Automate office files to PDF

Microsoft, LibreOffice or other?

OctoberPDFest
ONLINE

Intro

- “Office Open XML File Format” (Microsoft Office)
  - Standardized as ISO/IEC 29500-1:2016
  - Document creation, little guidance for rendering

- PDF is an essential part of all current digital processes
  - Reliable, robust rendering
  - Archiving

  - How can you get from Office to PDF, preferably in an automated way?

The easy approach: Printer Driver

- Works for every application that has a printer interface
- Limitations
  - Unicode (not needed)
  - Tagging not possible
  - Metadata (not needed and not easily possible)
  - 2 conversions instead of one: To the printer language and from there to PDF
  - PostScript has a lot of limitations, transparency, page size,
    on an office printer that usually is not a problem, but in a digital file this becomes obvious

“Office Open XML Format” to PDF

- Most office files are created with Microsoft Office
  - Quality of conversion is defined by similarity of appearance to its results

- Automated conversion via Microsoft Office is possible
  (various solutions are around)
  - Requires
    - Handling of dialogues
    - Stabilize the application if running for hours or days
  - Bound to the Windows platform

Alternatives to Microsoft Office

- LibreOffice (Open Office)
- Standalone converters or libraries
  various solutions available

- Rendering is not in the core of the “Office Open XML File Format” standard, no alternative will exactly match Microsoft Office results
- But LibreOffice has gotten much closer in recent versions…

At first glance: Office files vs PDF

- **Text processors: Word, Writer**
  - Text reflow on the fly
- **Spreadsheet: Excel, Calc**
  - No pages at all
  - Formulas
- **Slides: PowerPoint, Impress**
  - Close to PDF
- **Email**
  - Threads
  - External content
  - Header, Body, Attachments
- **Other**
  - Visio, Project…

Text processors: Reflow on the fly

- PDF: (mostly) places each object independent from other objects
- This is not the case in a re-flowable format (Word / Writer, also HTML)

- PDF is faster at least for docs with various pages
  - PDF: Direct access to each page (independent from previous pages)
  - Office: Needs to render all previous pages

Test case: long texts

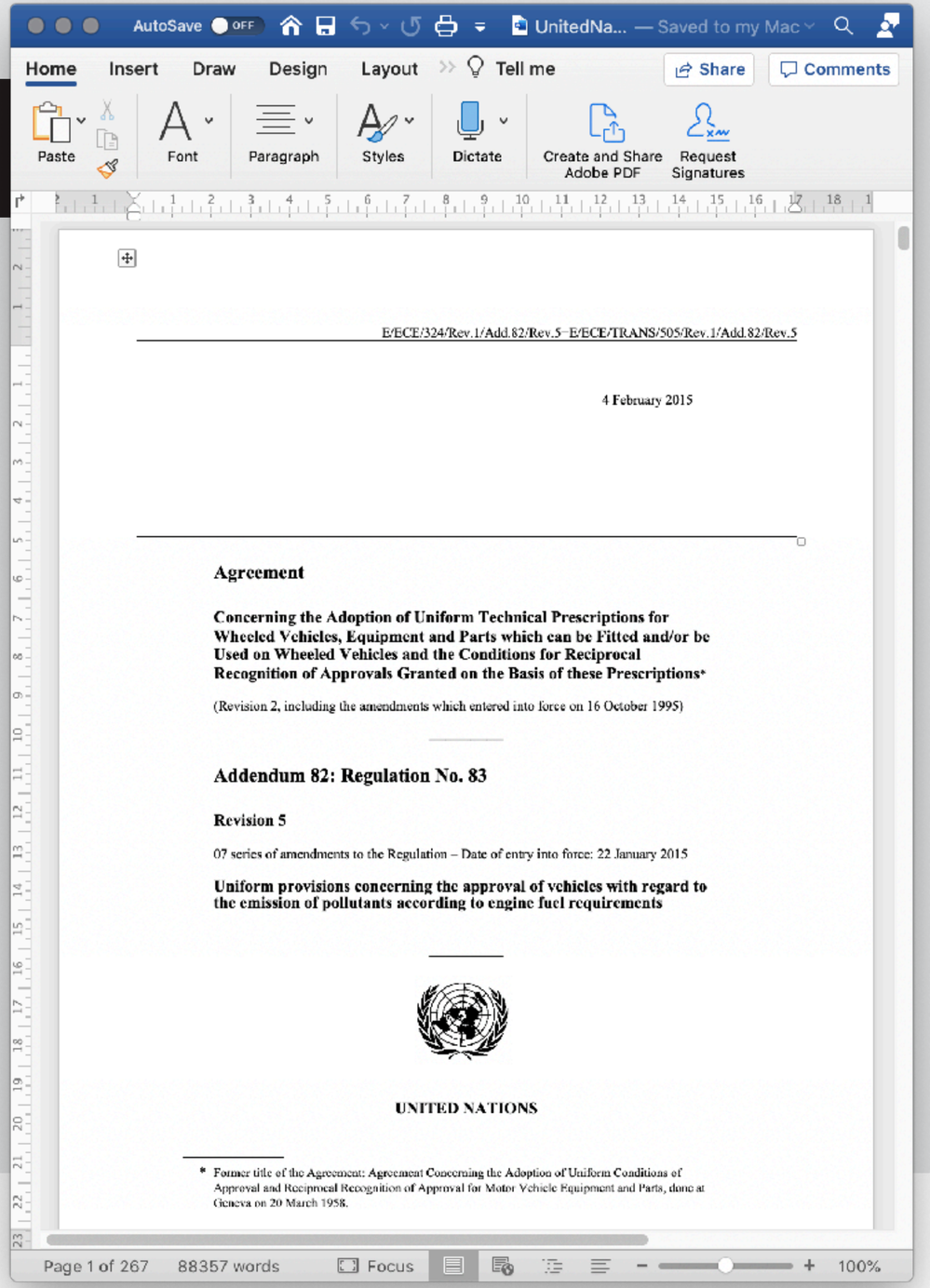
Longer texts from the same (MS Word) file

Slightly different text reflow results in:

MS Word 267 pages

LibreOffice 275 pages

Examples…



PDF association

Libre Office

from the dates given in paragraphs 12.2.3. and 12.2.4. of this Regulation for new type approvals and new vehicles respectively are given in Table A11/1:

Table A11/1: Final OBD threshold limits

		Reference mass (RM) (kg)	*Mass of carbon monoxide (CO) (mg/km)*		*Mass of non-methane hydrocarbons (NMHC) (mg/km)*		*Mass of oxides of nitrogen (NO_x) (mg/km)*		*Mass of particulate matter[1] (PM) (mg/km)*		*Number of particulates[1] (PN) (#/km)*	
Category	Class		PI	CI	PI	CI	PI	CI	CI	PI	CI	PI
M	—	All	1,900	1,750	170	290	90	140	12	12		
N_1	I	RM ≤ 1305	1,900	1,750	170	290	90	140	12	12		
	II	1305 < RM ≤ 1760	3,400	2,200	225	320	110	180	12	12		
	III	1760 < RM	4,300	2,500	270	350	120	220	12	12		
N_2	—	All	4,300	2,500	270	350	120	220	12	12		

Key PI Positive Ignition
CI Compression Ignition.

[1] Positive ignition particulate mass and number limits apply only to vehicles with direct injection engines

3.3.2.2. Until the dates specified in paragraphs 12.2.3. and 12.2.4. of this Regulation for new type approvals and new vehicles respectively, the OBD threshold limits in Table A11/2 shall be applied to vehicles that are type approved according to the emission limits set out in Table 1 in paragraph 5.3.1.4. of this Regulation, upon the choice of the manufacturer:

Table A11/2: Preliminary OBD threshold limits

		Reference mass (RM) (kg)	*Mass of carbon monoxide (CO) (mg/km)*		*Mass of non-methane hydrocarbons (NMHC) (mg/km)*		*Mass of oxides of nitrogen (NO_x) (mg/km)*		*Mass of particulate matter[1] (PM) (mg/km)*	
Category	Class		PI	CI	PI	CI	PI	CI	CI	PI
M	—	All	1,900	1,750	170	290	150	180	25	25
N_1	I	RM ≤ 1305	1,900	1,750	170	290	150	180	25	25
	II	1305 < RM ≤ 1760	3,400	2,200	225	320	190	220	25	25
	III	1760 < RM	4,300	2,500	270	350	210	280	30	30
N_2	—	All	4,300	2,500	270	350	210	280	30	30

Key PI Positive Ignition
CI Compression Ignition.

[1] Positive ignition particulate mass and number limits apply only to vehicles with direct injection engines

3.3.2.3. The OBD thresholds limits for compression ignition vehicles that comply with the emission limit values set out Table 1 in paragraph 5.3.1.4. of this Regulation and type-approved before the dates given in paragraph 12.2.1. of this Regulation are contained in the Table A11/3. These threshold limits shall cease to apply from the dates set out in paragraph 12.2.2. of this Regulation for new vehicles to be registered, sold or entered into service.

MS Office

Table A11/1: Final OBD threshold limits

		Reference mass (RM) (kg)	*Mass of carbon monoxide (CO) (mg/km)*		*Mass of non-methane hydrocarbons (NMHC) (mg/km)*		*Mass of oxides of nitrogen (NO_x) (mg/km)*		*Mass of particulate matter[1] (PM) (mg/km)*		*Number of particulates[1] (PN) (#/km)*	
Category	Class		PI	CI	PI	CI	PI	CI	CI	PI	CI	PI
M	—	All	1,900	1,750	170	290	90	140	12	12		
N_1	I	RM ≤ 1305	1,900	1,750	170	290	90	140	12	12		
	II	1305 < RM ≤ 1760	3,400	2,200	225	320	110	180	12	12		
	III	1760 < RM	4,300	2,500	270	350	120	220	12	12		
N_2	—	All	4,300	2,500	270	350	120	220	12	12		

Key PI Positive Ignition
CI Compression Ignition.

[1] Positive ignition particulate mass and number limits apply only to vehicles with direct injection engines

3.3.2.2. Until the dates specified in paragraphs 12.2.3. and 12.2.4. of this Regulation for new type approvals and new vehicles respectively, the OBD threshold limits in Table A11/2 shall be applied to vehicles that are type approved according to the emission limits set out in Table 1 in paragraph 5.3.1.4. of this Regulation, upon the choice of the manufacturer:

Table A11/2: Preliminary OBD threshold limits

		Reference mass (RM) (kg)	*Mass of carbon monoxide (CO) (mg/km)*		*Mass of non-methane hydrocarbons (NMHC) (mg/km)*		*Mass of oxides of nitrogen (NO_x) (mg/km)*		*Mass of particulate matter[1] (PM) (mg/km)*	
Category	Class		PI	CI	PI	CI	PI	CI	CI	PI
M	—	All	1,900	1,750	170	290	150	180	25	25
N_1	I	RM ≤ 1305	1,900	1,750	170	290	150	180	25	25
	II	1305 < RM ≤ 1760	3,400	2,200	225	320	190	220	25	25
	III	1760 < RM	4,300	2,500	270	350	210	280	30	30
N_2	—	All	4,300	2,500	270	350	210	280	30	30

Key PI Positive Ignition
CI Compression Ignition.

[1] Positive ignition particulate mass and number limits apply only to vehicles with direct injection engines

3.3.2.3. The OBD thresholds limits for compression ignition vehicles that comply with the emission limit values set out Table 1 in paragraph 5.3.1.4. of this Regulation and type-approved before the dates given in paragraph 12.2.1. of this Regulation are contained in the Table A11/3. These threshold limits shall cease to apply from the dates set out in paragraph 12.2.2. of this Regulation for new vehicles to be registered, sold or entered into service.

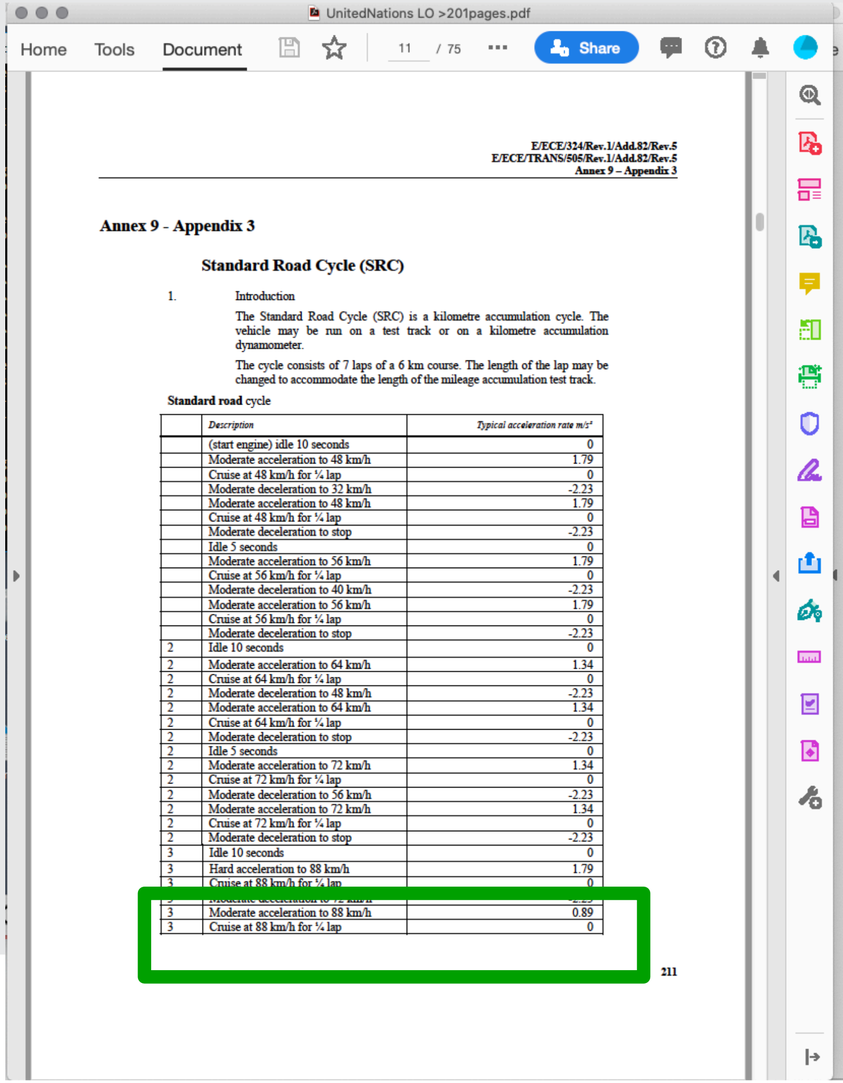
Table A11/3: Interim OBD threshold limits

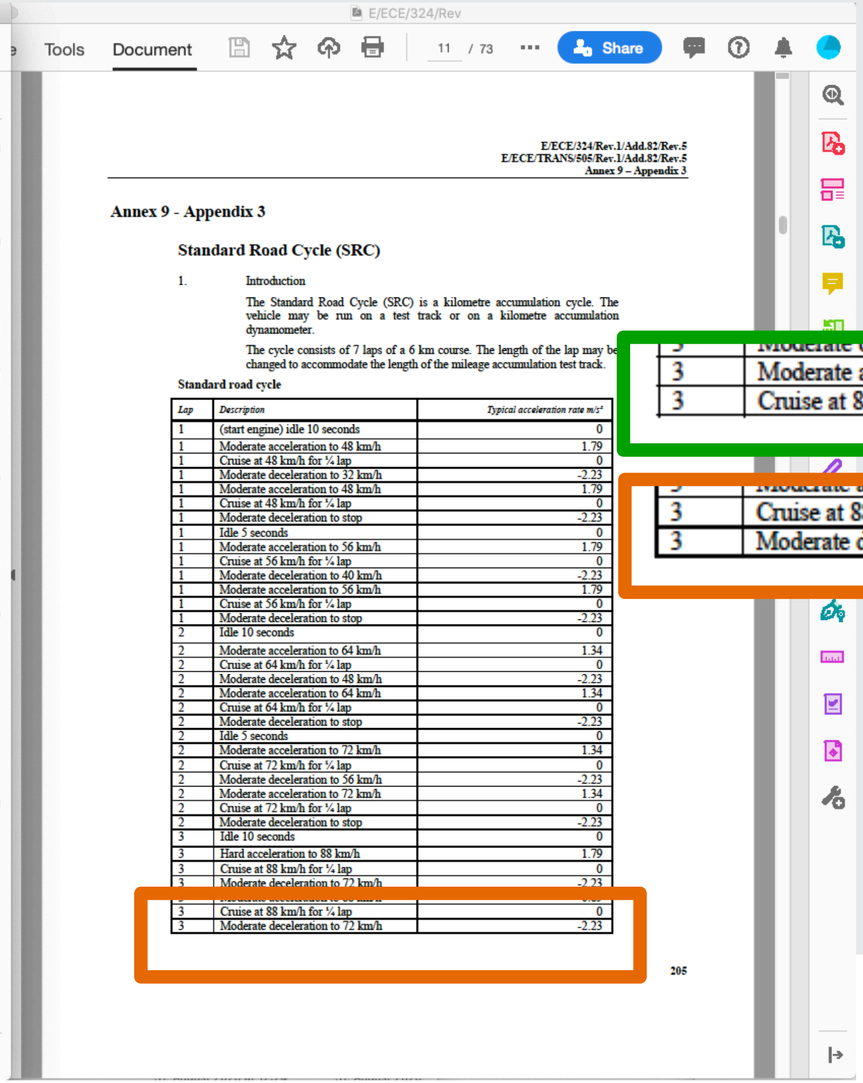
		Reference mass (RM) (kg)	*Mass of carbon monoxide (CO) (mg/km)*	*Mass of non-methane hydrocarbons (NMHC) (mg/km)*	*Mass of oxides of nitrogen (NO_x) (mg/km)*	*Mass of particulate matter (PM) (mg/km)*
Category	Class		CI	CI	CI	CI
M	—	All	1,900	320	240	50
N_1	I	RM ≤ 1305	1,900	320	240	50
	II	1305 < RM ≤ 1760	2,400	360	315	50
	III	1760 < RM	2,800	400	375	50
N_2	—	All	2,800	400	375	50

Key PI Positive Ignition
CI Compression Ignition.

Reflowable text and other objects

OctoberPDFest ONLINE





Libre Office

3	Moderate acceleration to 88 km/h	0.89
3	Cruise at 88 km/h for ¼ lap	0

MS Office
one more line
on the same page

3	Cruise at 88 km/h for ¼ lap	0
3	Moderate deceleration to 72 km/h	-2.23

Type: NG/biomethane

Characteristics	Units	Basis	Limits min.	Limits max.	Test Method
Reference fuel G_{20}					
Composition:					
Methane	per cent mole	100	99	100	ISO 6974
Balance[1]	per cent mole	-	-	1	ISO 6974
N_2	per cent mole				ISO 6974
Sulphur content	mg/m³ [2]	-	-	10	ISO 6326-5
Wobbe Index (net)	MJ/m³ [3]	48.2	47.2	49.2	
Reference fuel G_{25}					
Composition:					
Methane	per cent mole	86	84	88	ISO 6974
Balance [1]	per cent mole	-	-	1	ISO 6974
N_2	per cent mole	14	12	16	ISO 6974
Sulphur content	mg/m³ [2]	-	-	10	ISO 6326-5
Wobbe Index (net)	MJ/m³ [3]	39.4	38.2	40.6	

[1] Inerts (different from N_2) + C_2 + C_{2+}.
[2] Value to be determined at 293.2 K (20 °C) and 101.3 kPa.

222

Type: NG/biomethane

Characteristics	Units	Basis	Limits min.	Limits max.	Test Method
Reference fuel G_{20}					
Composition:					
Methane	per cent mole	100	99	100	ISO 6974
Balance[1]	per cent mole	-	-	1	ISO 6974
N_2	per cent mole				ISO 6974
Sulphur content	mg/m³ [2]	-	-	10	ISO 6326-5
Wobbe Index (net)	MJ/m³ [3]	48.2	47.2	49.2	
Reference fuel G_{25}					
Composition:					
Methane	per cent mole	86	84	88	ISO 6974
Balance [1]	per cent mole	-	-	1	ISO 6974
N_2	per cent mole	14	12	16	ISO 6974
Sulphur content	mg/m³ [2]	-	-	10	ISO 6326-5
Wobbe Index (net)	MJ/m³ [3]	39.4	38.2	40.6	

[1] Inerts (different from N_2) + C_3 + C_{2+}.
[2] Value to be determined at 293.2 K (20 °C) and 101.3 kPa.
[3] Value to be determined at 273.2 K (0 °C) and 101.3 kPa.

UnitedNations LO >201pages.pdf

Home Tools Document 23 / 75 Share

Libre Office

E/ECE/324/Rev.1/Add.82/Rev.5
E/ECE/TRANS/505/Rev.1/Add.82/Rev.5
Annex 10a

Characteristics	Units	Basis	Limits min.	Limits max.	Test Method

[3] Value to be determined at 273.2 K (0 °C) and 101.3 kPa.

1.3. Technical data of hydrogen for internal combustion engines

Type: hydrogen for internal combustion engines

Characteristics	Units	Limits minimum	Limits maximum	Test method
Hydrogen purity	% mole	98	100	ISO 14687-1
Total hydrocarbon	µmol/mol	0	100	ISO 14687-1
Water[1]	µmol/mol	0	[2]	ISO 14687-1
Oxygen	µmol/mol	0	[2]	ISO 14687-1
Argon	µmol/mol	0	[2]	ISO 14687-1
Nitrogen	µmol/mol	0	[2]	ISO 14687-1
CO	µmol/mol	0	1	ISO 14687-1
Sulphur	µmol/mol	0	2	ISO 14687-1
Permanent particulates[3]				ISO 14687-1

[1] Not to be condensed
[2] Combined water, oxygen, nitrogen and argon: 1.900 µmol/mol.
[3] The hydrogen shall not contain dust, sand, dirt, gums, oils, or other substances in an amount sufficient to damage the fuelling station equipment of the vehicle (engine) being fuelled.

1.4. Technical data of hydrogen for fuel cell vehicles

Type: Hydrogen for fuel cell vehicles

Characteristics	Units	Limits	Test method

E/ECE/324/Rev

Tools Document 21 / 73 Share

MS Office

E/ECE/324/Rev.1/Add.82/Rev.5
E/ECE/TRANS/505/Rev.1/Add.82/Rev.5
Annex 10

Type: Ethanol (E75)

Parameter	Unit	Limits [1] Minimum	Limits [1] Maximum	Test method [2]
Research octane number, RON		95	-	EN ISO 5164
Motor octane number, MON		85	-	EN ISO 5163
Density at 15 °C	kg/m³	report		EN ISO 12185
Vapour pressure	kPa	50	60	EN ISO 1 30 16-1 (DVPE)
Sulphur content [3, 4]	mg/kg	-	10	EN ISO 20846 EN ISO 20884
Oxidation stability	minutes	360	-	EN ISO 7536
Existent gum content (solvent washed)	mg/100ml	-	4	EN ISO 6246
Appearance shall be determined at ambient temperature or 15 °C whichever is higher.		Clear and bright, visibly free of suspended or precipitated contaminants		Visual inspection
Ethanol and higher alcohols [7]	% (V/V)	70	80	EN 1601 EN 13132 EN 1451 7
Higher alcohols (C_3 - C_8)	% (V/V)	-	2	
Methanol		-	0.5	
Petrol [5]	% (V/V)	Balance		EN 228
Phosphorus	mg/l	0.3[6]		EN 15487 ASTM D 3231
Water content	% (V/V)	-	0.3	ASTM E 1064 EN 15 489

Simple formula

OctoberPDFest ONLINE

UnitedNations LO >201pages.pdf

Home Tools Document 25 / 75 Share

Libre Office

2.11. A "*warm-up cycle*" means sufficient vehicle operation such that the coolant temperature has risen by a least 22 K from engine starting and reaches a minimum temperature of 343 K (70 °C).

2.12. A "*Fuel trim*" refers to feedback adjustments to the base fuel schedule. Short-term fuel trim refers to dynamic or instantaneous adjustments. Long-term fuel trim refers to much more gradual adjustments to the fuel calibration schedule than short-term trim adjustments. These long-term adjustments compensate for vehicle differences and gradual changes that occur over time.

2.13. A "*Calculated load value*" refers to an indication of the current airflow divided by peak airflow, where peak airflow is corrected for altitude, if available. This definition provides a dimensionless number that is not engine specific and provides the service technician with an indication of the proportion of engine capacity that is being used (with wide open throttle as 100 per cent);

$$CLV = \frac{\text{Current airflow}}{\text{Peak airflow (at sea level)}} \bullet \frac{\text{Atmospheric pressure (at sea level)}}{\text{Barometric pressure}}$$

2.14. "*Permanent emission default mode*" refers to a case where the engine management controller permanently switches to a setting that does not require an input from a failed component or system where such a failed component or system would result in an increase in emissions from the vehicle to a level above the limits given in paragraph 3.3.2. of this annex.

2.15. "*Power take-off unit*" means an engine-driven output provision for the purposes of powering auxiliary, vehicle mounted, equipment.

2.16. "*Access*" means the availability of all emission-related OBD data including all fault codes required for the inspection, diagnosis, servicing or repair of emissions-related parts of the vehicle, via the serial interface for the standard diagnostic connection (pursuant to paragraph 6.5.3.5. of Appendix 1 to this annex).

2.17. "*Unrestricted*" means:

2.17.1. Access not dependent on an access code obtainable only from the manufacturer, or a similar device; or

2.17.2. Access allowing evaluation of the data produced without the need for any unique decoding information, unless that information itself is standardised.

E/ECE/324/Rev

Home Tools Document 25 / 73 Share

MS Office

2.11. A "*warm-up cycle*" means sufficient vehicle operation such that the coolant temperature has risen by a least 22 K from engine starting and reaches a minimum temperature of 343 K (70 °C).

2.12. A "*Fuel trim*" refers to feedback adjustments to the base fuel schedule. Short-term fuel trim refers to dynamic or instantaneous adjustments. Long-term fuel trim refers to much more gradual adjustments to the fuel calibration schedule than short-term trim adjustments. These long-term adjustments compensate for vehicle differences and gradual changes that occur over time.

2.13. A "*Calculated load value*" refers to an indication of the current airflow divided by peak airflow, where peak airflow is corrected for altitude, if available. This definition provides a dimensionless number that is not engine specific and provides the service technician with an indication of the proportion of engine capacity that is being used (with wide open throttle as 100 per cent);

$$CLV = \frac{\text{Current airflow}}{\text{Peak airflow (at sea level)}} \bullet \frac{\text{Atmospheric pressure (at sea level)}}{\text{Barometric pressure}}$$

2.14. "*Permanent emission default mode*" refers to a case where the engine management controller permanently switches to a setting that does not require an input from a failed component or system where such a failed component or system would result in an increase in emissions from the vehicle to a level above the limits given in paragraph 3.3.2. of this annex.

2.15. "*Power take-off unit*" means an engine-driven output provision for the purposes of powering auxiliary, vehicle mounted, equipment.

2.16. "*Access*" means the availability of all emission-related OBD data including all fault codes required for the inspection, diagnosis, servicing or repair of emissions-related parts of the vehicle, via the serial interface for the standard diagnostic connection (pursuant to paragraph 6.5.3.5. of Appendix 1 to this annex).

2.17. "*Unrestricted*" means:

2.17.1. Access not dependent on an access code obtainable only from the manufacturer, or a similar device; or

2.17.2. Access allowing evaluation of the data produced without the need for any unique decoding information, unless that information itself is standardised.

Libre Office

Annex 14

Emissions test procedure for Hybrid Electric Vehicles (HEV)

1. Introduction

1.1. This annex defines the specific provisions regarding type-approval of a Hybrid Electric Vehicle (HEV) as defined in paragraph 2.21.2. of this Regulation.

1.2. As a general principle, for the tests of Types I, II, III, IV, V, VI and OBD, hybrid electric vehicles shall be tested according to Annexes 4a, 5, 6, 7, 9, 8 and 11 to this Regulation respectively, unless modified by this annex.

1.3. For the Type I test only, OVC vehicles (as categorized in paragraph 2. of this annex) shall be tested according to condition A and to condition B. The test results under both conditions A and B and the weighted values shall be reported in the communication form.

1.4. The emissions test results shall comply with the limits under all specified test conditions of this Regulation.

2. Categories of hybrid electric vehicles

Vehicle charging	*Off-vehicle charging*[1] *(OVC)*		*Not off-vehicle charging*[2] *(NOVC)*	
Operating mode switch	Without	With	Without	With

[1] Also known as "externally chargeable"
[2] Also known as "not externally chargeable"

3. Type I test methods

3.1. Externally chargeable (OVC HEV) without an operating mode switch

3.1.1. Two tests shall be performed under the following conditions:

Condition A: Test shall be carried out with a fully charged electrical energy/power storage device.

Condition B: Test shall be carried out with an electrical energy/power storage device in minimum state of charge (maximum discharge of capacity).

The profile of the State Of Charge (SOC) of the electrical energy/power storage device during different stages of the Type I test is given in Appendix 1 to this annex.

3.1.2. Condition A

3.1.2.1. The procedure shall start with the discharge of the electrical energy/power storage device of the vehicle while driving (on the test track, on a chassis dynamometer, etc.):

(a) At a steady speed of 50 km/h until the fuel consuming engine of the HEV starts up;

MS Office

Annex 14

Emissions test procedure for Hybrid Electric Vehicles (HEV)

1. Introduction

1.1. This annex defines the specific provisions regarding type-approval of a Hybrid Electric Vehicle (HEV) as defined in paragraph 2.21.2. of this Regulation.

1.2. As a general principle, for the tests of Types I, II, III, IV, V, VI and OBD, hybrid electric vehicles shall be tested according to Annexes 4a, 5, 6, 7, 9, 8 and 11 to this Regulation respectively, unless modified by this annex.

1.3. For the Type I test only, OVC vehicles (as categorized in paragraph 2. of this annex) shall be tested according to condition A and to condition B. The test results under both conditions A and B and the weighted values shall be reported in the communication form.

1.4. The emissions test results shall comply with the limits under all specified test conditions of this Regulation.

2. Categories of hybrid electric vehicles

Vehicle charging	*Off-vehicle charging*[1] *(OVC)*		*Not off-vehicle charging*[2] *(NOVC)*	
Operating mode switch	Without	With	Without	With

[1] Also known as "externally chargeable"
[2] Also known as "not externally chargeable"

3. Type I test methods

3.1. Externally chargeable (OVC HEV) without an operating mode switch

3.1.1. Two tests shall be performed under the following conditions:

Condition A: Test shall be carried out with a fully charged electrical energy/power storage device.

Condition B: Test shall be carried out with an electrical energy/power storage device in minimum state of charge (maximum discharge of capacity).

The profile of the State Of Charge (SOC) of the electrical energy/power storage device during different stages of the Type I test is given in Appendix 1 to this annex.

3.1.2. Condition A

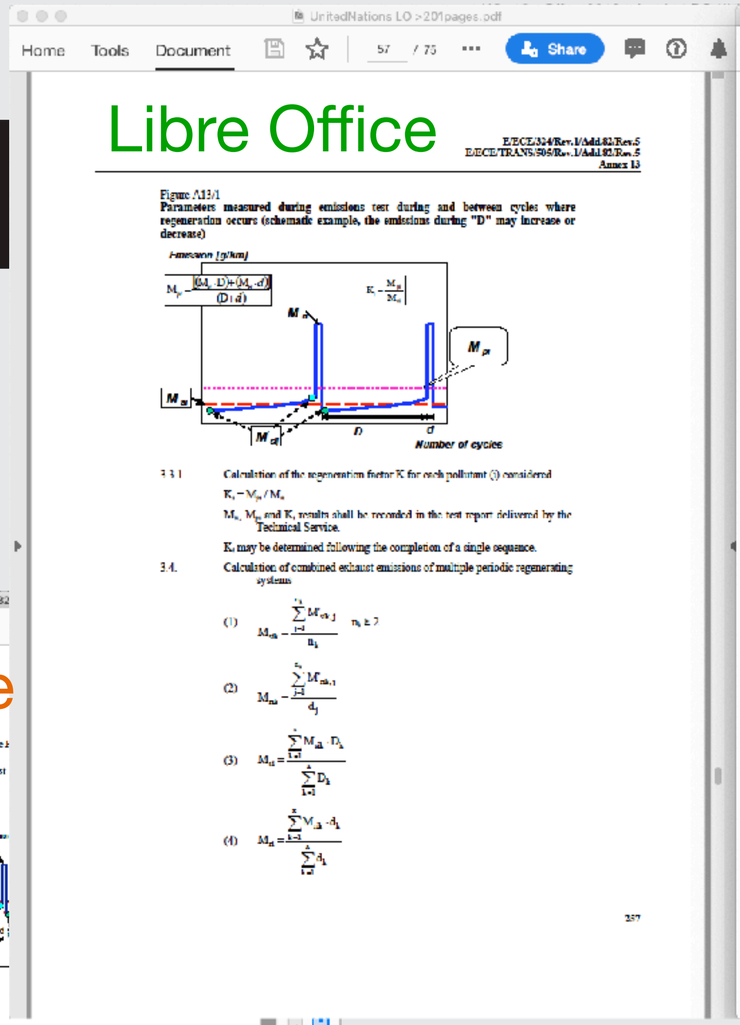
3.1.2.1. The procedure shall start with the discharge of the electrical energy/power storage device of the vehicle while driving (on the test track, on a chassis dynamometer, etc.):

(a) At a steady speed of 50 km/h until the fuel consuming engine of the HEV starts up;

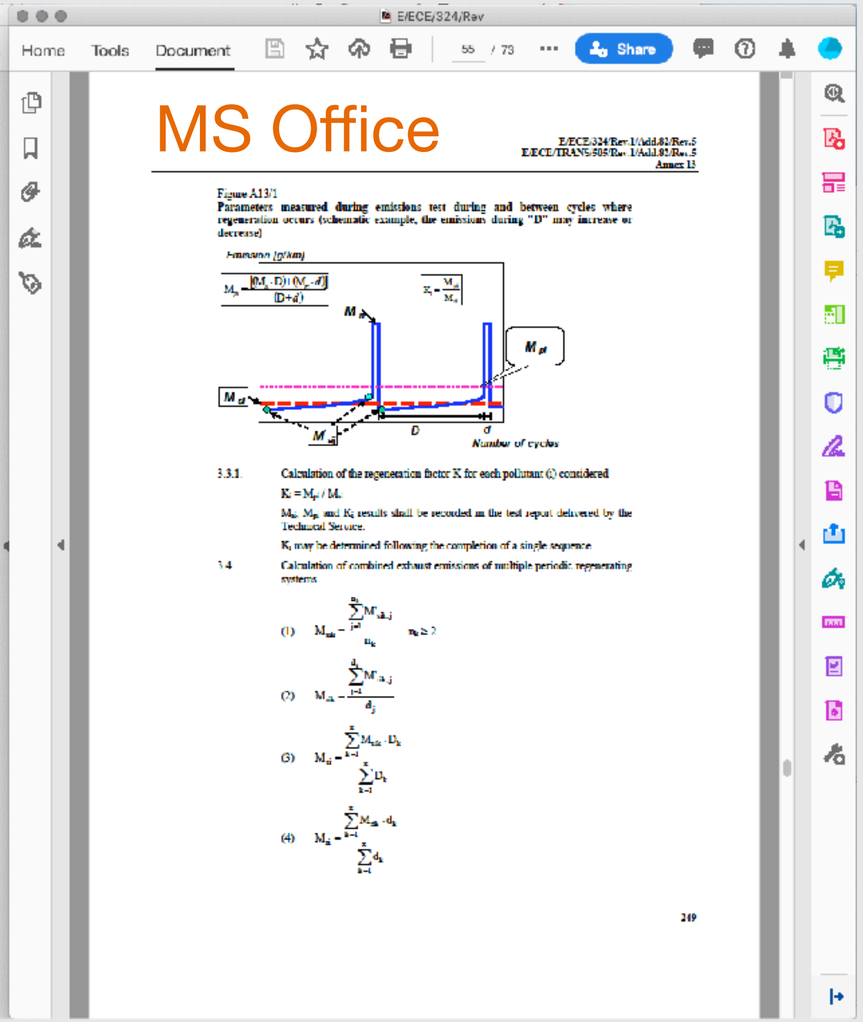
Vector objects

Identical (Vector graphics)

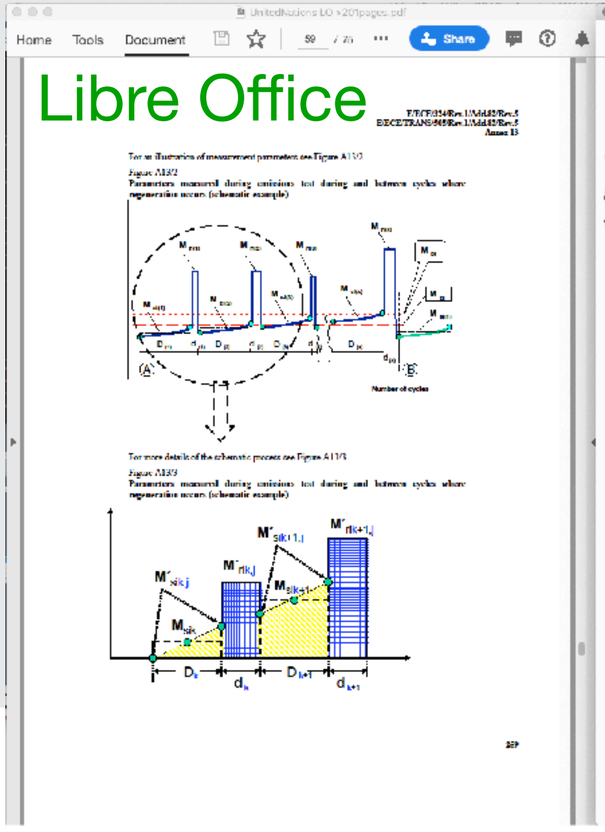
Libre Office



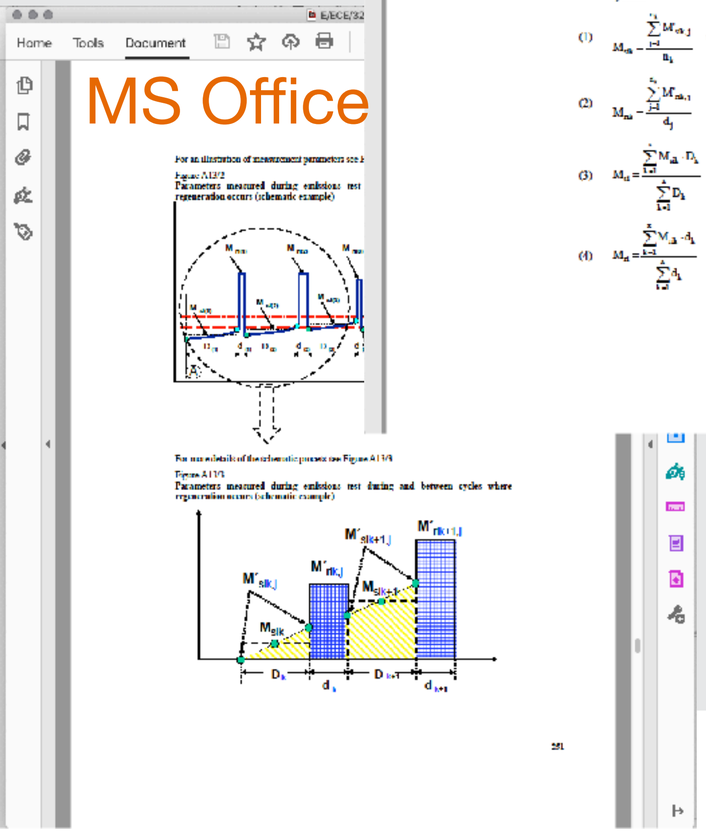
MS Office



Libre Office



MS Office

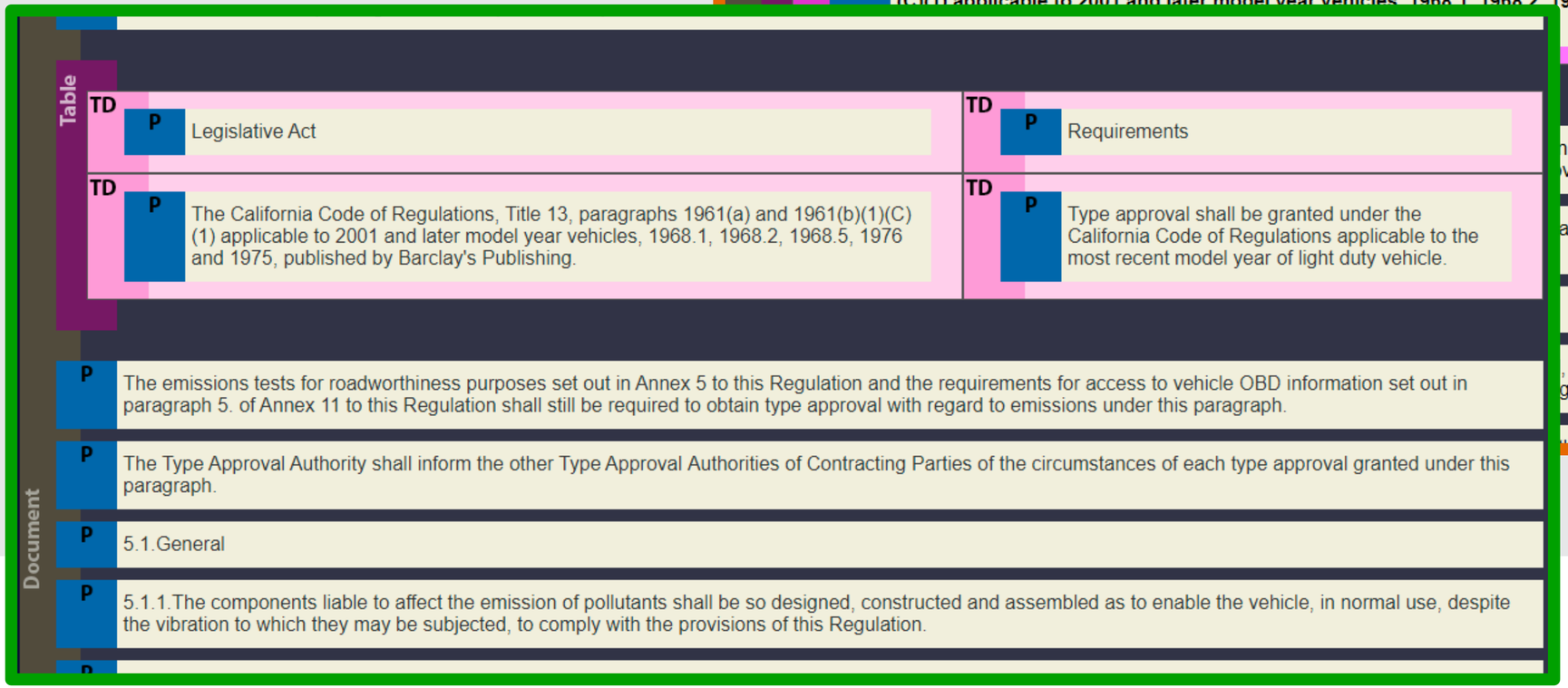


Tagging

Works - if prepared as needed

MS Office

Libre Office

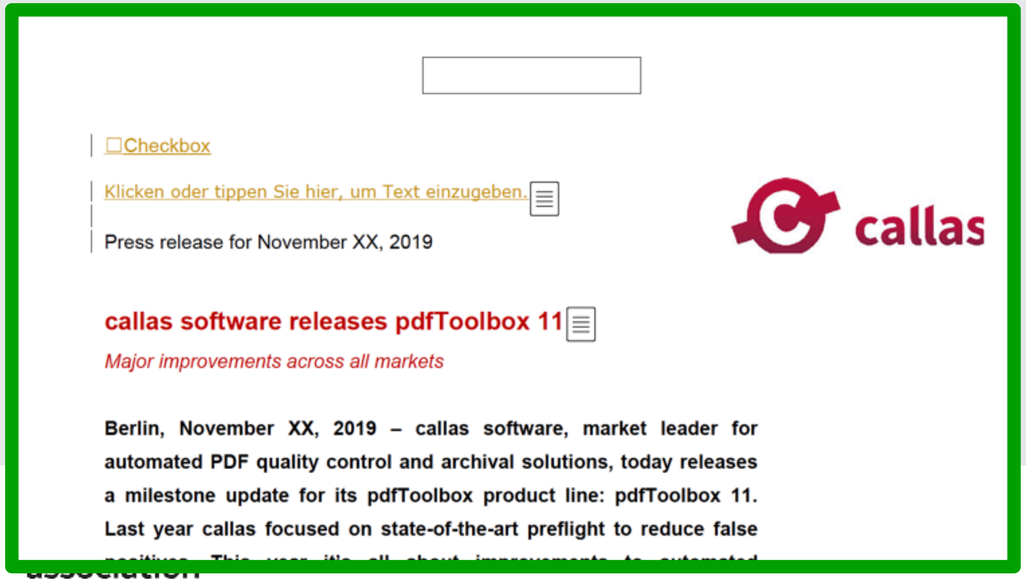


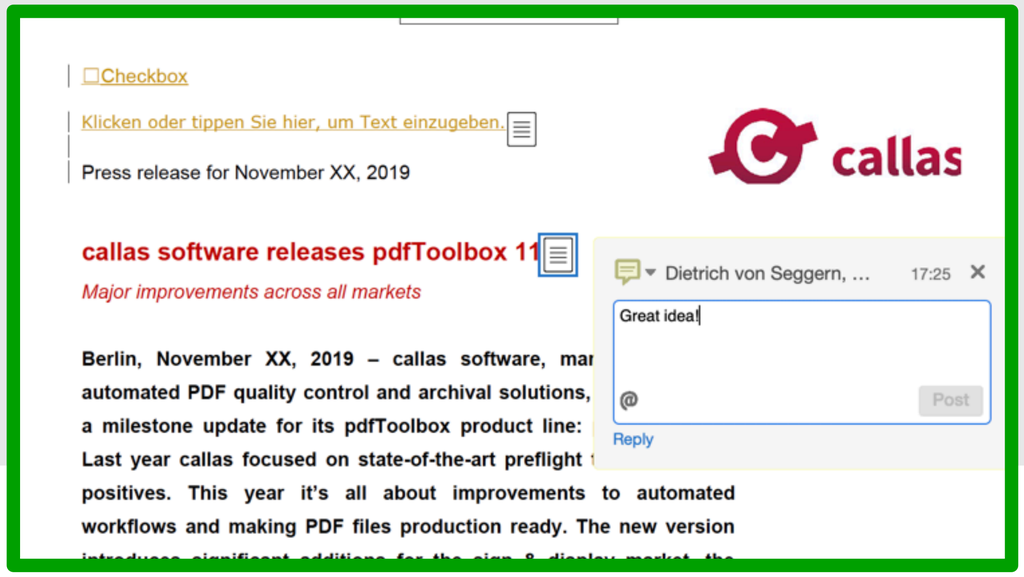
Annotations

MS Office

- MS Word: Flattened
- LibreOffice: Optional as PDF annotations

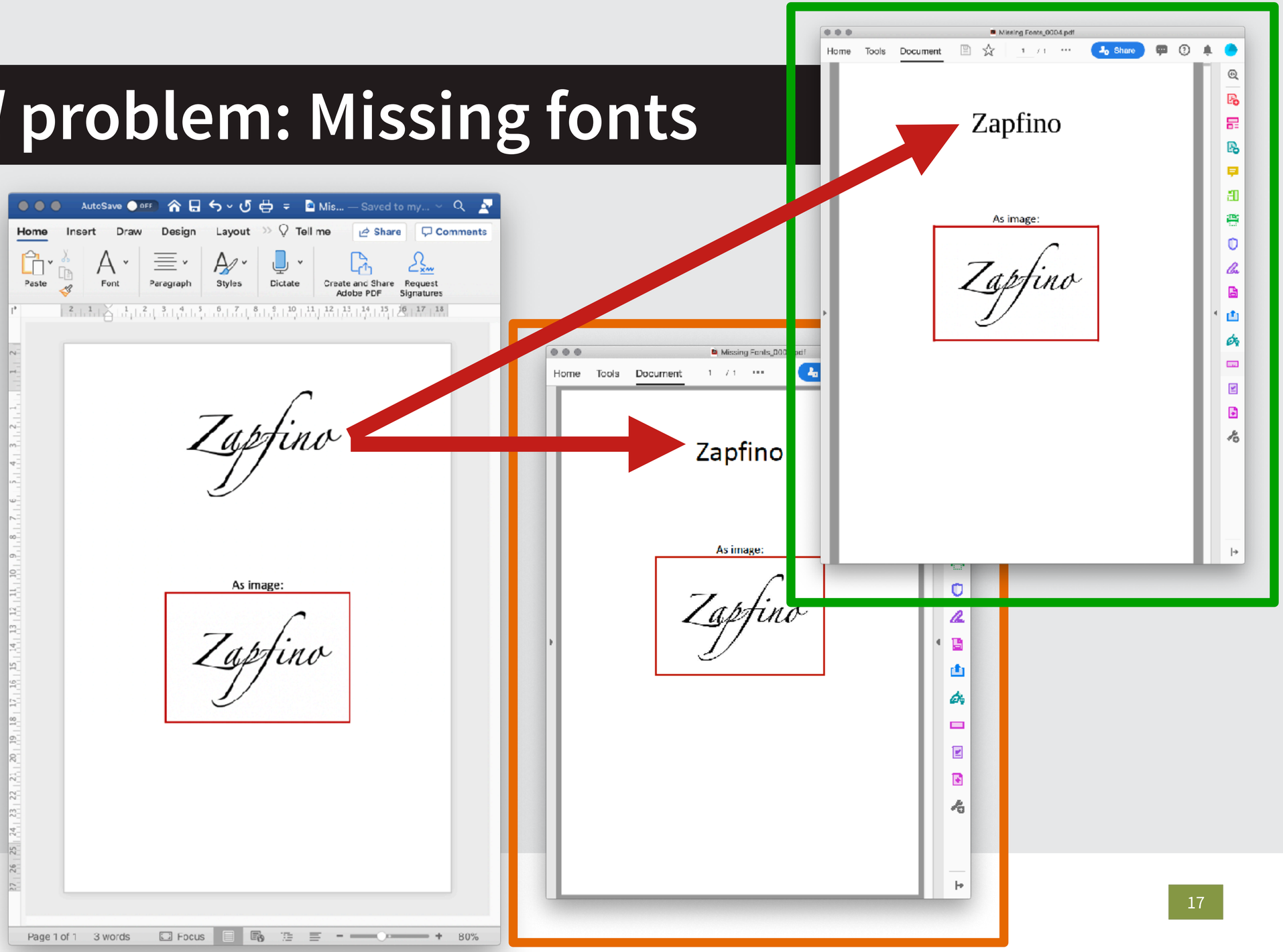
Libre Office





General problem: Missing fonts

No warnings (without special developments)

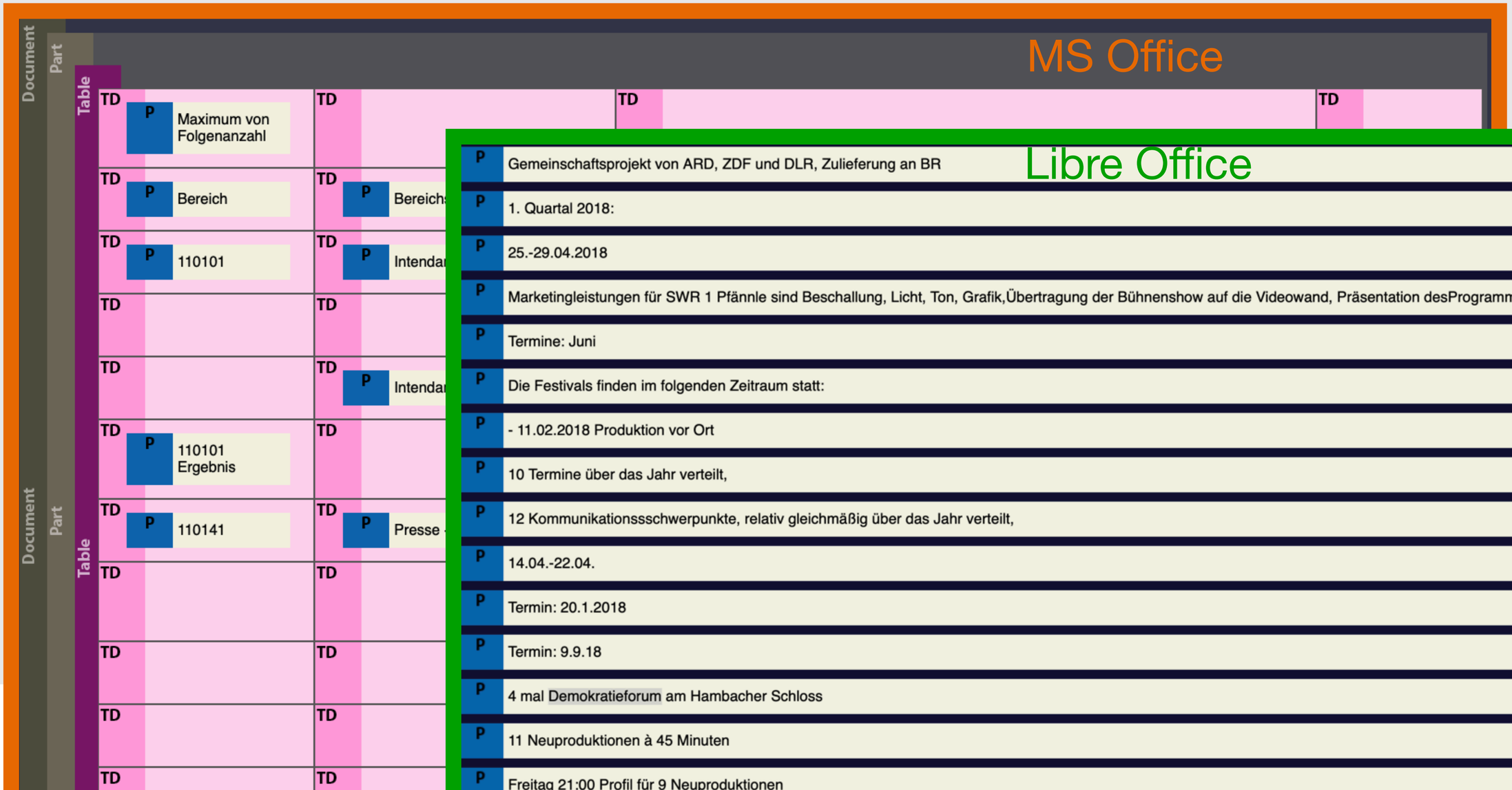


Spreadsheets: No pages, formulas

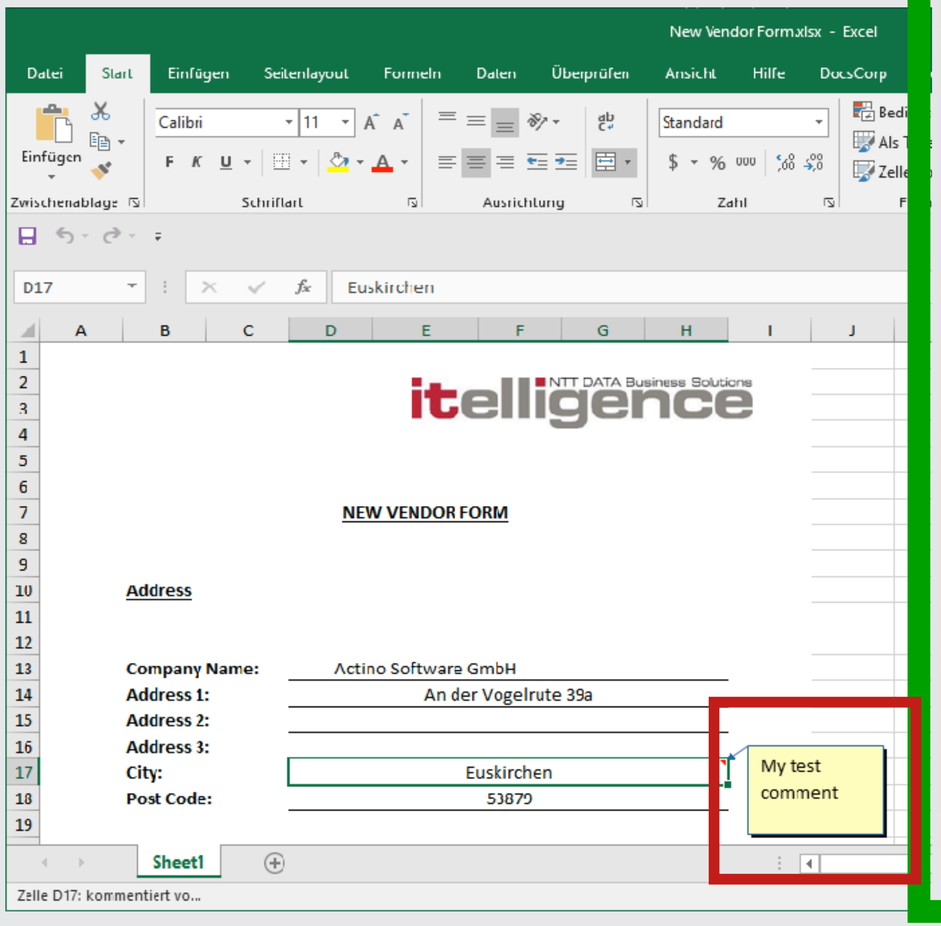
No real page concept

Not a document similar to Word / Writer and PDF

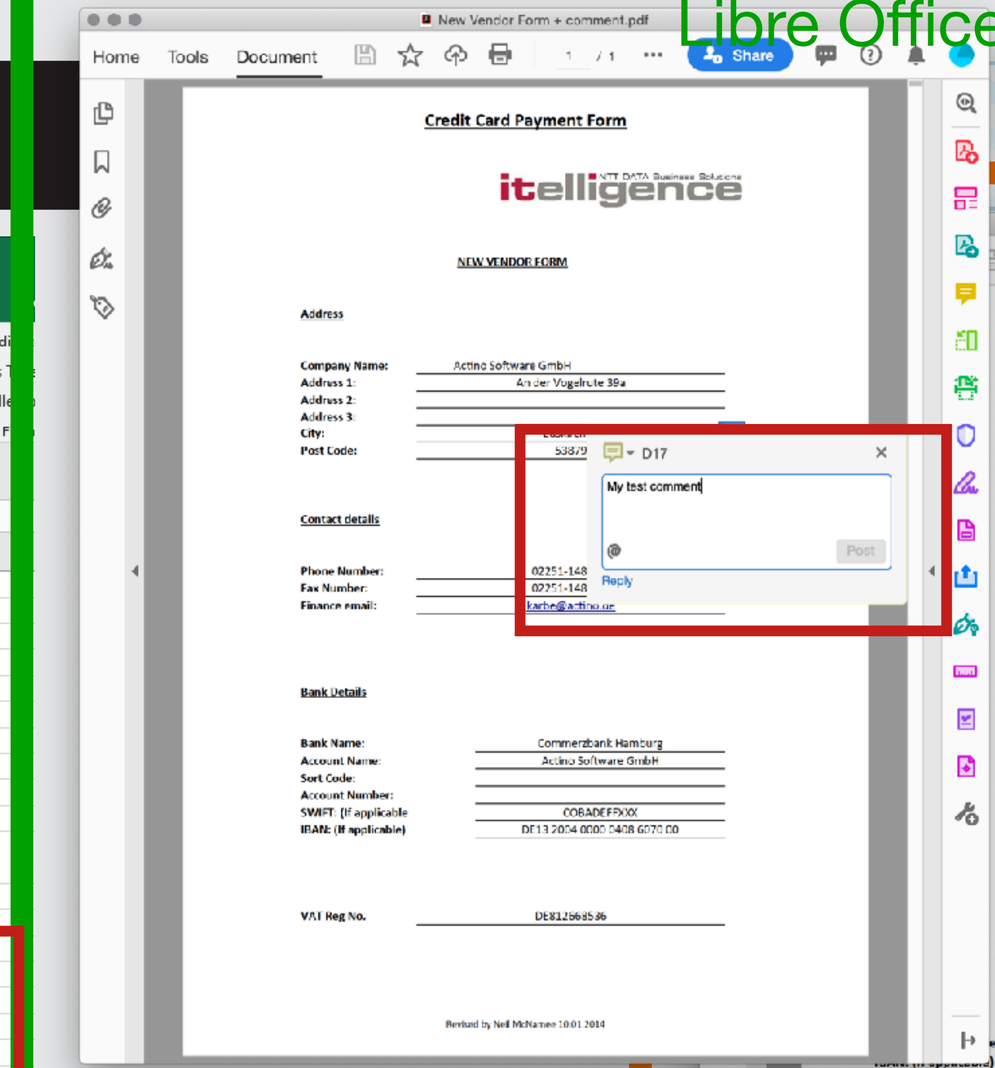
Tagging to a certain degree



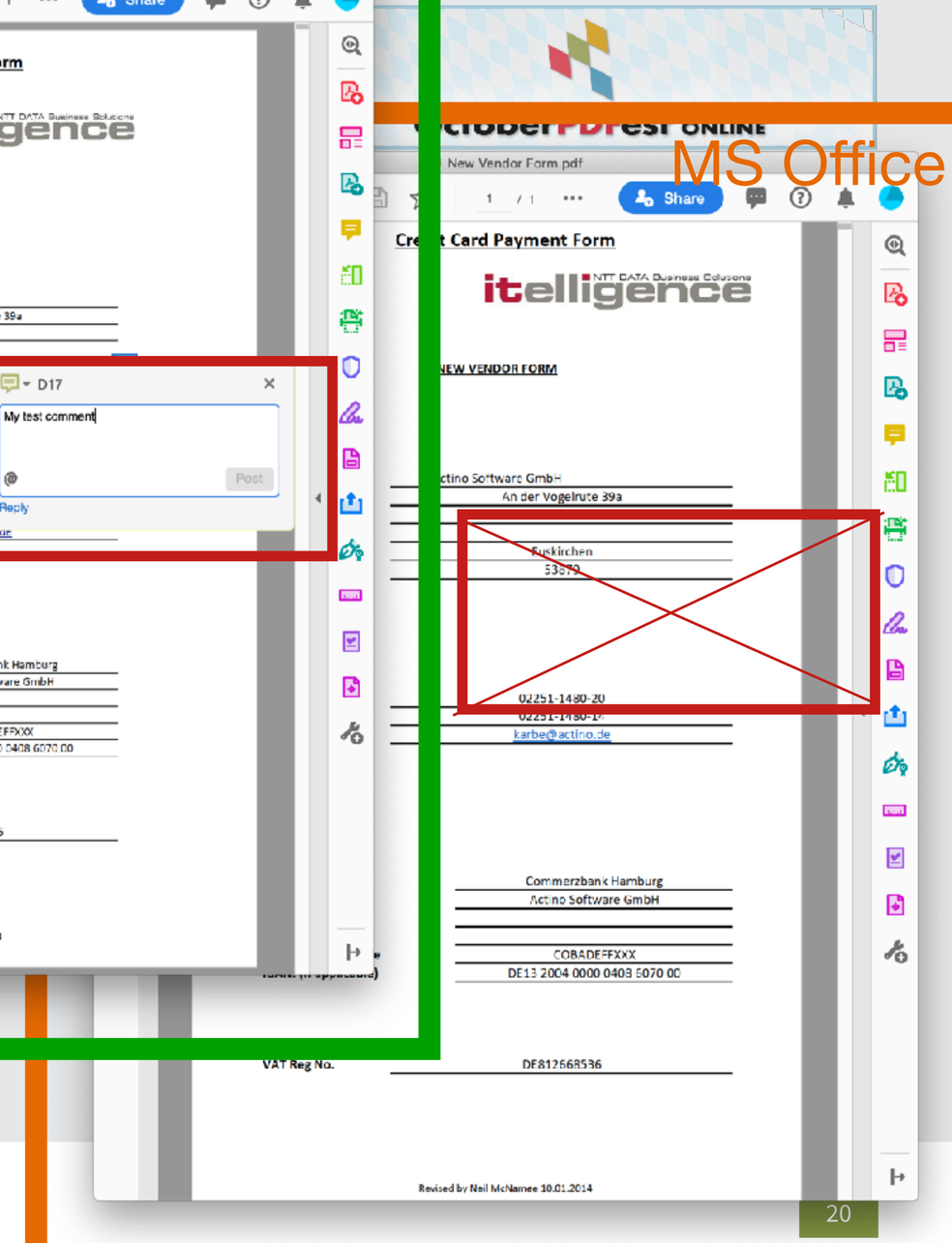
Annotations



Libre Office

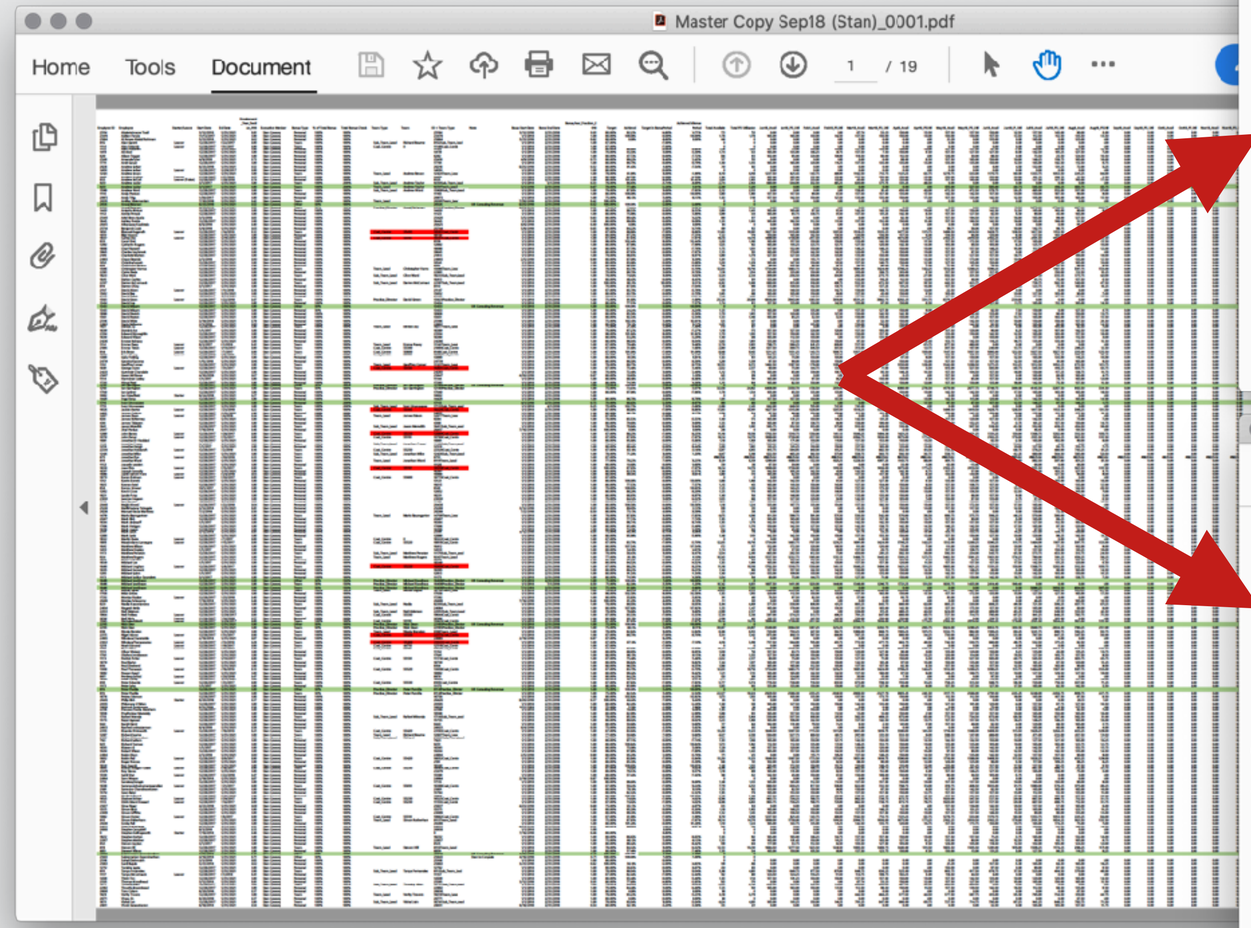


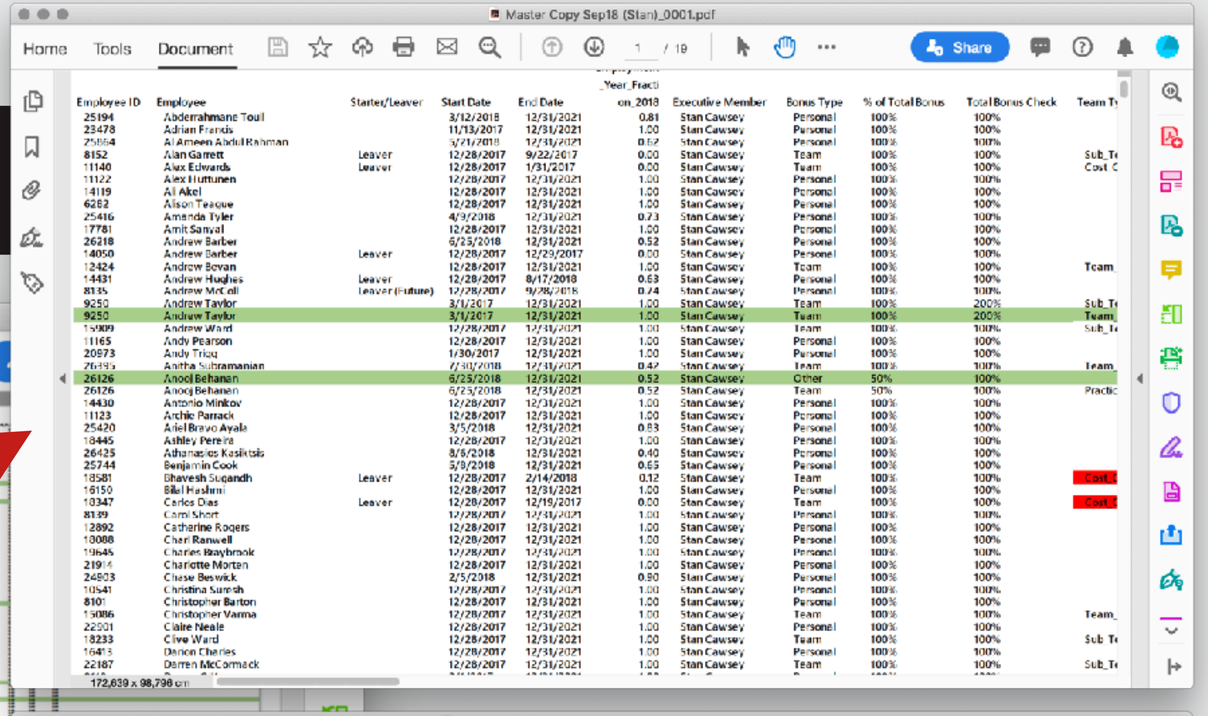
MS Office

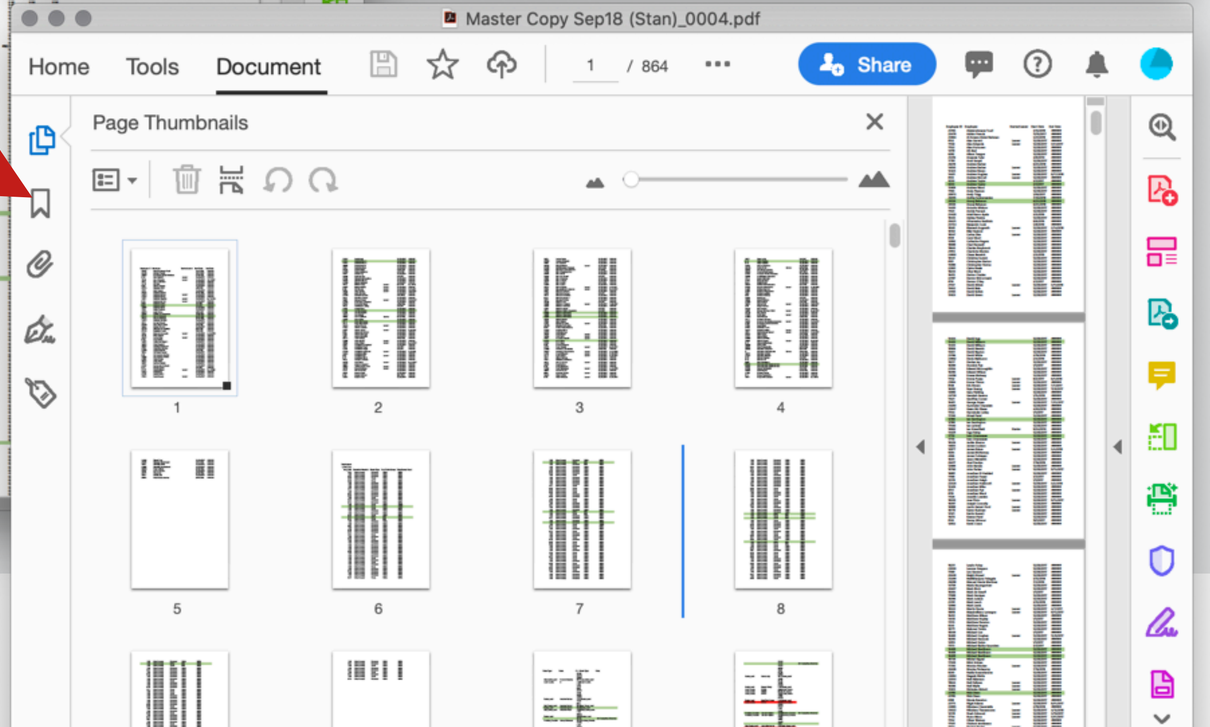


PDF association

General: Page breaks





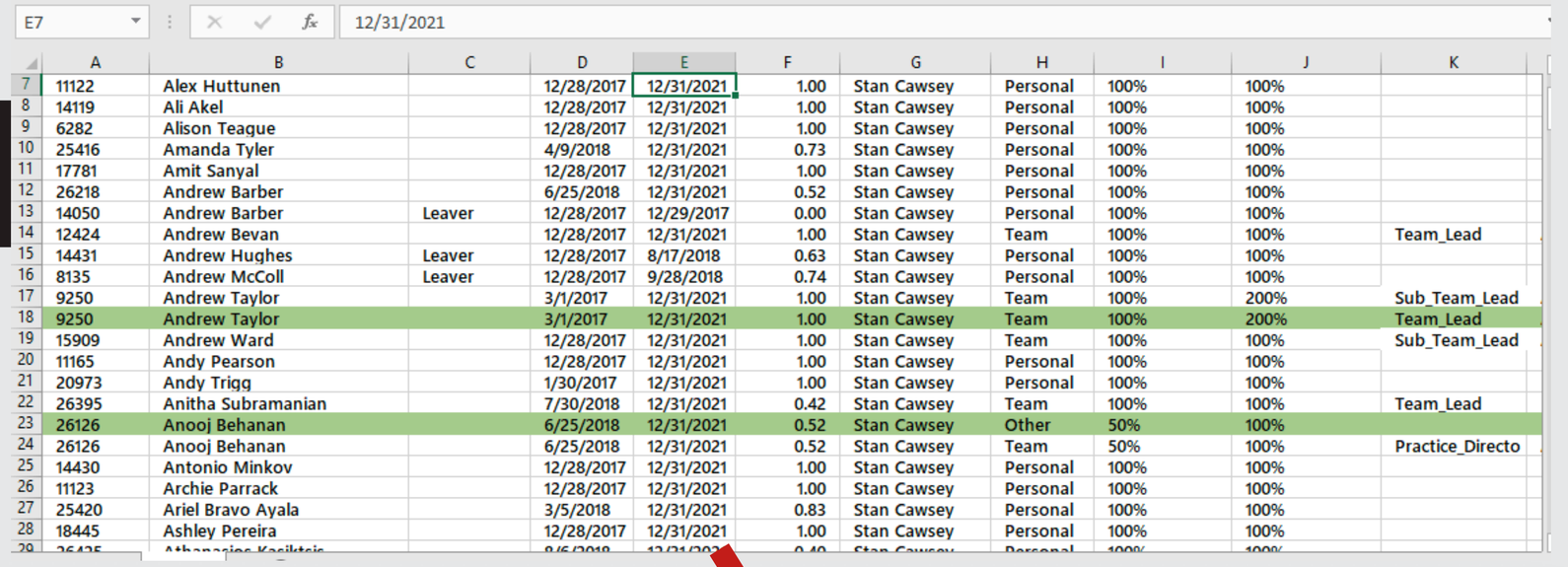


PDF association

General: hashs

Slight changes (e.g. a different font used for PDF conversion)

may result in slightly longer texts which will result in hashes “#######”



E7 | 12/31/2021

	A	B	C	D	E	F	G	H	I	J	K
7	11122	Alex Huttunen		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
8	14119	Ali Akel		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
9	6282	Alison Teague		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
10	25416	Amanda Tyler		4/9/2018	12/31/2021	0.73	Stan Cawsey	Personal	100%	100%	
11	17781	Amit Sanyal		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
12	26218	Andrew Barber		6/25/2018	12/31/2021	0.52	Stan Cawsey	Personal	100%	100%	
13	14050	Andrew Barber	Leaver	12/28/2017	12/29/2017	0.00	Stan Cawsey	Personal	100%	100%	
14	12424	Andrew Bevan		12/28/2017	12/31/2021	1.00	Stan Cawsey	Team	100%	100%	Team_Lead
15	14431	Andrew Hughes	Leaver	12/28/2017	8/17/2018	0.63	Stan Cawsey	Personal	100%	100%	
16	8135	Andrew McColl	Leaver	12/28/2017	9/28/2018	0.74	Stan Cawsey	Personal	100%	100%	
17	9250	Andrew Taylor		3/1/2017	12/31/2021	1.00	Stan Cawsey	Team	100%	200%	Sub_Team_Lead
18	9250	Andrew Taylor		3/1/2017	12/31/2021	1.00	Stan Cawsey	Team	100%	200%	Team_Lead
19	15909	Andrew Ward		12/28/2017	12/31/2021	1.00	Stan Cawsey	Team	100%	100%	Sub_Team_Lead
20	11165	Andy Pearson		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
21	20973	Andy Trigg		1/30/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
22	26395	Anitha Subramanian		7/30/2018	12/31/2021	0.42	Stan Cawsey	Team	100%	100%	Team_Lead
23	26126	Anooj Behanan		6/25/2018	12/31/2021	0.52	Stan Cawsey	Other	50%	100%	
24	26126	Anooj Behanan		6/25/2018	12/31/2021	0.52	Stan Cawsey	Team	50%	100%	Practice_Directo
25	14430	Antonio Minkov		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
26	11123	Archie Parrack		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	
27	25420	Ariel Bravo Ayala		3/5/2018	12/31/2021	0.83	Stan Cawsey	Personal	100%	100%	
28	18445	Ashley Pereira		12/28/2017	12/31/2021	1.00	Stan Cawsey	Personal	100%	100%	

Master Copy Sep (Stan)_0003.pdf

Home Tools Document / 19 Share

Employee ID	Employee	Starter/Leaver	Start Date	End Date	Employment Year_Fracti on_2018	Executive Member	Bonus Type	% o
25194	Abderrahmane Touil		3/12/2018	#######	0.81	Stan Cawsey	Personal	10
23478	Adrian Francis		11/13/2017	#######	1.00	Stan Cawsey	Personal	10
25864	Al Ameen Abdul Rahman		5/21/2018	#######	0.62	Stan Cawsey	Personal	10
8152	Alan Garrett	Leaver	12/28/2017	9/22/2017	0.00	Stan Cawsey	Team	10
11140	Alex Edwards	Leaver	12/28/2017	1/31/2017	0.00	Stan Cawsey	Team	10
11122	Alex Huttunen		12/28/2017	#######	1.00	Stan Cawsey	Personal	10
14119	Ali Akel		12/28/2017	#######	1.00	Stan Cawsey	Personal	10
6282	Alison Teague		12/28/2017	#######	1.00	Stan Cawsey	Personal	10
25416	Amanda Tyler		4/9/2018	#######	0.73	Stan Cawsey	Personal	10
17781	Amit Sanyal		12/28/2017	#######	1.00	Stan Cawsey	Personal	10
26218	Andrew Barber		6/25/2018	#######	0.52	Stan Cawsey	Personal	10
14050	Andrew Barber	Leaver	12/28/2017	#######	0.00	Stan Cawsey	Personal	10
12424	Andrew Bevan		12/28/2017	#######	1.00	Stan Cawsey	Team	10
14431	Andrew Hughes	Leaver	12/28/2017	8/17/2018	0.63	Stan Cawsey	Personal	10
8135	Andrew McColl	Leaver (Future)	12/28/2017	9/28/2018	0.74	Stan Cawsey	Personal	10
9250	Andrew Taylor		3/1/2017	#######	1.00	Stan Cawsey	Team	10
9250	Andrew Taylor		3/1/2017	#######	1.00	Stan Cawsey	Team	10
15909	Andrew Ward		12/28/2017	#######	1.00	Stan Cawsey	Team	10
11165	Andy Pearson		12/28/2017	#######	1.00	Stan Cawsey	Personal	10
20973	Andy Trigg		1/30/2017	#######	1.00	Stan Cawsey	Personal	10
26395	Anitha Subramanian		7/30/2018	#######	0.42	Stan Cawsey	Team	10
26126	Anooj Behanan		6/25/2018	#######	0.52	Stan Cawsey	Other	50
26126	Anooj Behanan		6/25/2018	#######	0.52	Stan Cawsey	Team	50
14430	Antonio Minkov		12/28/2017	#######	1.00	Stan Cawsey	Personal	10
11123	Archie Parrack		12/28/2017	#######	1.00	Stan Cawsey	Personal	10
25420	Ariel Bravo Ayala		3/5/2018	#######	0.83	Stan Cawsey	Personal	10

PDF association

Slides: Close to PDF

- Has pages
- Mostly absolute positioning of objects

Power Point / Impress

- Not supported (in MS Office or LibreOffice)
  - Tagging
  - Slide transitions (to PDF transitions)
  - Metadata
  - Layout templates
    - Presenter notes
    - Handout (several slides)

S	name	*(Optional)* The *transition style* that shall be used when moving to this page from another during a presentation. Default value: *R*. *Split* — Two lines sweep across the screen, revealing the new page. The lines may be either horizontal or vertical and may move inward from the edges of the page or outward from the centre, as specified by the **Dm** and **M** entries, respectively. *Blinds* — Multiple lines, evenly spaced across the screen, synchronously sweep in the same direction to reveal the new page. The lines may be either horizontal or vertical, as specified by the **Dm** entry. Horizontal lines move downward; vertical lines move to the right. *Box* — A rectangular box sweeps inward from the edges of the page or outward from the centre, as specified by the **M** entry, revealing the new page. *Wipe* — A single line sweeps across the screen from one edge to the other in the direction specified by the **Di** entry, revealing the new page. *Dissolve* — The old page dissolves gradually to reveal the new one. *Glitter* — Similar to *Dissolve*, except that the effect sweeps across the page in a wide band moving from one side of the screen to the other in the direction specified by the **Di** entry. *R* — The new page simply replaces the old one with no special transition effect; the **D** entry shall be ignored. *Fly* — *(PDF 1.5)* Changes are flown out or in (as specified by **M**), in the direction specified by **Di**, to or from a location that is offscreen except when **Di** is *None*. *Push* — *(PDF 1.5)* The old page slides off the screen while the new page slides in, pushing the old page out in the direction specified by **Di**. *Cover* — *(PDF 1.5)* The new page slides on to the screen in the direction specified by **Di**, covering the old page. *Uncover* — *(PDF 1.5)* The old page slides off the screen in the direction specified by **Di**, uncovering the new page in the direction specified by **Di**. *Fade* — *(PDF 1.5)* The new page gradually becomes visible through the old one.

Email conversion

- No support in Libre Office at all
- Microsoft Outlook very difficult to automate

➡ Preferably direct file format conversion

Email: A complex topic of its own

- Header: into XMP metadata of the PDF
- Body: the richest variant (HTML, RTF, Text) into PDF pages
  - What to do with external references?
- Attachments: As PDF attachments either converted to PDF if possible or not
  - Special problem: “digital only (XML) files”
    (calendar attachments, vCard attachments)

Summary: Office to PDF

- **Text processors: Word, Writer**
  - No content missing
  - Differences in positioning/reflow
- **Spreadsheet: Excel, Calc**
  - Results good in both cases
  - General pagination problems
- **Slides: PowerPoint, Impress**
  - Results good in both cases
- **Email**
  - A complex topic of its own
  - Direct file format conversion

Automate office files to PDF

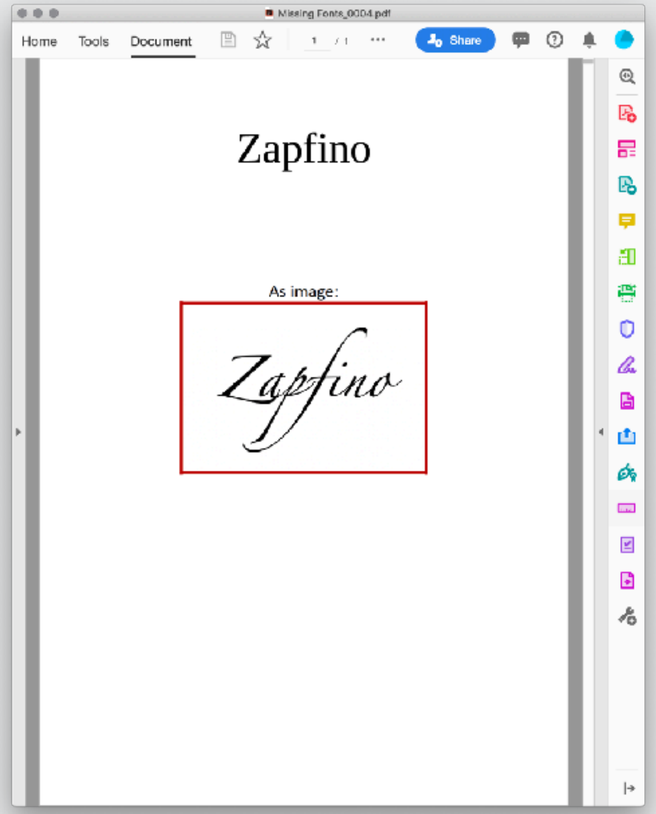
Microsoft, LibreOffice or other?

Office to PDF with callas pdfaPilot

- Select office application (MS Office or LibreOffice / OpenOffice) for conversion
- Create an as rich as possible intermediate result and post process result PDF
  - Support for PDF/A-4 right away
- Optionally embed original as PDF attachment

Special params: Text processors

- Log font substitutions
- Control whether or not to update fields
- Control quality (image resolution)
- Control comments

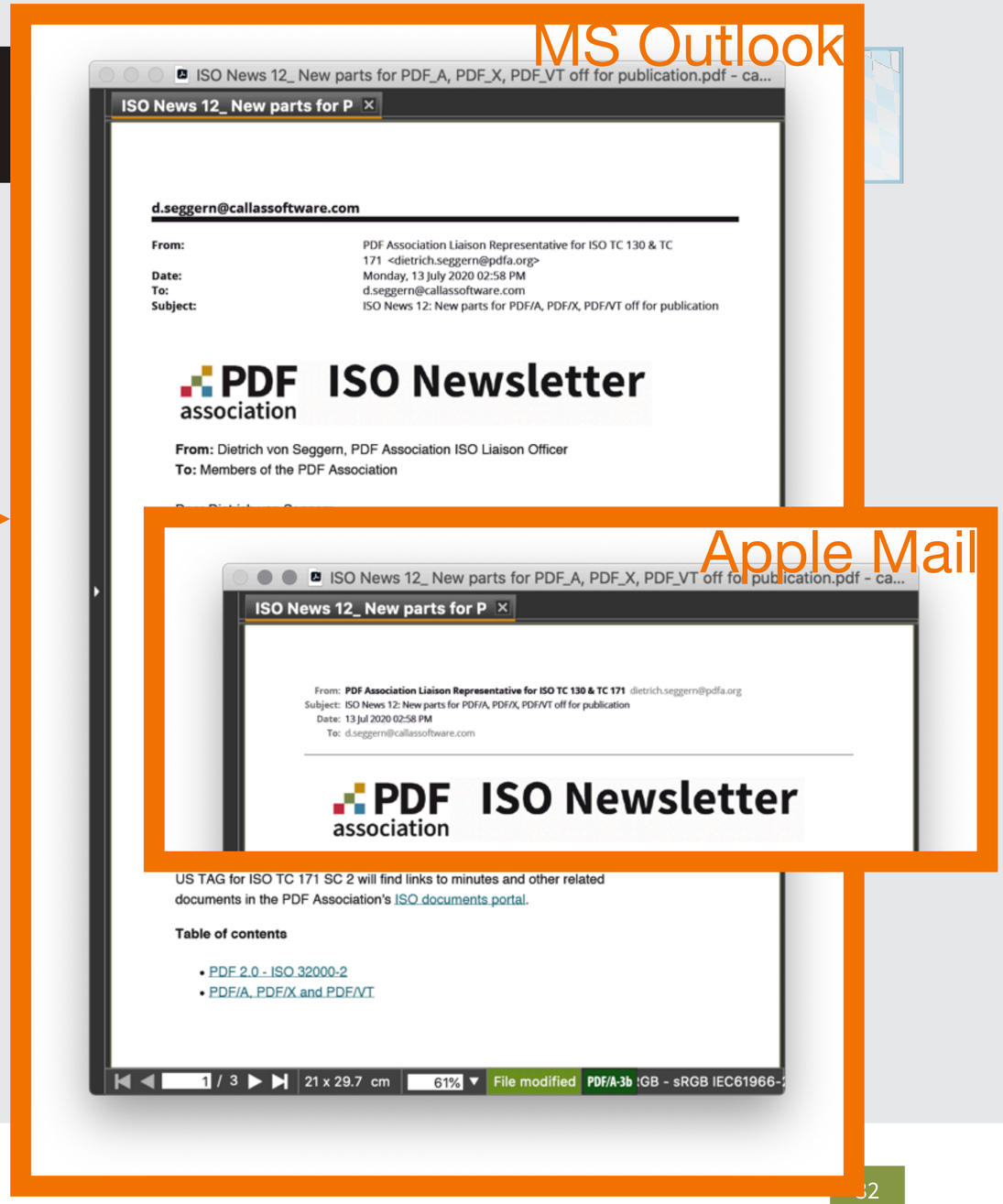


Special parameters: Spreadsheets

- One page per table in a spreadsheet: all filled cells on a single page or “use Excel page layout”
- Include hidden columns
- Shrink content in cells to fit cell dimensions (avoid #######)

Special parameters: Email

- Email parameters
  - Control remote content (for HTML emails)
  - Templates for page layout
  - Embed attachments
    - Original format
    - Converted to PDF
    - Both
  - Add all header fields to PDF XMP metadata



Automate office files to PDF

Microsoft, LibreOffice or other?

General requirements

- Filesize of results
- Quality of PDF code
- Embedded originals
- Unicode
- Conversion / Automation
  - Performance
  - Platforms
  - Unattended
  - Parallel processing

for PDF/A or other requirements it is often better to do additional PDF work, e.g. font embedding or transparency flattening

--topdf_parameter Custom parameters, valid values are:
ShowHiddenColumns (for MS Excel conversion only)
ShrinkToFit (for MS Excel conversion only)
PrintQualityAndComments (for MS Word conversion only)
UpdateChangedFields (for MS Word conversion only)
DoNotHideOffice (for MS Office only)
NoMemoryOptimization (for MS Office only)
CSV_IMPORT=FieldSeparator,TextDelimiter,CharacterSet (for LibreOffice only)
FieldSeparator: Set field seperator as ASCII value, e.g. 44 for comma, 59 for semi-colon, 44/59 for both
TextDelimiter: Set TextDelimeter as ASCII value, e.g. 34 for double quotes and 39 for single quotes
CharacterSet: Set character set. 0 = Unknown/System, 1 = Windows-1252, 9 = System Default, 11 = Ascii, 76 = UTF-8
NoBitmapMissingFonts (for MS Word conversion only)

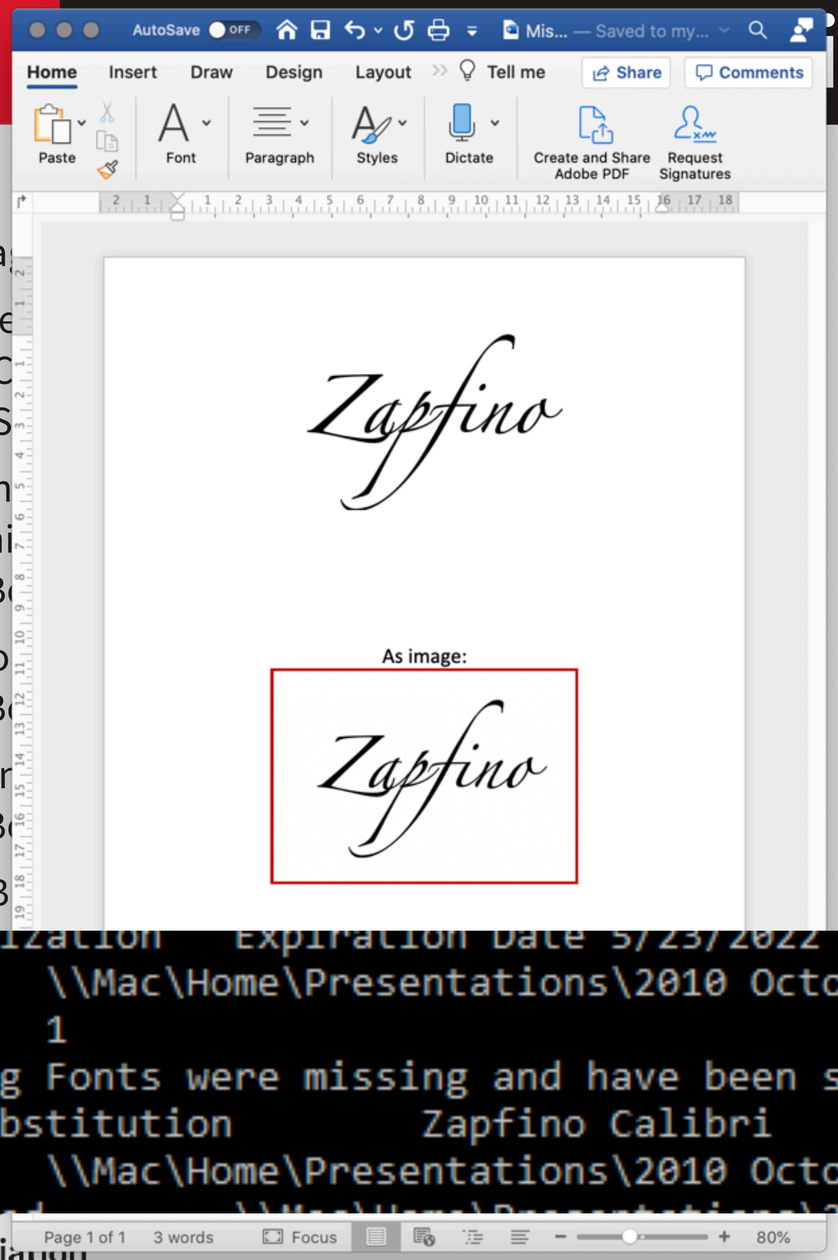
Missing Fonts… are logged as return value so that an application can act on that information

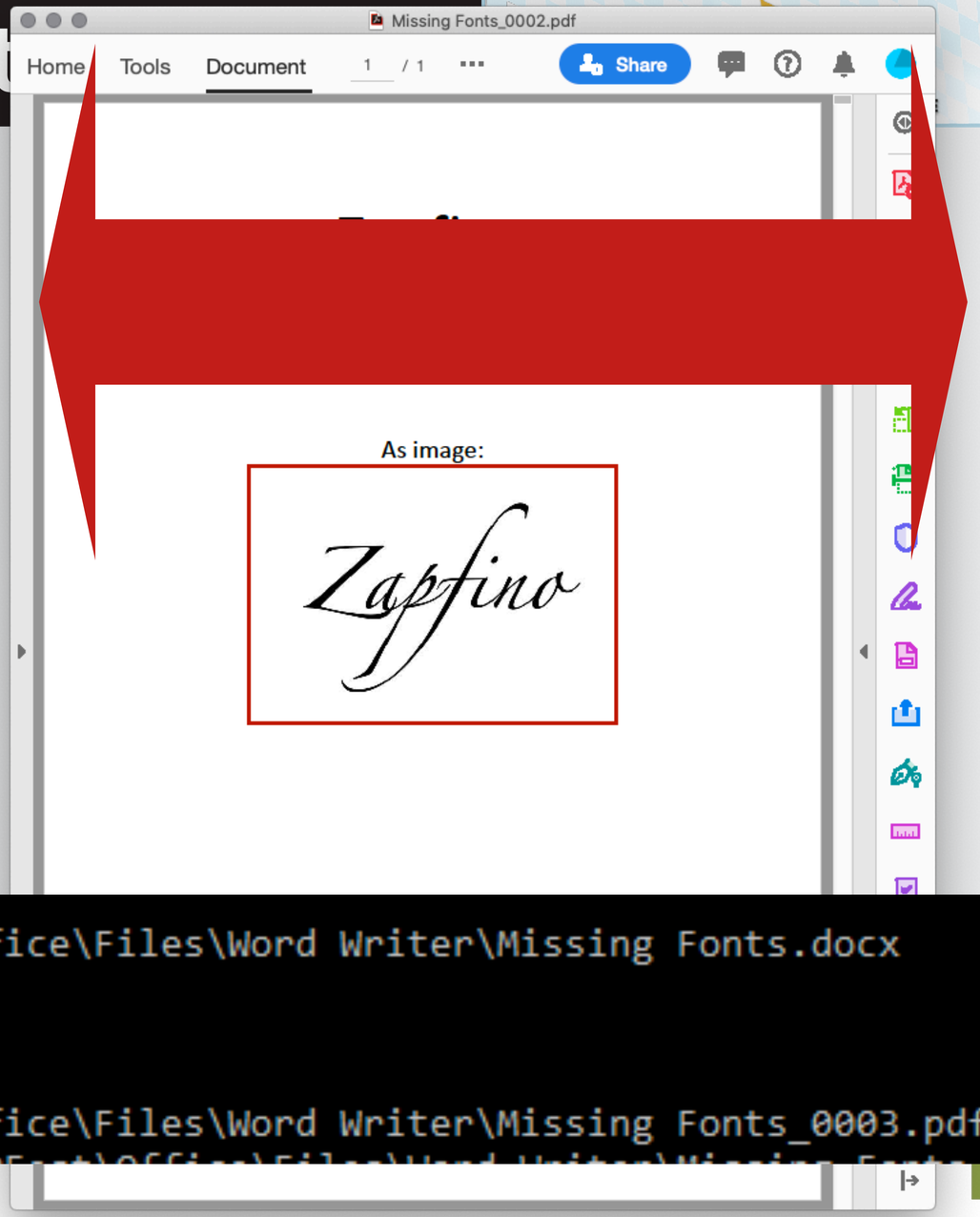
Some tests
(not representative)

- Filesize of results / Quality of PDF code
  - No differences found?
- Embedded originals
  - LibreOffice or PDF side
- Unicode
  - In general no problem, what about bullet points from Libre Office?
- Conversion / Automation
  - Performance
    - No differences found?
  - Platforms
    - LibreOffice and other converters are available on Linux, results are - other than for MS Office - the same on MacOS
  - Unattended
    - Manageable (dialogue management, restart the app once in a while)
    - Without user context: Requires preparation
  - Parallel processing
    - .NET easier

irement

- Ta
- Me / C / S
- An /ni /B
- Fo /B
- Lir /B
- “B






```
Serialization   Expiration Date 5/23/2022
Input   \\Mac\Home\Presentations\2010 OctoberPDFest\Office\Files\Word Writer\Missing Fonts.docx
Pages   1
Warning Fonts were missing and have been substituted
FontSubstitution        Zapfino Calibri
Output  \\Mac\Home\Presentations\2010 OctoberPDFest\Office\Files\Word Writer\Missing Fonts_0003.pdf
```

Some tests with .doc / .docx (not representative)

Some tests (not representative)

MS Word file
=> MS Word
=> LibreOffice

Problems
Longer text
Tables
Complex constructions

Conceptual requirements Word

Metadata
Annotations
Forms
Links

“Bugs”
Missing fonts
- Fonts need to be present in the system for office files
- If not: layout changes without notification
- In PDF: Embedded

Page breaks

Conceptual requirements Excel

Tagging (no)

Metadata

Annotations

Pagination

“Bugs”

Cells that are too small

MS Excel file

=> MS Word

=> LibreOffice

Problems are the same with both Apps?

Control of pagination
with print output

Cells that are too small

(Foxit: LibreOffice is better for CSV input)

Conceptual requirements Excel

- Tagging (no)
- Metadata
- Annotations
- Pagination
- “Bugs”
  - Cells that are too small

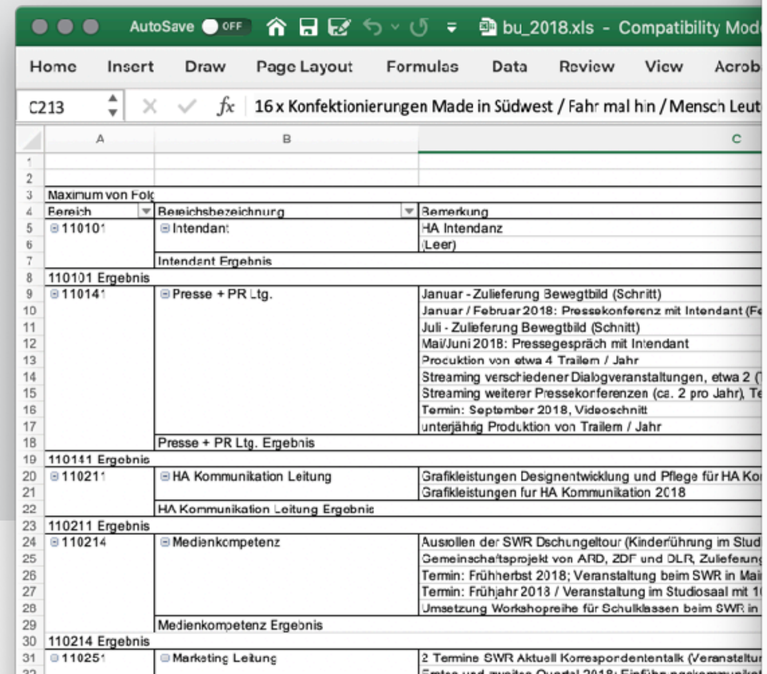
--topdf_parameter Custom parameters, va
ShowHiddenColumns (for MS Excel conversion only)
ShrinkToFit (for MS Excel conversion only)
CSV_IMPORT=FieldSeparator,TextDelimiter,CharacterSet (for LibreOffice only)
FieldSeparator: Set field separator as ASCII value, e.g. 44 for comma, 59 for semi-colon, 44/59 for both
TextDelimiter: Set TextDelimeter as ASCII value, e.g. 34 for double quotes and 39 for single quotes
CharacterSet: Set character set. 0 = Unknown/System, 1 = Windows-1252, 9 = System Default, 11 = Ascii,

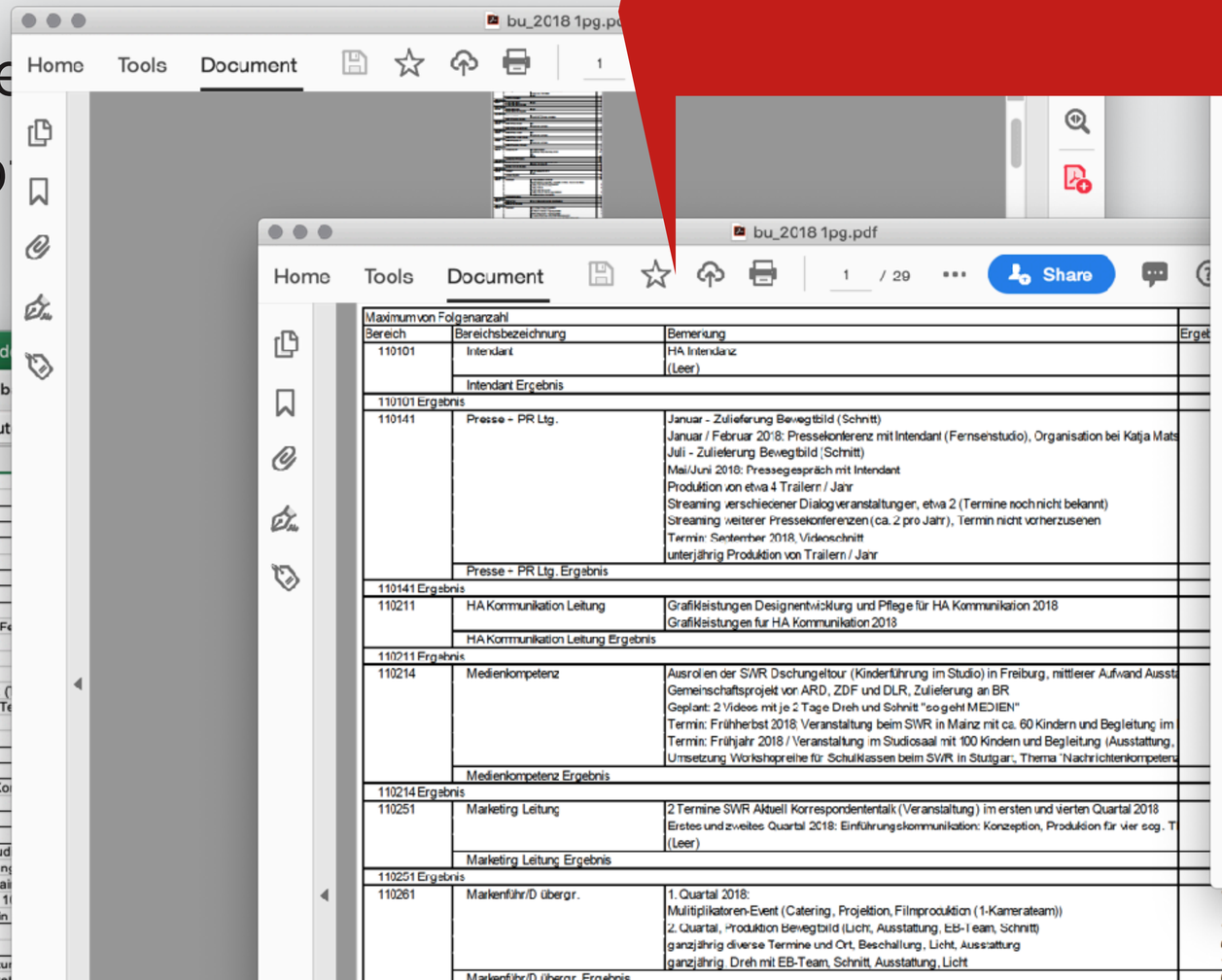
Missing Fonts… are logged as return value so that an application can act on that information

Spreadsheets: Excel

something complete

no real page concep





Conversion via applications

Conceptual requirements Email

Header information (Metadata)
Referenced images
Attachments
Tagging (no)

“Bugs”
Pagination
winmail.dat (TNEF)

Some tests (not representative)

Email file
=> MS Word (Outlook)
=> ~~LibreOffice~~

Outlook can’t be as easily accessed as the other MS Office applications
More or less impossible
Only native converters
Compared to the other formats no as strict design requirements
Biggest problems are attachments