.org/fair-principles/

ROTHAMSTED RESEARCH

Hmm, e-RA isn't FAIR…

How do we become FAIR

Datasets:

- ✓ Conditions of access for datasets
- ✗ No DOI (except for some aggregated datasets)
- ✗ No structured, machine readable metadata and metadata not in context
- ✗ Appropriate (semantic) annotation of data and fields
- ✗ Datasets poorly structured and data fields poorly described, no community standards adopted
- ✗ Non-standard formats

e-RA Website:

- ✓ Lots of content
- ✗ but presented as semi-structured narratives and inconsistent across experiments

ROTHAMSTED RESEARCH

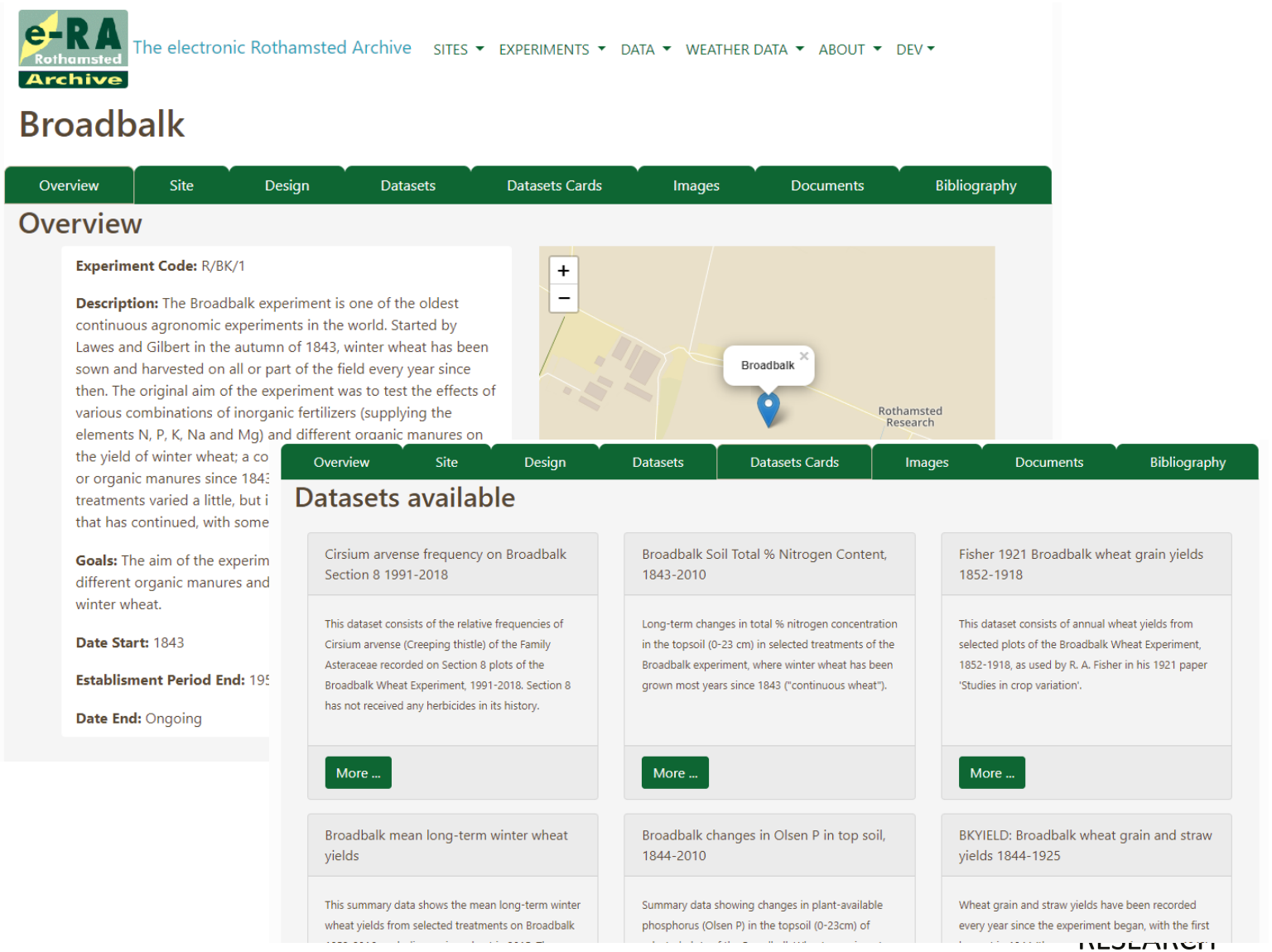
Improving LTE Metadata

GL TEN

+

schema.org

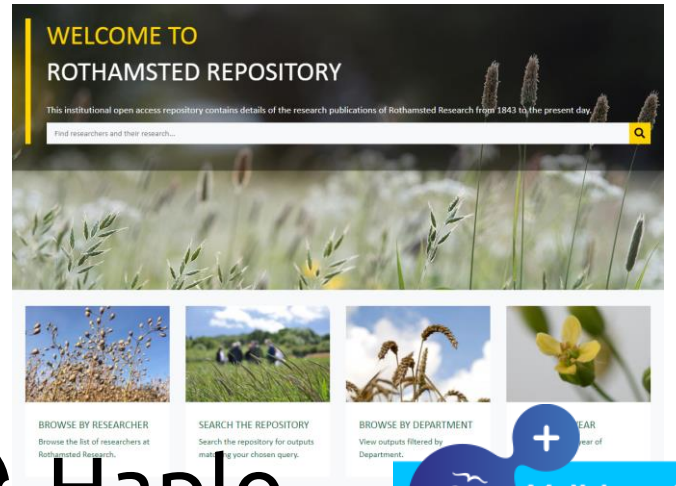
Schemas adopted for structured Long-term Experiment metadata



Publishing Findable, Accessible and Re-usable Datasets

Structured DataCite metadata

Harvested into Rothamsted's Institutional Repository



Dataset: Broadbalk changes in Olsen P in top soil, 1844-2010

DOI: 10.23637/keyrefoabkolsenp

Experiment: Broadbalk

Files:
- Olsen P data and treatments
- Open Access BK Olsen P

Version: 1

Creation Date: 2016-08-22

Publication Date: N/A

Last updated: 2016-08-22

Keywords: long term experiments - Broadbalk long-term experiment - soil - phosphorous - fertilizer - farmyard manure

Authors: Margaret Glendining, Sarah Perryman

Publisher: Rothamsted Research

Summary data showing changes in plant-available phosphorus (Olsen P) in the topsoil (0-23cm) of selected plots of the Broadbalk Wheat experiment, 1844-2010.

Methods

The figure shows changes in plant-available phosphorus (Olsen P) in the topsoil (0-23cm) of selected plots of the Broadbalk Wheat experiment. Olsen P, which is commonly used to assess plant-available P, is measured by extracting soil with a solution of 0.5M NaHCO3, buffered at pH 8.5. Until 2000, most of the Broadbalk plots had received superphosphate P (35 kg P ha-1) each year since 1852. Since 2001, applications of P have been withheld on some plots. The accumulation of available P was greatest on the FYM plot, reflecting the larger total P inputs in FYM (approx 45 kg P ha-1) compared to that applied as inorganic fertilizer or where no P was applied. The greater Olsen P concentrations on plot 5 (PKMg) compared to plot 8 (N3PKMg), despite their similar P inputs, were a result of the higher yields and greater P offtakes in the crop on plot 8, due to the applied N. Plot 18 (N4 1/2PKMg) received P at half the standard rate between 1852 and 2001, resulting in relatively low Olsen P concentrations. Since 2001 fertilizer P has been applied at the full rate on this plot, and there has been a corresponding increase in Olsen P. Olsen P concentrations were least on plots receiving no (or very little) P inputs in fertilizer (plot 3, given no fertilizer or manure). Since 2001, fertilizer P has not been applied to plots 5 and 8 (and others not shown), due to the high reserves of plant-available P in the soil. This is reviewed each year.

Related Documents

Access Use and Conditions

YOU MUST CITE AS: Rothamsted Research (N/A). Broadbalk changes in Olsen P in top soil, 1844-2010 *Electronic Rothamsted Archive* 10.23637/keyrefoabkolsenp

Rothamsted relies on the integrity of users to ensure that Rothamsted Research receives suitable acknowledgment as being the originators of these data. This enables us to monitor the use of each dataset and to demonstrate their value. Please send us a link to any publication that uses this Rothamsted data.

X Cancel ✓ Capture

Page has schema.org structured data (@type: Dataset)

Access conditions and CC-BY License

Haplo

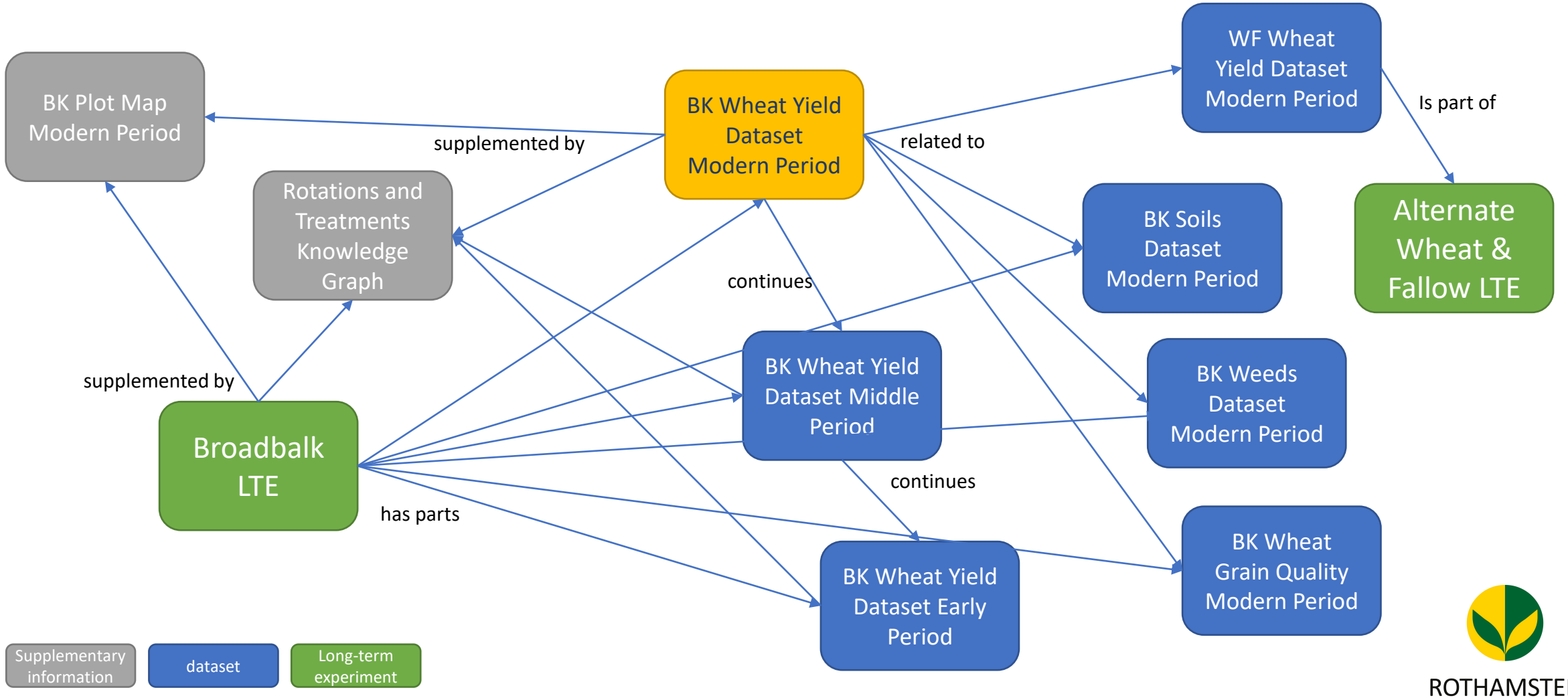
Validated


```
<script type="application/ld+json">{
    "@context": "https://schema.org/",
    "@type": "Dataset",
    "identifier": "10.23637/KeyRefOABKyields",
    "name": "Broadbalk mean long-term winter wheat yields",
    "shortName": "OAWWYields",
    "url": "http://www.era.rothamsted.ac.uk/Broadbalk/bbk_open_access_yields",
    "description": [
        {
            "inLanguage": "en",
            "descriptionType": "Abstract",
            "description": "This summary data shows the mean long-term winter wheat yields from selected treatments on Broadbalk 1852-2016, exclud
agronomic practices introduced on Broadbalk such as modern cultivars, better control of pests, diseases and weeds, especially since the 1960s."
        },
        {
            "inLanguage": "en",
            "descriptionType": "Methods",
            "description": "To control weeds, the experiment was divided into five sections in the 1920s and one section bare fallowed each year.
manure have remained at around 1 t ha-1. In 1968 a rotation was introduced on part of the experiment, so that it is now possible to compare the yi
been used, which has allowed us to exploit the greater grain yield potential of modern cultivars. From 1985, two higher N rates have been tested,
from fertilizer alone exceeding those from FYM alone, and the combination of FYM + 96 kgN ha-1 (144 kgN ha-1 since 2005) often exceeding both. The
Hereward to Crusoe. The greatest yields were not always achieved with the highest N rate. The figure shows the mean greatest first wheat yields ac
1984)."
        }
    ],
    "publisher": {
        "type": "organization",
        "name": "Rothamsted Research"
    },
    "datePublished": null,
    "dateCreated": "2020-01-17",
    "dateModified": null,
    "inLanguage": "en",
    "version": "1",
    "keywords": [
        "Broadbalk long-term experiment",
        "long term experiments",
        "crop yield",
        "wheat",
        "fertilizer",
        "farmyard manure",
        "cropping system"
    ],
    "creator": [
        {
            "type": "organization",
```


ROTHAMSTED RESEARCH

DOI Knowledge Graphs

Experiments, Datasets and Supplementary information are published in context.



ROTHAMSTED RESEARCH

Refactoring the data

- Simple tabular CSV format
- Relationships between data tables possible

Positional data, identifies plots in time and space

Treatment factors – what was applied to the plots

Observed and Measured data

sample_year	strip	old_section	section	soil_dep	cropping	straw_incorporation	herbicide	fungicide	n_rate	other_fertiliser_combination	soil_organic_carbon	soil_total_nitrogen	olsen_p	bulk_density	calculated_soc_content
1988	3	Ia	0	0-23	Continuous wheat	S1	H1	F1	N0	Nil	0.93	0.105	6	1.252	26.78
1987	3	Ib	1	0-23	Continuous wheat	S0	H1	F1	N0	Nil	0.98	0.108	4	1.252	28.22
1988	3	II	2	0-23	Wheat in rotation	S0	H1	F1	N0	Nil	0.88	0.088	8	1.252	25.34
1988	3	II	3	0-23	Wheat in rotation	S0	H1	F1	N0	Nil	0.77	0.078	9	1.252	22.17
1987	3	III	4	0-23	Wheat in rotation	S0	H1	F1	N0	Nil	0.79	0.092	9	1.252	22.75
1988	3	III	5	0-23	Wheat in rotation	S0	H1	F1	N0	Nil	0.85	0.082	8	1.252	24.48
1987	3	IV	6	0-23	Continuous wheat	S0	H1	F0	N0	Nil	ns	ns	ns	1.252	ns
1987	3	IV	7	0-23	Wheat in rotation	S0	H1	F1	N0	Nil	0.83	0.090	6	1.252	23.90
1987	3	Va	8	0-23	Continuous wheat	S0	H0	F1	N0	Nil	ns	ns	ns	1.252	ns
1987	3	Vb	9	0-23	Continuous wheat	S0	H1	F1	N0	Nil	0.83	0.095	5	1.252	23.90
1988	5	Ia	0	0-23	Continuous wheat	S1	H1	F1	N0	PKMg	0.97	0.102	102	1.252	27.93
1987	5	Ib	1	0-23	Continuous wheat	S0	H1	F1	N0	PKMg	0.94	0.108	72	1.252	27.07
1988	5	II	2	0-23	Wheat in rotation	S0	H1	F1	N0	PKMg	0.77	0.085	86	1.252	22.17
1988	5	II	3	0-23	Wheat in rotation	S0	H1	F1	N0	PKMg	0.77	0.085	88	1.252	22.17
1987	5	III	4	0-23	Wheat in rotation	S0	H1	F1	N0	PKMg	0.77	0.092	70	1.252	22.17
1988	5	III	5	0-23	Wheat in rotation	S0	H1	F1	N0	PKMg	0.82	0.084	91	1.252	23.61
1987	5	IV	6	0-23	Continuous wheat	S0	H1	F0	N0	PKMg	ns	ns	ns	1.252	ns
1987	5	IV	7	0-23	Wheat in rotation	S0	H1	F1	N0	PKMg	0.87	0.095	67	1.252	25.05
1987	5	Va	8	0-23	Continuous wheat	S0	H0	F1	N0	PKMg	ns	ns	ns	1.252	ns
1987	5	Vb	9	0-23	Continuous wheat	S0	H1	F1	N0	PKMg	0.88	0.099	82	1.252	25.34
1988	6	Ia	0	0-23	Continuous wheat	S1	H1	F1	N1	PKMg	1.02	0.104	102	1.252	29.37
1987	6	Ib	1	0-23	Continuous wheat	S0	H1	F1	N1	PKMg	1.00	0.117	65	1.252	28.80
1988	6	II	2	0-23	Wheat in rotation	S0	H1	F1	N1	PKMg	0.82	0.091	91	1.252	23.61
1988	6	II	3	0-23	Wheat in rotation	S0	H1	F1	N1	PKMg	0.90	0.091	92	1.252	25.92
1987	6	III	4	0-23	Wheat in rotation	S0	H1	F1	N1	PKMg	0.79	0.095	80	1.252	22.75
1988	6	III	5	0-23	Wheat in rotation	S0	H1	F1	N1	PKMg	0.88	0.085	98	1.252	25.34
1987	6	IV	6	0-23	Continuous wheat	S0	H1	F0	N1	PKMg	ns	ns	ns	1.252	ns
1987	6	IV	7	0-23	Wheat in rotation	S0	H1	F1	N1	PKMg	0.90	0.098	85	1.252	25.92
1987	6	Va	8	0-23	Continuous wheat	S0	H0	F1	N1	PKMg	ns	ns	ns	1.252	ns
1987	6	Vb	9	0-23	Continuous wheat	S0	H1	F1	N1	PKMg	1.02	0.108	92	1.252	29.37

Rigorous metadata with semantic annotation

"Primary key" values

Human readable labels

Semantic annotations for data

name	factor_	unit	treatment	treatment_rdfType	treatment_form	treatment_form_rdfType	application_timing	description
N0	0		nitrogen	http://purl.obolibrary.org/obo/PECO_0007284				between 1968-85 applied as calcium ammonium nitrate (Nitro
N1	48	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	ammonium nitrate (Nitram, 34.5%N)	http://aims.fao.org/aos/agrovoc/c_352	spring	between 1968-85 applied as calcium ammonium nitrate (Nitro
N2	96	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	ammonium nitrate (Nitram, 34.5%N)	http://aims.fao.org/aos/agrovoc/c_352	spring	between 1968-85 applied as calcium ammonium nitrate (Nitro
N3	144	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	ammonium nitrate (Nitram, 34.5%N)	http://aims.fao.org/aos/agrovoc/c_352	spring	between 1968-85 applied as calcium ammonium nitrate (Nitro
N4	192	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	ammonium nitrate (Nitram, 34.5%N)	http://aims.fao.org/aos/agrovoc/c_352	spring	between 1968-85 applied as calcium ammonium nitrate (Nitro
N5	240	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	ammonium nitrate (Nitram, 34.5%N)	http://aims.fao.org/aos/agrovoc/c_352	spring	between 1968-85 applied as calcium ammonium nitrate (Nitro
N6	288	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	ammonium nitrate (Nitram, 34.5%N)	http://aims.fao.org/aos/agrovoc/c_352	spring	between 1968-85 applied as calcium ammonium nitrate (Nitro
N*1	48	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	sodium nitrate	http://aims.fao.org/aos/agrovoc/c_7148	spring	applied between 1852 and 1967 only
N*2	96	kgN ha^{-1} yr^{-1}	nitrogen	http://purl.obolibrary.org/obo/PECO_0007284	sodium nitrate	http://aims.fao.org/aos/agrovoc/c_7148	spring	applied between 1852 and 1967 only
P	35	kgP ha^{-1} yr^{-1}	phosphorous	http://purl.obolibrary.org/obo/PECO_0007397	triple superphosphate	http://aims.fao.org/aos/agrovoc/c_7521	autumn	35kgP as triple superphosphate until 2000
(P)	0		phosphorous	http://purl.obolibrary.org/obo/PECO_0007397				35kgP as triple superphosphate until 2000; no longer applied d
K	90	kgK ha^{-1} yr^{-1}	potassium	http://purl.obolibrary.org/obo/PECO_0007293	potassium sulphate	http://aims.fao.org/aos/agrovoc/c_6142	autumn	
Mg	12	kgMg ha^{-1} yr^{-1}	magnesium	http://purl.obolibrary.org/obo/PECO_0007288	Kieserite since 2001	http://purl.obolibrary.org/obo/CHEBI_32599	autumn	11kgMg as magnesium sulphate until 1973; 35kgMg every third
Na	16	kgNa ha^{-1} yr^{-1}	sodium		sodium sulphate	http://purl.obolibrary.org/obo/CHEBI_32149	autumn	16kgNa as sodium sulphate until 1973
(Na)	0		sodium					16kgNa as sodium sulphate until 1973, then no longer applied
FYM	35	t ha^{-1} yr^{-1}	FYM	http://purl.obolibrary.org/obo/PECO_0007087	farmyard manure	http://aims.fao.org/aos/agrovoc/c_2810	autumn	Farmyard manure from cattle 1968-2000 only
(FYM)	0		No FYM	http://purl.obolibrary.org/obo/PECO_0007087				
S0	0		wheat straw not incorporated		wheat straw not incorporated			straw not incorporated
S1	1		wheat straw incorporated		wheat straw incorporated	http://aims.fao.org/aos/agrovoc/c_7441	autumn	chopped wheat straw from previous crop incorporated since 1
H0	0		no herbicides	http://purl.obolibrary.org/obo/PECO_0007183	herbicides not applied			herbicides never applied
H1	1		herbicides	http://purl.obolibrary.org/obo/PECO_0007183	herbicides	http://aims.fao.org/aos/agrovoc/c_3566	spring and summer	herbicides applied as required since 1957
F0	0		no fungicides	http://purl.obolibrary.org/obo/PECO_0007268	spring and summer fungicides not applied			spring and summer fungicides not applied
F1	1		fungicides	http://purl.obolibrary.org/obo/PECO_0007268	spring and summer fungicides	http://aims.fao.org/aos/agrovoc/c_3566	spring and summer	spring and summer fungicides applied as required since 1979

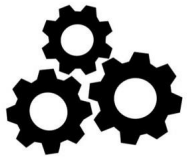
ROTHAMSTED RESEARCH

Do our changes measure up to FAIR?

Findable

- ✓ Persistent Identifier provided (DOI)
- ✓ Extensive Metadata provided
- ✓ Metadata registered in a searchable resource

Interoperable



- ✓ Non-proprietary, open format
- ✓ Metadata uses a formal, accessible and shared language
- ✓ Data annotated using standard vocabularies
- ✓ Data linked to other data to provide context

Accessible

- ✓ Metadata is retrievable by its identifier
- ✓ Metadata provides URLs to physical files
- ✓ Metadata available even when data are not

Reusable

- ✓ Metadata have plurality of relevant attributes
- ✓ Data are released with a clear and accessible data usage license (CC-BY)
- ✓ Metadata are associated with provenance
- ✓ Metadata meet domain relevant community standards

ROTHAMSTED RESEARCH

What have we learned from Going FAIR

1. You can’t change culture overnight
    - This is a new way of thinking about our LTEs and data both for our data managers and people using the data, be sensitive to their needs but change should make life easier!

2. Budgets are not unlimited – trade-off between ambitions and financial/human resources

3. Address skills gaps, invest in and train research data managers who understand the data
    - Develop and use best practices – Data Carpentries

ROTHAMSTED RESEARCH

But why Go FAIR in the first place?

Putting FAIR into practice means people can more easily discover and re-use the data.

Using DOIs means our data can be recognised and attributed.

Machine readable and interoperable datasets can be accessed and used in new ways with other datasets

1. Democratises access to a unique data collection
2. Data producers and curators are recognised
3. Supports reproducible science
4. Enables new (data) science approaches
5. Increases impact of the LTEs and their data
   - impact is more easily measured
   - demonstrates return on investment to funders

ROTHAMSTED RESEARCH

Thanks

Nathalie Castells-Brooke (LTE Computer Scientist)
Margaret Glendining (LTE Data Curator)
Andy Macdonald (LTE Project Leader)
Sarah Perryman (LTE Data Curator)

The Rothamsted Long-term Experiments National Capability is funded by UK BBSRC and Lawes Agricultural Trust